

Department of Epidemiology and Biostatistics
School of Public Health
Faculty of Medicine
Imperial College London

Characterising urban environments in Sub-Saharan Africa with satellite imagery and unsupervised deep learning

Antje Barbara Metzler

Thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy (PhD), August 2023

Declaration of originality and copyright

Declaration of originality

I hereby declare that this thesis is my own original research and work. Any other work that is not my own, I have cited appropriately.

Some work of this thesis has been published in the peer-reviewed publication titled "Phenotyping urban built and natural environments with high-resolution satellite images and unsupervised deep learning" in the Science of the Total Environment (STOTEN) journal, 2023 [1]. DOI: <https://doi.org/10.1016/j.scitotenv.2023.164794>

Copyright

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Abstract

Sub-Saharan Africa is the fastest urbanising region globally, with cities experiencing rapid changes in infrastructure such as roads, buildings, vegetation and land use. Monitoring these developments is crucial for informing measures that make urban development sustainable and benefit city residents. Using high-resolution satellite imagery and unsupervised machine learning, this thesis presents a new framework to characterise urban environments in Accra and three other Sub-Saharan cities.

In Accra, the method identified distinct and interpretable urban phenotypes. Clusters captured either single urban features, such as water bodies, or intricate combinations like buildings surrounded by vegetation. With its nuanced analysis of Accra, including sensitivity analyses on scale and cluster number, this approach set a strong foundation for expanding to other cities in Sub-Saharan Africa.

In the second part, the method was extended to multiple cities, using a two-step feature learning and clustering approach. It was applied to four Sub-Saharan cities: Accra, Dakar, Dar es Salaam and Kigali. The framework identified shared urban phenotypes across all cities, including densely populated areas and informal settlements both within and outside the city core. It also captured city-specific features such as patterns of urban vegetation and urban agriculture in Dar es Salaam and Kigali.

This work revealed common growth patterns across African cities. Specifically, Accra, Dakar, and Dar es Salaam share similar urban features, largely due to their colonial histories, even though they differ in climate and historical backgrounds. Kigali differs in its lack of a strong colonial footprint, yet shares other features, such as mixed urban-rural landscapes. This thesis demonstrates that high-resolution satellite images, coupled with unsupervised deep learning provide an interpretable and scalable solution for tracking urban development in near real-time, particularly in low-income countries where traditional data are scarce.

Ghana by Air

Let the dry savannah winds
grab you by the hair
and drag you
stumbling
through tough clumps
of cow-nibbled grass.

Did you see the children streak by
on bikes - two or sometimes three on the same?
Did you see the striking garments
energised by night dances?

Now the green of the forest
will sweep you across the secret
passages of
Ashanti pride
into the rich red
depths of palm nut soup.

Did you see a woman drag her husband
home by the cuff - insulting him all the way?
Did you hear the market traders
hurling prices down the street?

The lights of Accra
will confound you as you circle
dreamily round
haphazard
houses arranged to fit
sprouting family trees.

Did you hear the blaring
horns - sometimes angry, sometimes friendly?
Did you see two lovers
heading for Labadi beach?

Follow them
lift the hem
of the sea
you have reached
the centre
of the world

Nii Ayikwei Parkes, 2020

Contents

Abstract	3
1 Introduction	23
1.1 Research aims and objectives	24
1.1.1 Objectives	25
1.2 Structure of thesis	25
2 Background	27
2.1 Overview	27
2.2 Urbanisation in Sub-Saharan Africa	27
2.2.1 Opportunities of growing cities in Sub-Saharan Africa	29
2.2.2 Challenges of growing cities in Sub-Saharan Africa	30
2.3 Post-colonial urban inheritance and policy in Sub-Saharan Africa	31
2.3.1 Colonial city planning in the study cities	33
2.3.2 Urban development trends and policies in the study cities	34
2.4 Data gaps in Sub-Saharan Africa	35
2.5 Studying cities	36

2.5.1	Complexity of cities	36
2.6	Using images for studying cities	38
2.6.1	Street-level images	38
2.6.2	Satellite images	39
2.6.3	Computer vision for satellite images	40
2.6.4	Deep learning for satellite images	41
2.7	Unsupervised learning	42
2.7.1	Clustering	43
2.7.2	Unsupervised image analysis for studying cities	44
2.7.3	Unsupervised machine learning for satellite images	44
2.7.4	Unsupervised deep learning for studying cities with satellite images	45
2.8	Summary	46
3	Study areas, data sources and data lessons	47
3.1	Overview and objectives	47
3.2	Study areas	48
3.2.1	Selection of cities used for analysis	48
3.3	City profiles	48
3.3.1	Accra, Ghana	48
3.3.2	Dakar, Senegal	50
3.3.3	Dar es Salaam, Tanzania	51
3.3.4	Kigali, Rwanda	52

3.4	City boundaries	52
3.5	Satellite images	53
3.5.1	Maxar satellite images	54
3.6	Data for cluster interpretation	59
3.6.1	Overview	59
3.6.2	Data on the built and natural environment	60
3.6.3	Summary	74
4	Clustering one city: Accra	76
4.1	Overview and objectives	76
4.2	Methods	77
4.2.1	Data sources and pre-processing	77
4.2.2	Analysis framework	80
4.2.3	DeepCluster: Combined feature extraction and clustering	81
4.2.4	Cluster interpretation	85
4.2.5	External cluster analysis	85
4.2.6	Internal evaluation of clusters	89
4.3	Results	91
4.3.1	External cluster analysis	91
4.3.2	Internal evaluation of clusters	98
4.4	Discussion	102
4.4.1	Implications for tracking and enhancing sustainable urban develop- ment	103

4.5	Summary	106
5	Sensitivity analyses	107
5.1	Overview and objectives	107
5.2	Methods	108
5.2.1	Data sources and pre-processing	108
5.2.2	Methods of sensitivity analyses	109
5.3	Results	112
5.3.1	Sensitivity to image tile dimensions	112
5.3.2	Sensitivity to zoning	114
5.3.3	Sensitivity to number of clusters	114
5.4	Discussion	122
5.4.1	Finding a suitable scale of analysis	123
5.4.2	Finding a suitable K	125
5.5	Summary	126
6	Clustering multiple cities	127
6.1	Overview and objectives	127
6.2	Methods	128
6.2.1	Data sources and pre-processing	128
6.2.2	Inference on Accra model	129
6.2.3	Model considerations for the multi-city analysis	130
6.2.4	Cluster evaluation	133

6.2.5	Number of clusters	135
6.3	Results	137
6.3.1	Cluster results for joint, separate and hybrid models	137
6.4	Discussion	155
6.5	Summary	157
7	Phenotyping natural and built environments in cities in SSA	162
7.1	Overview and objectives	162
7.2	Methods	163
7.2.1	Data and pre-processing	163
7.2.2	Hybrid two-stage approach	163
7.2.3	External analysis of clusters	164
7.2.4	Internal analysis of clusters	165
7.2.5	Sensitivity to number of clusters (K)	166
7.3	Results	167
7.3.1	Clusters of the urban environment	167
7.3.2	Cluster interpretation with feature prediction	174
7.3.3	Feature space analysis	177
7.4	Discussion	185
7.5	Summary	189

8 Discussion and conclusion	190
8.1 Summary	190
8.2 Strengths and limitations	191
8.2.1 Strengths and novelty	191
8.2.2 Limitations	192
8.3 Impact and implications for tracking urban development and policy	193
8.3.1 Leveraging unsupervised machine learning and satellite images	194
8.3.2 Implications for urban policy and planning	195
8.3.3 Urban phenotypes in SSA and their implications	196
8.4 Future work	199
8.5 Conclusion	200
8.6 Open science and code availability	201
Bibliography	201
A	228
A.1 Maxar Open Data	228
A.2 Maxar internal use license	228
A.3 Satellite image datasets	236
A.4 XGBoost hyperparameters, Accra model	236
A.5 Hyperparameter experiments for multi-city clustering	237
A.5.1 Approach A	237
A.5.2 Approach B	239

A.5.3	Approach C	241
A.5.4	Cluster names, Approach A with K=16.	244
A.6	XGBoost hyperparameters, combined approach	245
A.7	ESA land cover co-occurrence	246
A.8	Sensitivity to cluster number, multi-city analysis	251
A.9	Inter- and intra-cluster distances	253

List of Tables

3.1	List of boundary shapefiles used for defining the study cities.	53
3.2	Satellite image basemaps purchased from Maxar.	57
3.3	Data on population, built and natural environment. NDVI: Normalised Difference vegetation index	61
6.1	Satellite images used for the cluster analysis of four cities.	128
6.2	Table of clusters across cities with K=8, Approach A.	138
6.3	Table of clusters across cities, Approach B.	143
6.4	Table of clusters across cities, Approach C.	145
6.5	Comparing the overall MAD scores per city. Bold text indicates the best scores for each city. DeS: Dar es Salaam. Approach A refers to the K=8 model.	147
6.6	Selected learning rates across all approaches.	154
7.1	Accuracy scores of the classifiers trained on urban characteristics to predict specific cluster membership for each city.	165
A.1	List of available satellite basemap imagery by Maxar Open Data [2]	228

A.2	Best hyperparameters for XGBoost model predicting cluster membership in Accra.	236
A.3	Approach A, across all cities. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics. Analyses with a * return inherent image stitching of the Accra image that is part of the commercial pre-processing. Running the model with K=16 and LR of 0.1 resulted in an error.	238
A.4	Approach B, Dakar. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.	239
A.5	Approach B, Des. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.	239
A.6	Approach B, Kigali. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.	240
A.7	Approach C, across all cities. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.	241
A.8	Approach C, Accra. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.	241
A.9	Approach C, Dakar. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.	242
A.10	Approach C, Dar es Salaam. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.	242
A.11	Approach C, Kigali. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.	243
A.12	Best hyperparameter for XGBoost model predicting cluster membership of four cities.	245

List of Figures

3.1	Satellite basemap of Accra, Ghana.	55
3.2	Satellite basemap of Dakar, Senegal.	56
3.3	Satellite basemap of Dar es Salaam, Tanzania.	56
3.4	Purchased satellite basemap of Dakar, captured in 2021/2.	58
3.5	Purchased satellite basemap of Dar es Salaam, captured in 2021.	59
3.6	Purchased satellite basemap of Kigali, captured in 2020	60
3.7	Building area and building count in all study cities.	63
3.8	Difference in building orientation in degrees (°) plotted on top of a satellite image of the neighbourhoods <i>Nima</i> and <i>Mamobi</i> in Accra, Ghana.	64
3.9	Average building size and building orientation in all study cities.	65
3.10	Length of roads and major roads.	67
3.11	Distance to all roads and major roads.	68
3.12	Population density and NDVI of all study cities.	71
3.13	Correlation of environmental and demographic variables of Accra.	72
3.14	Correlation of environmental and demographic variables of Dakar.	72
3.15	Correlation of environmental and demographic variables of Dar es Salaam.	73

3.16	Correlation of environmental and demographic variables of Kigali.	73
3.17	ESA land cover map of Accra, Dakar, Dar es Salaam and Kigali. 10m resolution.	75
4.1	256 x 256 grid tiles plotted on top of a satellite image segment.	79
4.2	Examples of tiles that were excluded from the dataset.	80
4.3	Overview of data management, analysis, and interpretation.	84
4.4	Screenshot of webpage for cluster interpretation input.	88
4.5	Map of image-driven clusters in the Greater Accra Metropolitan Area. . . .	93
4.6	Built and natural environment and demographic characteristics of clusters.	94
4.6	Built and natural environment and demographic characteristics of clusters.	95
4.7	Difference in building orientation captured by the unsupervised analysis. .	95
4.8	Built and natural environment and demographic variables as predictors of membership in image-based clusters.	97
4.9	Top 9 activated images and filter extracted from Conv5.	99
4.10	Class activation maps of DeepCluster.	100
4.11	Cluster assignment and image tiles seen in the CNN extracted feature space. Part A.	101
4.12	Cluster assignment and image tiles seen in the CNN extracted feature space. Part B.	102
5.1	Segment of the satellite basemap with a 128 x 128 pixel (green), 384 x 384 pixel (yellow) and 512 x 512 pixel (red) pixel grid indicating the tile sizing of the sensitivity analysis.	109
5.2	Cluster results of varying tile size plotted on a map of GAMA.	113

5.3	Impact of image tile size on urban cluster characteristics.	117
5.4	Influence of tile size on cluster membership.	119
5.5	Sensitivity to zoning.	119
5.6	Impact of cluster number on the clustering outcome.	119
5.6	Impact of feature extraction hyperparameter k on learning of deep features (DF): DF_{k8} , DF_{k50} and DF_{k100} (left to right).	121
5.7	Influence of feature extraction hyperparameter k on learning of deep fea- tures (DF): DF_{k8} , DF_{k50} and DF_{k100} (left to right).	121
5.8	Elbow plot and CH score from K=2 to K=15 of DF_{k8}	122
5.9	Elbow plot and CH score from K=2 to K=15 of DF_{k50}	122
5.10	Elbow plot and CH score from K=2 to K=15 of DF_{k100}	123
6.1	Transferring the Accra model to satellite images of Dakar.	130
6.2	Cluster approaches A, B, and C.	132
6.3	Approach A: Joint one-stage clustering with K=8.	139
6.4	Heatmap of MedAD scores per cluster and urban characteristic for Ap- proach A.	139
6.5	Approach A: Joint one-stage clustering with K=16.	140
6.6	Shared cluster across cities for Approach A with K=16.	141
6.7	Common clusters found across cities for Approach B. Each row represents one city; from top to bottom: Accra, Dakar, Dar es Salaam and Kigali. Dynamic mosaic ©2022 Maxar Technologies.	144
6.8	Examples of city-specific clusters for Approach B).	144

6.9	Common clusters found across cities for Approach C. Each row represents one city; from top to bottom: Accra, Dakar, Dar es Salaam and Kigali. Dynamic mosaic ©2022 Maxar Technologies.	146
6.10	Examples of city-specific clusters for Approach C).	147
6.11	Water and vegetation clusters for Approach B and Approach C.	150
6.12	Densely populated areas clusters for Approach B and Approach C.	151
6.13	Cluster change of Approach B (left) and Approach C (right), Accra.	158
6.14	Cluster change of Approach B (left) and Approach C (right), Dakar.	159
6.15	Cluster change of Approach B (left) and Approach C (right), Dar es Salaam.	160
6.16	Cluster change of Approach B (left) and Approach C (right), Kigali.	161
7.1	Two-stage multi-city clustering with K=8. Clusters plotted on the map of Accra.	168
7.2	Two-stage multi-city clustering with K=8. Clusters plotted on the map of Dakar.	169
7.3	Two-stage multi-city clustering with K=8. Clusters plotted on the map of Dar es Salaam.	169
7.4	Two-stage multi-city clustering with K=8. Clusters plotted on the map of Kigali.	170
7.5	Pie charts representing proportions of cluster membership per city.	171
7.6	Normalised co-occurrence plots (by rows) of Accra, Dakar, DeS and Kigali.	173
7.7	SHAP summary for Accra.	175
7.8	SHAP summary for Dakar.	176
7.9	SHAP summary for Dar es Salaam.	177

7.10	SHAP summary for Kigali.	178
7.11	Cluster centroids plotted on the first, second and third principal component of the feature space.	179
7.12	Clustermap of inter-cluster distances.	180
7.13	Clustermap of inter-cluster distances with z-score normalisation.	182
7.14	Silhouette and CH score of Accra and Dakar.	183
7.15	Silhouette and CH score of Dar es Salaam and Kigali.	184
A.1	List of labelled satellite image datasets considered for transfer learning. The table was created in 2020 and some datasets have been updated since.	236
A.2	ESA land cover classes, Kigali.	246
A.3	Two clusters of Kigali.	247
A.4	Dark dense vegetation cluster.	248
A.5	Bare areas.	249
A.6	Densely populated areas.	250
A.7	Clustergram of image representations of Accra.	251
A.8	Clustergram of image representations of Dakar.	251
A.9	Clustergram of image representations of Dar es Salaam.	252
A.10	Clustergram of image representations of Kigali.	252
A.11	Inter- and intra-distances of image representations with cluster assignment in the feature space.	253

Acronyms

AI Artificial Intelligence.

AMA Accra Metropolitan Area.

CAMs Class Activation Maps.

CH Calinski-Harabasz.

CNN Convolutional Neural Network.

CNNs Convolutional Neural Networks.

ESA European Space Agency.

GAMA Greater Accra Metropolitan Area.

GDAL Geospatial Data Abstraction Library.

GDP gross domestic product.

LiDAR Light Detection and Ranging.

MAD Median Absolute Deviation.

MAUP Modifiable Areal Unit Problem.

MedAD Median Absolute Deviation.

ML Machine Learning.

NDVI Normalised Difference Vegetation Index.

NMI Normalised Mutual Information.

NUP National Urban Policy.

OSM Open Street Map.

PC Principle Component.

PIC Power Iteration Clustering.

SAR Synthetic-Aperture Radar.

SHAP SHapley Additive exPlanations.

SSA Sub-Saharan Africa.

UN United Nations.

VHR very high resolution.

Acknowledgements

First and foremost, I would like to extend my thanks to my supervisors, Majid Ezzati, Viktoriia Sharmanska and Wenjia Bai. I am very grateful for your time, insight and intellectual support. Your collective wisdom has been a source of inspiration and has significantly shaped both this work and my broader research approach. Particular gratitude goes to Majid for your guidance and pushing me to never lose sight of the bigger picture (but at the same time schooling my eye for detail); to Viktoriia for your enthusiasm and new perspectives; and to Wenjia for your machine learning expertise and consistent support.

I am indebted to colleagues from both the *Pathways to Equitable Healthy Cities* consortium and the *Global Environmental Health* group. Your expertise and kindness greatly enriched my PhD experience. Special thanks go to my PhD peers: Emily Muller, Theo Rashid and Victor Lhoste. I am grateful I could always rely on your intellect and encouragement. A big thank you is attributed to Ricky Nathvani, your insights and feedback were invaluable. Thank you also to James Bennett, Vishwanath Doreswamy, Bin Zhou and Esra Suel for your help, especially relating to topics of statistics and machine learning. Sierra Clark and Honor Bixby, thank you for setting a formidable precedent, and Annie Sheehan and Aina Roca-Barcelo, for being such brilliant companions in the concluding year of my PhD. My appreciation extends to all my colleagues from *NCD-RisC*; our Wednesday meetings and lunch breaks were always a highlight. To my PhD colleagues from Ghana University, namely Charles Agyei-Asabere, Dina Adjei-Boadi, Elvis Kyere-Gyeabour and Jacob Doku Tetteh: thank you for your valuable input and for taking the time to repeatedly look at clusters with me. Thomas Wright and Dragana Vuckovic, our Zoom calls brought much-needed structure and connection in the depths of the pandemic. Many thanks to Abeer Arif and Giulia Mangiameli, your dedication to the *Pathways* project is unmatched, and your contributions are and were pivotal to the group.

I gratefully acknowledge the financial support from the Imperial College PhD president's scholarship scheme and the Wellcome Trust's Our Planet Our Health grant, which funded the larger *Pathways* project.

I owe so much to my family, who have supported this journey by filling me with love and fostering my curiosity since day one. *Liebsten Dank* to my mum, Antje, thank you for always believing in and trusting my choices. *Tausend Dank* to my dad, Hans-Peter, thank you for lighting my enthusiasm for computers and the sciences. *Allergrößten Dank* to my sisters, Laura and Marie-Christine. I am so grateful for your unconditional support and I am profoundly proud to be your sister. Thank you for always paving my way and continuously inspiring me.

Thank you to my London family, who gave me the warmth of a home away from home and always nourished my love for the sciences as well as the arts. I have so much gratitude for Eva Havelka and Martina Rimmele, whose support is all-encompassing. Million thanks to Martina, for your enthusiasm and understanding (and our countless coffee/tea breaks at White City campus); and to Evie, for your wisdom and constant encouragement. Many thanks also to Dorian Henning, Tillman Hönig, Hannah Backmeister, Theresa Eggler, Giulia Carlini, Karen Laanem and Vici Grabenwöger. Thank you for rooting for me.

And last but certainly not least, to George Ssali. Words might fall short, but know that my gratitude for our conversations, your unwavering support and sense of humour is boundless. This journey would have been significantly harder without you by my side.

Chapter 1

Introduction

In the past and present, African cities are estimated and projected to experience the highest rates of population growth globally. Over the last 50 years, the urban population in Sub-Saharan Africa (SSA) has increased almost tenfold, rising from 51 million in 1970 to more than 450 million today [3]. And this trend is anticipated to continue. According to United Nations (UN) projections, the population of African cities will grow by an additional 950 million people over the next three decades [3].

However, due to their rapid urban growth and limited enforcement of urban planning policies in SSA, cities in SSA are expanding and changing in an irregular nature [4]. This development has led to heterogeneous land use and housing patterns, creating pockets of informal settlements and varied neighbourhoods scattered across large city boundaries [4]. Despite challenges related to lacking infrastructure and provisions, cities also offer substantial opportunities for fighting poverty and propelling sustainable development [5]. The UN ‘World Cities Report 2020’ underscores the necessity of developing data-driven strategies to harness the potential of urbanisation and, in doing so, foster sustainable and resilient urban landscapes [3]. However, the lack of data, especially in SSA, restricts the implementation of these strategies.

Addressing these data gaps in SSA presents a significant challenge. Much of data-driven urban research focuses on the Global North, resulting in a lack of representation for the unique urban development realities in SSA [6]. While recent initiatives by private companies and public organisations have made strides in collecting datasets capturing urban features such as vegetation, roads, buildings, and population, comprehensive datasets are still lacking [7, 8, 9, 10]. In addition, data accessibility issues persist, and the available data often have varying spatial and temporal coverage and resolutions [11, 12, 13].

High-resolution satellite images combined with deep learning have emerged as a powerful tool to fill these data gaps and provide detailed insights into the urban form, including the built and natural environment [14, 15, 16]. Among these, unsupervised machine learning algorithms, particularly clustering, can identify patterns in large datasets without significant prior knowledge, assumptions or labels [17, 18]. It is a data-driven process analogous to categorising a collection of objects based solely on observed attributes [17, 19]. In urban and environmental studies, unsupervised learning can aid in unravelling the complexities of cities and provide valuable insights on trends in urban development [20]. With the advancements of deep learning, more nuanced and detailed studies of urban environments have become possible [21]. However, the application of unsupervised deep learning to remote sensing images in urban environment analysis remains understudied [22].

1.1 Research aims and objectives

This PhD research is housed within the *Pathways to Equitable Healthy Cities* Study, which is a global partnership of leading academic and policy institutions from six countries with the main goal of improving population health, enhancing health equity, and ensuring environmental sustainability in cities.

The overall aim of my PhD thesis was to measure and analyse urban development at different scales in Sub-Saharan Africa (SSA) with satellite images. The objectives of my

PhD research were as follows:

1.1.1 Objectives

1. To apply unsupervised clustering to visual features derived from high-resolution satellite images of Accra, Ghana.
2. To evaluate how the scale of analysis changes the clusters of the unsupervised deep learning model when applied to urban environments in Accra, Ghana.
3. To interpret the clusters in relation to external demographic and environmental features.
4. To extend the clustering framework to additional cities in SSA and to identify clusters, and their extent and arrangement, across cities that are similar and clusters that are city-specific.

1.2 Structure of thesis

In Chapter 2, I give an overview of the current literature on urbanisation with a focus on cities in SSA. Additionally, I summarise advances in Machine Learning (ML), especially unsupervised ML applied to satellite images for studying cities. In Chapter 3, I introduce the different data sources and formats that were used in my work, and state the intricacies of high-resolution satellite images. In Chapter 4, I describe a combined feature extraction and deep learning clustering framework and apply it to satellite images of Accra, Ghana. Chapter 5 includes the sensitivity analyses of the framework, which investigates various model choices such as the scale of analysis and the number of clusters. In Chapter 6, I describe how the deep learning clustering framework can be effectively scaled and applied to multiple cities in SSA. In Chapter 7, I apply a combined feature extraction and clustering approach to four cities in SSA and describe the resultant urban clusters. Lastly,

Chapter 8 provides a comprehensive discussion of the findings, drawing final conclusions and elucidating the broader implications stemming from the research.

Chapter 2

Background

2.1 Overview

In this chapter, I summarise the literature on the intersection of urban development, computer vision, unsupervised machine learning and remote sensing. The chapter opens with a summary of urbanisation trends and urban policy in Sub-Saharan Africa (SSA), with a focus on the cities studied in this thesis. Secondly, I review the literature on how complex urban environments are studied and how remote sensing data can be used to aid the analysis. I further describe how computer vision and deep learning algorithms have been applied to satellite images to study urban environments. Finally, I outline the literature on unsupervised machine learning and clustering methods and explain how they are used to study urban environments together with images.

2.2 Urbanisation in Sub-Saharan Africa

For most of human history, people lived in rural areas. Over the last century and especially the recent decades, this has shifted dramatically. In 2007, for the first time in

history, more people lived in cities compared to their rural counterparts. Since then, the urban population has been increasing steadily. Today, about 4.3 billion people live in cities, which represents about 55% of the world's population [3]. In Western Europe, the Americas, Australia, Japan, and the Middle East, over 80% of the population resides in urbanised regions. In low-income countries, especially SSA, there are still more people living in rural areas compared to cities. However, this percentage is changing rapidly through migration to urban areas and changing population structures [4]. The United Nations (UN) forecasts that by 2030, the urban population of Africa will comprise approximately 50% of its total population [3]. Within this century, numerous African cities will evolve into some of the world's largest metropolises. By 2100, Africa is expected to be home to a third of the global population [23].

Certain aspects of urbanisation in SSA are unique to the region. Firstly, rural-urban migration in SSA is typically a result of 'push' factors such as deteriorating farming conditions worsened by climate change, fluctuating agricultural prices, natural disasters, and even violent rural conflicts [24, 25]. These forces drive people towards urban areas, despite not necessarily corresponding to an increase in overall productivity [25, 26, 27]. Simultaneously, urban infrastructure has not been able to match the pace of rapid urbanisation. As a result, less than half of the urban population in SSA has access to essential services such as improved sanitation as outlined in Section 2.2.2. Thirdly, urbanisation in SSA often takes the form of urban concentration, with a large portion of a country's urban population living in a single, rapidly growing city [28]. Today, key emerging megacities include Dar es Salaam, Lagos, Kinshasa, Luanda, and Johannesburg. Forecasts suggest that by 2100, Africa will house thirteen out of the world's twenty largest urban agglomerations [23]. The fastest growing cities in Africa include Nairobi as well as Accra and Dakar [29].

2.2.1 Opportunities of growing cities in Sub-Saharan Africa

Urbanisation in Africa has brought significant advantages to city dwellers. Although there are some negative aspects associated with such a massive transformation as mentioned in Section 2.2.2, the benefits in terms of economic opportunities and improved living conditions provided by African cities far outweigh the drawbacks [30]. The urban populations in SSA tend to outperform rural populations on multiple aspects. First, the availability of job opportunities makes cities attractive places to live. In comparison to rural areas, hourly wages in cities are more than double and underemployment is less prevalent due to urban workers working longer hours per week (about 30% more hours per week) [30, 31]. Education opportunities are also more accessible in African cities, with young residents receiving between 2.5 to 4 years more education than their rural counterparts. This not only generates economic and social benefits throughout their lives but also allows for the transition to a skilled economy [30, 32]. Secondly, cities also offer better infrastructure and services, such as access to electricity, home ownership, and banking. About 80% of Africans living in cities are connected to the grid, compared to only 20% of rural households [33, 30]. Cities serve as incubators for social and cultural change, with lower fertility rates and more favourable dependency ratios than rural areas [30]. Thirdly, health and life expectancy are usually higher in cities [34, 35, 36]. This is often connected to access to health facilities and overall better sanitation [37, 38, 39], but also opportunities for better nutrition and public safety [40].

There is a well-documented relationship between economic growth and urbanisation [41, 42]. Although urbanisation is a multifaceted phenomenon, successful urban development can yield numerous advantages such as concentrated economic activity, shorter trade routes, access to jobs, shared infrastructure, and division of labour [43, 44]. Nonetheless, urbanisation alone does not guarantee economic growth; the nature of the urbanisation process plays a critical role. It is vital that growth is paralleled by advancements in infrastructure, efficient public service delivery, and competent governance to ensure that

urbanisation positively impacts broader economic development [27, 45, 41].

2.2.2 Challenges of growing cities in Sub-Saharan Africa

On average, life in the city is usually healthier and more convenient [36] as outlined in Section 2.2.1. However, the *average* urban dweller rarely exists per se. This is because high inequalities prevail in cities in SSA. Globally, SSA experiences the second-greatest income disparity, surpassed only by Latin America, and South African cities are recognised as the most unequal [46]. Even though the percentage of the urban population living in informal settlements in SSA has decreased from 64% to 51% over the last two decades [47], there is still a significant percentage of the population living in inadequate housing, without access to improved water and sanitation. Additionally, these communities face the highest environmental risks, such as floods or droughts, which are amplified due to climate change [48, 49].

In contrast to the vertical growth observed in China and Latin America, cities across SSA are sprawling outwards, a dynamic that is shaping both socio-economic and spatial realities [50, 51]. These socio-economic inequalities manifest spatially, meaning many cities in SSA, have fragmented spatial structures with pockets of slums neighbouring high-rise business buildings and scattered fringe developments [50, 4]. The World Bank reports that between 46% and 77% of city growth in Africa has been due to expansion, rather than densification. A notable feature of this sprawling growth is *leapfrogging*, a process where development often bypasses the city centre, leaving patches of undeveloped land in between, further enhancing the scattered and fragmented nature of these cities [52]. This horizontal growth, however, carries considerable economic implications [51]. It is generally costlier to provide essential services such as water, power, and transportation to widespread cities compared to compact ones. As a result, residents often resort to alternative solutions, such as private water trucks and waste collectors, increasing living costs [48]. Furthermore, the scattered structure hampers efficient labour markets. Compact

cities and those with efficient transport systems enable easier access to job opportunities, facilitating better matching of skills with vacancies. On the macro scale, the economic burden is palpable. The relative cost experienced by African households in terms of their per capita GDP is higher than their counterparts in other regions due to the scattered nature of these cities [4, 51]. This further exacerbates the socio-economic inequalities within the region. While the overall life expectancy and health are higher in cities in SSA, living in urban environments also comes with specific health effects. Air and noise pollution, often exacerbated by rapid urbanisation, are significant concerns, associated with various health issues, from respiratory diseases to stress-related illnesses [53, 54, 55]. Similarly, inadequate housing, marked by substandard sanitation and ventilation, is associated with various health issues. These range from mental health disorders to respiratory and diarrhoeal diseases, to vector-borne diseases such as malaria [56, 57, 58]. The availability of healthcare services can be a hurdle in sprawling peri-urban areas. Even with superior facilities compared to rural areas, uneven distribution of health facilities makes access difficult for residents of informal settlements or city fringes, negatively impacting their health [59]. Lastly, evidence suggests that the decreasing access to urban green spaces has a negative impact on health and liveability [60, 61].

2.3 Post-colonial urban inheritance and policy in Sub-Saharan Africa

Much of the existing literature has long perpetuated the perspective of African cities as dangerous and dysfunctional [62]. Literature on African cities after the turn of the millennium represents a more nuanced view and emphasises the economic rationality of numerous activities in African cities, viewing them as the fount of cultural creativity and innovation [62], as well as places of opportunity and overall better health as outlined in Section 2.2.1.

From a perspective of the Global North, cities in SSA are very easily classified into *formal* and *informal* [63], often representing the dichotomy of the ‘traditional’ or ‘African’ and ‘modern’ way of living. The former is often represented by unstructured and spontaneous forms of living and building, and the latter assumes the shape of well-built (as defined by colonisers) city centres that often mirrored the built structures in their homelands [63]. Yet, these parts of the city deemed informal by the Global North, because housing has been constructed without permission, often contain very similar structures compared to formal types of living. In many cases, they share the same city structures (street network and housing patterns) as the officially acknowledged neighbourhoods next to them [64, 63]. Throughout this thesis, I make the distinction between informal and formal settlements, but I want to highlight that these distinctions mainly impact land governance and policy. Informal and thus unregulated does in most cases not imply disorder, since urban settlers of ‘informal’ settlements often follow strict rules and regulations, for example, street vendors paying ‘taxes’ to specific ‘owner of a place’ [63].

Colonial urban planning patterns, such as segregated residential areas, central business districts, and infrastructure, continue to shape the spatial organisation and development of many cities in the region [63, 65]. In the post-colonial era, African governments have sought to reformulate urban policies to promote more inclusive, sustainable, and equitable urban environments. These efforts include upgrading informal settlements, investing in public transportation systems, fostering participatory planning processes, and preserving green spaces [66]. However, the persistence of the colonial legacy, coupled with rapid urbanisation, population growth, and limited resources, presents significant challenges and vary depending on the region [66]. In addition to lack of resourcing, issues related to governance persist, such as incentives at different scales of government, which in cases can create social conflict for urban policy [67, 68, 69]. In the following section, I describe regional differences in colonial urban inheritance and today’s urban policy efforts in East and West Africa, with a focus on the cities studied in this thesis, namely Accra (Ghana),

Dakar (Senegal), Dar es Salaam (Tanzania), and Kigali (Rwanda).

2.3.1 Colonial city planning in the study cities

The influence of colonial histories varies across cities in SSA, with their imprint on urban development and spatial configurations being more pronounced in certain locations. In West Africa, for instance, urban growth is significantly shaped by its colonial past [70, 65, 71]. Accra, which was under British colonial rule, epitomises this, showcasing a town planning rooted in racial considerations, and a mixture of older towns influenced by various ethnic groups. ‘New Town’, often the first port of call for city migrants, exemplifies these historical influences [72]. Contrastingly, West African cities under Francophone colonial rule, such as Dakar, bear a different colonial imprint. They tend to be more compact, less sprawled, and display more intense development in the city centre [70]. During the colonial era, the French government enforced greater centralisation and uniformity in the management of urban areas, with included implementing standardised land use planning throughout the city, and the use of an integrated road grid system (*Plan hippodamien*) as seen in Dakar [70].

The colonial impact on city development extends to East Africa as well. Dar es Salaam’s urban landscape mirrors Accra in many ways, reflecting the segregation principles of its British and German colonial past. Elements such as waterfront areas reserved for colonisers, architecture emulating European styles, and separate sections for ‘Indian and Arab-style’ residences attest to this history [73, 74].

In contrast, colonial city planning is less prominent in Kigali, another East African city studied in this thesis. Having become Rwanda’s capital only post-independence in 1962, and having undergone extensive damage during the Rwandan genocide, Kigali’s urban landscape differs significantly from cities with deeper colonial imprints and is mostly influenced by its geographic landscape.

2.3.2 Urban development trends and policies in the study cities

Urban landscapes across SSA display both common trends in urban development as well as unique features stemming from regional geography, historical context, and government structures. As such, city-specific policies tailored to address each city's unique challenges are essential. Urbanisation in West Africa is notably characterised by the emergence of vast 'ribbon' developments or urban corridors along coastal regions. These areas often extend beyond national borders, marked by high levels of commuting and cross-border migration [66]. As these urban stretches further expand inland relative to the coast, they incorporate additional localities, culminating in crowded agglomerations teeming with vibrant business districts. The increasing demand for prime real estate in these West African cities has driven up property prices, leading to the marginalisation of low-income communities through evictions and challenges in secure land ownership [75, 76]. This situation has underscored the pressing need for upgrading informal settlements, with community-led organisations lobbying for reforms in urban management and enhanced service provision [66].

Meanwhile, East Africa, the global leader in urbanisation rates, experiences similar challenges [3]. The region grapples with critical issues such as high unemployment, growth of densely populated informal settlements, inadequate infrastructure, and environmental degradation [66]. For example, Dar es Salaam is one of a few capital cities globally where infant mortality rates at birth exceed the rural averages [77]. This pattern highlights the potential adverse outcomes associated with rapid urban growth that is not effectively managed or properly coordinated. Similar to West Africa, challenges of urbanisation in East Africa are often amplified by administrative inefficiencies and bureaucratic complexities. Nevertheless, recognising the necessity for sustainable and equitable urbanisation, national governments have introduced National Urban Policies (NUPs) [66]. Despite their increasing adoption across SSA, only 17 out of 44 countries had established explicit NUPs

as of 2015, though countries such as Ghana, Rwanda, and Tanzania have made considerable strides in this direction [66]. Ghana has distinguished itself with a comprehensive NUP framework and an action plan that translates the framework’s goals into tangible initiatives [78]. Similarly, Rwanda has demonstrated its commitment to urban issues through its ambitious ‘Kigali Masterplan’ and the adoption of its NUP in 2015 [79]. Tanzania, on the other hand, employs a national master plan and an urban planning act to manage its urbanisation challenges, illustrating a more localised approach, as exemplified by the ‘Dar es Salaam Masterplan’ [80, 81, 66]. Despite expressing interest in developing a NUP, Senegal has yet to make significant progress [66]. While considerable progress has been made in policy formulation, the urban policy landscape in SSA remains diverse and inconsistent, with some countries lacking an explicit NUP. Limited financial resources continually hinder the timely achievement of urban development goals, even in nations with established NUPs such as Ghana. Nonetheless, several countries, including Rwanda, have displayed notable commitments, both effort-wise and financially, towards enhancing urban planning and living conditions [66].

2.4 Data gaps in Sub-Saharan Africa

Several international frameworks and organisations have underlined the need for more data and research on urban areas specifically, including a Sustainable Development Goal (SDG) dedicated to making ”cities and human settlements inclusive, safe, resilient and sustainable (Goal 11)” [82]. Despite the recognition of necessity to forge new knowledge to answer complex urban challenges, urban sciences of today are very fragmented and often focus on the Global North [6]. Datasets that capture information on urban features of the built and natural environments, such as roads, buildings, and green and blue space, especially measured on consistent spatial scales and time frequencies, are also limited. While there has been a notable shift towards collecting and analysing this data in recent years, driven by sustainable development initiatives and endeavours by major

companies such as Google and Microsoft [7, 8, 9], gaps in data availability still exist. Even when data is accessible in SSA, its spatial and temporal consistency often varies [11, 12, 13]. Addressing these challenges is essential for fostering inclusive urban planning and building resilient cities that can accommodate the region’s growing population [13]. Effective solutions require the integration of diverse knowledge sources, ensuring a holistic understanding of urban complexities [83].

2.5 Studying cities

2.5.1 Complexity of cities

The city is one of the most complex cultural products that has expanded all over the globe. Like language, it is never static, constantly evolving by being put to use and at the same time shaping how we express ourselves through vocabulary and grammar. In the same way, cities influence how we live. We shape them through moulding the urban spaces, with form and function being heavily intertwined [84]. Urban *morphology* explains just this form-function complexity, a term coined by the German writer Goethe and originally used in bioscience to describe the form of biological systems and the relationship between its constituent parts [85]. A city is therefore not just the sum of all of its parts (in form of buildings and empty spaces), but also how humans interact with them (their function). The spatial signatures of cities can reveal different stories of history, society, and technology, all of which are encoded in the composition of built and natural environments [86]. Cities tend to share morphological-spatial configurations with their geographical or cultural neighbours [87]. For example, European cities share a common makeup that shows high structural (and functional) building variability with the urban density decreasing from the city centre to the periphery. Eastern European cities have a mix of patterns from Western capitalist urbanisation as well as from their socialist past [88]. Cities in the United States, share similarities in terms of connectivity and high density outside of

the city centre and are in general highly influenced by automobile transport since they often blossomed in the second half of the 20th century [89]. These cities also share a resemblance with some cities in Asia and South America that have a colonial or dominion heritage, such as Singapore and São Paulo [87]. Cities in East and Southeast Asia have grown vertically, with average building height increasing over time [51], in contrast to cities in SSA, which have a high share of compact, low-rise buildings, displaying indications of city sprawl that reflect stark contrasts, such as business developments next to informal settlements and other indications of spatial inequalities [90, 4]. While they are similar in these regards, they also have some region-specific intricacies discerning cities in East Africa from cities in West or Central Africa and vice versa. East African cities tend to have higher structural variability in building size compared to Central African cities [87], while, both have a large share of low compact, low-rise buildings. Cities in West Africa share some similarities with some Asian cities (especially cities with British colonial history, e.g. Jaipur, Bangalore, and Baghdad), with a larger share of large low-rise buildings compared to the other cities in SSA. Large low-rise buildings could also be an indicator of commercial, institutional, and industrial activities occurring in a specific city.

There are multiple ways of studying the urban complexity of the built and natural environment in a temporal, spatial, visual, scaling or connectivity context [91]. All these different quantifiable features of a city reflect and influence the well-being of its dwellers [92, 93]. However, as described in Section 2.4, data on various features of cities' environments, such as vegetation, roads, buildings, and population settlements, is limited and infrequently gathered in low-income countries, or if data are available, have inconsistent spatial and temporal coverage and resolution. Consequently, studies on urban environments often focus on individual features, such as the extent of urban areas [10, 94], land use and land cover [95, 96, 16, 97], including green and blue spaces [98, 99], roads and connectivity [100, 101, 102], and population density [103]. These elements are interrelated and display complex patterns across different scales which are relevant for policy

implementation at different administrative levels.

2.6 Using images for studying cities

Research suggests that multidimensional factors of urban development can be measured from images of the urban environment, which provide new opportunities for studying cities [104, 105, 106]. The idea that images can be used as proxies that represent the development of a city at multiple scales is based on the premise that these factors can be either measured through direct visual signals or indirect factors that influence the urban environment [72, 107, 108]. Recent advances in Artificial Intelligence (AI) and remote sensing have leveraged the potential that lies in urban modelling with imagery data from different sources and at various temporal and spatial scales [15, 16, 109, 110].

In Section 2.6.1, I provide a short summary of the literature pertaining to urban analysis via street-level images. Subsequently, in Section 2.6.2, I explore in greater depth the existing body of work on the use of satellite images and computer vision for studying cities.

2.6.1 Street-level images

The increasing availability of street-level data has opened up new possibilities for understanding and predicting various aspects of urban life. Recent studies have turned to the application of modern Machine Learning (ML) methods on street-level images to discern visual city attributes and extrapolate insights about a wide range of urban phenomena. Arietta et al. [105] used non-linear support vector regression to identify visual city characteristics in street-level images, thereby predicting non-visual city attributes such as crime rates, housing prices, and population density. Likewise, Gebru [106] identified a correlation between the make and model of cars detected in Google Street View images and the election results in the United States. Suel et al. [111] used street-level images to predict

health and socioeconomic inequalities across London. These works illustrate that image data, combined with visual heuristics extracted from these images, can be a powerful tool to complement more labour-intensive approaches in detecting socioeconomic trends with fine spatial and temporal resolution.

Furthermore, researchers have used street-level images to measure environmental factors that impact health such as urban green and blue spaces in cities. These either focus on measuring near-road greenery or blue/green spaces [112, 113] or look at specific types of vegetation such as tree coverage [114, 115]. Additional research, focuses on urban green in relation to walkability or physical activity [116, 117]. Further studies examine the relationship between the natural and built environment and health [118]. Additionally, street-level images have been used to map environmental exposures, such as air and light pollution [119, 120, 121]. Other researchers made use of street-level images to predict urban appearance and perceptions of safety [122, 123, 124, 125].

Generally, most studies adopt a supervised approach, focusing on aspects such as the built and natural environment [126], exposures [127], and perceptions [128] using images. Beyond street-level images, researchers have also explored alternative data sources for studying cities, such as scraped social media images of cities or GPS data such as traffic counts [129, 127, 130, 131].

2.6.2 Satellite images

Remote sensing makes use of the electromagnetic energy that physical subjects reflect and uses sensors to measure parts of the electromagnetic spectrum. Urban remote sensing has been developing over the last fifty years, with the first land cover maps created from a colour infrared film from a hand-held camera from the Apollo and Gemini mission in 1965 [132]. Later in 1972, the launch of the Landsat-1 satellite moved urban remote sensing from airborne to satellite level. Research in urban remote sensing increases significantly in 1999 with the advent of satellite sensors that capture images at a high spatial resolu-

tion (5- 0.5 meters), such as IKONOS and QuickBird [132]. Today, very-high-resolution images can capture images with a spatial resolution of < 1 meters [133], and together with the advancements in ML have opened up new opportunities for tackling real-life problems at larger scale [109].

Satellite images can be used to gather various layers of information about an area at different scales and across flexible points in time, ranging from detailed analysis of individual buildings or objects to studies of entire continents. The use of automated methods for data processing and image analysis makes it possible to quickly and accurately infer information for use in supporting decision-making [134]. Remotely sensed data, collected from a range of sources and covering a wide range of spatial, temporal, and spectral scales, can be used to map and understand important environmental characteristics of urban areas [135]. These include mapping neighbourhoods, urban extent or land use [136, 94, 137, 72], their indices (such as Normalised Difference Vegetation Index (NDVI) or socio-economic status) [138, 139, 140], and model human and physical processes [135, 141, 142].

2.6.3 Computer vision for satellite images

Early efforts in computer vision for remote sensing were focused on enhancing images, extracting features, and performing basic classification tasks, such as land cover mapping and change detection. Traditionally, the information was derived through approaches such as spectral indices [143] and pattern texture extraction methods [144] on a pixel level. For instance, the NDVI is a spectral index frequently used to assess vegetation coverage by measuring the difference between near-infrared and red light reflected by the vegetation in a given area [94, 138]. Similarly, texture extraction techniques, such as the Grey Level Co-occurrence Matrix (GLCM), have been employed to analyse the spatial arrangement and frequency of pixel brightness values, thereby providing valuable insights into land-use patterns [145].

The evolution of satellite imaging technology, coupled with the advent of high-resolution satellite sensors able to capture data at less than 1 meter per pixel, has precipitated a significant shift in image classification methodologies, transitioning from a pixel-based approach to an emphasis on object and scene-level analysis, thus challenging the adequacy of pixel-based semantics [146, 147]. In 2001, Blaschke and Strobel stated that analysing satellite images on the object level is the way forward, and indeed object-based analysis has been dominating the remote sensing analysis in the past 20 years [148, 146]. Furthermore, with the publications of the first large remote sensing scene classification datasets, such as UC Merced in 2010, together with the developments in deep learning, a new paradigm of scene-level analysis was introduced [21]. The emergence of deep learning and Convolutional Neural Networks (CNNs) has further revolutionised the field of computer vision for remote sensing, enabling the automation of complex tasks, such as object recognition, semantic segmentation, and large-scale environmental monitoring, described in Section 2.6.4. Today, computer vision for remote sensing plays a crucial role in various urban applications, including urban planning, natural resource management, disaster response, and climate change research.

2.6.4 Deep learning for satellite images

Advances in computing power have unlocked the potential of neural networks run on imagery data. Neural networks are algorithms that consist of connected nodes, which are designed to recognise patterns [149]. Deep networks are wider networks that have multiple layers of nodes that are capable of arbitrarily approximating most functions [150]. In a supervised setting, the process involves guiding the neural networks during training by providing labelled examples, allowing the algorithms to learn and generalise patterns. Label allocation, a crucial aspect, however, can be time and labour intensive, demanding human expertise to assign predefined categories to a subset of the data. Despite this challenge, the main advantages of deep learning algorithms in remote-sensing applications

are the ability to process large datasets as well as to distil information to create richer representations that allow further complex modelling [22].

The development of CNNs has aided the understanding of cities with machine eyes and thus highly influenced urban analytics [14, 15, 16, 109, 110]. For example, Zhao and Steele, among others, analysed urban temporal and spatial change using satellite remote sensing with ML and geo-spatial statistical tools to predict poverty and slum development [151, 152, 153, 108, 154, 72]. Additionally, several researchers applied supervised deep learning tools to imagery datasets to map complex spatial and temporal patterns. Huang and Zhang mapped urban land use and cover classification with deep neural networks and satellite imagery data [16, 155]. Furthermore, Jean and Babenko combined medium and high-resolution satellite imagery with ML tools to predict poverty in Africa and Mexico respectively [140, 156]. In 2020, Yeh together with Burke [157], published a paper that analysed economic well-being with supervised neural networks in Africa. Moreover, satellite images together with deep learning algorithms have been used to measure environmental exposures, such as noise and air pollution or wild fire smoke [158, 159, 160, 161]. Additionally, researchers measured levels of urban green across in countries and cities [162, 163] as well as greenery relating to health [127, 164].

2.7 Unsupervised learning

Unsupervised learning refers to the exploration and interpretation of data without reliance on external guidance or predefined categories. It is a process analogous to categorising a collection of objects based solely on observed attributes, without prior knowledge or labels. This approach is agnostic to any external or previous knowledge, allowing for the identification of inherent structures and patterns within the data itself, uncoloured by external biases or assumptions [17, 19].

2.7.1 Clustering

Unsupervised algorithms can help to find patterns in data without tendencies in any particular outcome, which can resemble what we may do as humans in finding visual similarities and thus forming hypotheses about the data observed [17, 18]. While human interpretation is possible on small scales, the increasing amount of information in the development of ‘big data’ has shifted the focus onto machines to make inferences from these large amounts of data. The terms unsupervised learning and clustering are often used interchangeably, however, differ slightly. While clustering is the one main form of unsupervised analysis, unsupervised analyses also include dimensionality reduction techniques or anomaly detection methods. This thesis, however, is focused solely on clustering as a form of data structuring, or grouping parts of the data into subcategories.

The idea of clustering is based on the premise that data points in one cluster are as similar as possible to the data points in the same cluster, and likewise, as dissimilar to features in other clusters [165]. Generally, clustering methods do not require labelled data; the definition of similarity and the loss function is the main determining factors of the algorithm. Distance (dissimilarity) and similarity are the fundamental concepts of creating clustering algorithms. When quantitative data are clustered, distance is the preferred method to measure its relationship. And vice versa, similarity is usually put to use when dealing with qualitative data features [19, 166, 165]. Clustering algorithms have previously been applied to image data, such as for low-level tasks like border detection or image segmentation by colour intensities [167]. These early computer vision tools detect patterns across a number of pixels, however, fail at recognising the context of pixel groupings. Today, a CNN can successfully create visual feature representations that can capture not only objects but their context in a picture [168, 144], as described in Section 2.6.4.

2.7.2 Unsupervised image analysis for studying cities

Unsupervised learning has proven to be effective in analysing and deciphering the complexity of cities, with applications ranging from measuring urbanisation to studying the built environment or urban dynamics [20]. In urban studies, ML techniques that rely on labelled data (supervised techniques) are most commonly used [169, 170]. Although supervised methods have been effective in many urban studies applications and datasets, they may not be suitable for all research questions due to the challenges of obtaining labelled data in real-world urban settings [171]. Unsupervised methods can highlight trends solely based on the features present in the dataset [172], providing further perspectives outside the assumed outcomes in a supervised analysis. While unsupervised approaches are very common in morphometric characterisation [173, 174], i.e. using clustering methods to classify the typology organically with information based on numerical features of the urban environment, it is hardly used for remote sensing images. Yet, with the advent of high-resolution imagery, this is predicted to change since researchers are able to capture the complexities and variations of land cover on a very small spatial scale [175].

2.7.3 Unsupervised machine learning for satellite images

Unlike in supervised analysis where labels are present, unsupervised analysis groups the pixels of an image without any external information. Until recently, the majority of unsupervised analysis or clustering of satellite images is applied on a pixel level, where pixels are grouped together that have similar pixel intensities [176, 177, 178, 179]. The most common pixel-level algorithms used are k-means [180], and its variant Iterative Self-Organising Data Analysis (ISODATA) [181] [144], or fuzzy approaches where the pixels can be assigned to multiple clusters [182]. With the developments in deep learning, the analysis has moved from pixel-level to patch level, since deep learning approaches can learn features of different levels of abstractions [21].

2.7.4 Unsupervised deep learning for studying cities with satellite images

Little research has been conducted in the field of unsupervised deep learning in combination with remote sensing to analyse urban environments and related outcomes [22]. Most of the deep clustering approaches in the literature either use a pre-trained network for feature extraction or an autoencoder to learn the features of the images, which are subsequently clustered [183]. Block et al. [184] used a CNN feature extractor together with k-means clustering for applications in demographic analysis across six cities. Other researchers used CNN feature extraction together with clustering to map schools in Liberia as well as monitor the development of a refugee camp in Bangladesh with high-resolution satellite images [185, 186]. Additionally, unsupervised deep learning has been used for change detection in longitudinal analysis of satellite images [187, 188].

Recently, self-supervised learning methods have emerged as a promising avenue to address the limited availability of labelled remote sensing data, using ‘pretext tasks’ derived from the inherent structure of the data for model training, such as multi-temporal image ordering [189] or finding *neighbourhoods* in datasets [190]. Additional explorations in this domain include unsupervised learning algorithms for clustering Synthetic-Aperture Radar (SAR) images [191] and analysis of Light Detection and Ranging (LiDAR) data for geo-demographic predictions [192].

Overall, unsupervised and self-supervised learning methods present compelling prospects for advancing the analysis of remote sensing data, especially considering the challenges associated with the lack of labelled datasets. Nonetheless, these promising approaches warrant further investigation to affirm their robustness across diverse settings and their compatibility with satellite images obtained from various satellite sensors. A critical aspect to consider is the trade-off between the explainability of unsupervised clusters and their need. While larger datasets enhance the explainability and validation of clusters, they might paradoxically reduce the urgency for unsupervised learning due to the data

richness. There exists a potential sweet spot where sufficient data allows for cluster validation, while the derived clusters offer significant insights or unveil meaningful categorisations across dimensions, as explored in this work. One feasible strategy is to develop and refine these unsupervised methods in a data-rich environment and subsequently adapt them to scenarios with limited data, ensuring there's enough for cluster validation - a concept that mirrors the multi-city clustering approach pursued in this thesis. Through this thesis, I seek to evaluate the potential utility of unsupervised deep learning with satellite images, to showcase its application in an urban development context and across multiple cities.

2.8 Summary

Urbanisation is rapidly increasing worldwide, with SSA playing a significant role in driving this global trend. Even though the overall living standards in cities are higher compared to those in rural areas, cities in SSA are facing multiple challenges due to their crowded, costly and poorly connected nature. Yet, urban growth also creates opportunities for fighting poverty and sustaining development [5]. To create policies that unlock the potential of expanding cities, it is important to understand the nature of urban growth and how it is linked to urban development and well-being at different scales. But cities are complex systems and data on urban environments are often fragmented and not frequently updated. Recent developments in computer vision, together with increasingly available high resolution satellite images have opened up possibilities for studying cities in new ways and on larger scales. Clustering algorithms allow researchers to find patterns in data without tendencies in any particular outcome and analyse complex combinations of urban features and characteristics and their relationships. However, only little research has been conducted in the field of unsupervised deep learning in combination with remote sensing to analyse urban environments and related outcomes.

Chapter 3

Study areas, data sources and data lessons

3.1 Overview and objectives

I identified, processed and collated data for clustering satellite images of four cities in Sub-Saharan Africa (SSA). This allowed me to implement the single-city analysis of Accra described in Chapter 4 and Chapter 5, as well as the multi-city clustering in Chapter 6 and Chapter 7.

In this chapter, I describe three different sets of data used for the analysis. First, I give an overall overview of the study cities. Second, I describe the high-resolution satellite data of all study cities, and data lessons learned. Finally, I outline the data used for cluster interpretation, which includes data of the built environment and vegetation, demographic data and land cover maps.

3.2 Study areas

3.2.1 Selection of cities used for analysis

This thesis focuses on four cities in SSA, namely Accra (Ghana), Dakar (Senegal), Dar es Salaam (Tanzania) and Kigali (Rwanda). The cities selected for analysis have shown rapid population growth in the past and are projected to continue growing [3]. The choice of cities is based on the estimated and projected population change in the past and spatial location. The analysis of Accra, Dakar and Dar Es Salaam gives a perspective on cities that are large seaports, with the former two located in West Africa and the latter in East Africa. Kigali is land-locked Eastern African city that changed rapidly in the past century. Additionally, the analysis captures a diverse set of landscapes and climate zones, ranging from tropical arid (Dakar) in the West, to tropical and humid in the East (Kigali).

3.3 City profiles

3.3.1 Accra, Ghana

Accra is the administrative, economic and political capital of Ghana, with a population of 5 million inhabitants [3] and an area of $1,500 \text{ km}^2$. The population of Greater Accra Metropolitan Area (GAMA) increased by almost 90 per cent from 2000 to 2021 [193]. Foreign investment, including in natural resources such as oil, have made it an international hub for trade, technology and education in SSA. Despite its recent development Accra still struggles with poverty, especially in terms of housing, health and sanitation, which persists particularly in informal communities [194].

The neighbourhoods of the city are, like most West African coastal cities, highly influenced by their colonial past [72]. The urban area of Accra and its neighbourhoods reflect features such as town planning based on race until 1923, older towns dominated by dif-

ferent ethnic groups that merged into the city, as well as parts such as ‘New Town’ which have been the first entry point into the city for migrants [72]. Today, the city’s diverse population includes members from all major ethnic groups of Ghana as well as people from neighbouring countries such as Nigeria, Togo and the Sahel region [195]. The city is mainly following a trajectory of economic segregation as described by Agyei-Mensah [195]. Additionally, Accra has outgrown its colonial infrastructure, and little urban planning combined with unnamed roads has increased the importance of neighbourhoods to navigate the city [196]. Nevertheless, it is quite difficult to delineate those neighbourhoods from each other, especially when place, poverty and health of city dwellers are so interlinked [197].

Pathways Workshop, 2019

In order to gain a local perspective on the issues related to inequalities in the city of Accra, I attended a workshop hosted by the *Pathways to Healthy Equitable Cities* Group at the University of Ghana that was attended by a diverse group of local policy and civil society stakeholders, along with members of the international research team. The workshop was organised with the purpose of aligning Pathways’ projects’ themes to policymakers’ and decision-makers’ priorities. Participating in the workshop and being involved in the discussions, I found that one occurring theme that was highlighted by the participants was the problem of housing and listed land ownership. This issue has also been mentioned by Byamugisha [198], who states that only 10 per cent of all land SSA is registered, as well as Toulmin [199] who estimates that only 2 to 3 per cent of land in West Africa is recorded by the government. Accra, like a lot of cities in SSA, is growing rather outward than upward, with built areas predicted to increase by 277 per cent from 2001 to 2025 [200, 201]. These implications of population growth also intersect with additional issues highlighted by the participants, including lack of clean water, sanitation, effective transportation, and inadequate formal affordable housing. Analysing urban development and its manifestation

in the urban form will thus help to understand the spatial component of city well-being.

3.3.2 Dakar, Senegal

Senegal is ahead of the average urbanisation rate in SSA, and about half of Senegal's population lives in urban areas. Its capital Dakar is a city in West Africa, located on a peninsula surrounded by 70 kilometres of coastline [202]. About 3.5 million people live in the Dakar city region and the population is expected to grow to 5.1 million in 2035 [3]. The city is the administrative and economic capital, as 80% of firms and modern employment opportunities, and 87% of local finances are located on the peninsula [203]. Senegal's capital is a hub for shipping and transport and most of the economic gains are due to fish and phosphates exports and services [204]. Like many other cities in West Africa, Dakar has a history of colonialism by Europeans and was used as a centre for the cross-Atlantic Slave trade but was later taken over by the French in the 19th century [205].

The colonial past of Dakar has had a significant influence on its urban development and architecture, with the city showcasing a blend of traditional African and European architectural styles. The historic district of Dakar, known as Plateau, bears the marks of French colonial rule with wide boulevards, public squares, and colonial-era buildings [206, 207]. Following Senegal's independence in 1960, several attempts at designing master plans have largely been ineffective due to a lack of consideration for local realities and inadequate management of disaster risks. Despite Senegal lacking a National Urban Policy (NUP) as mentioned in Section 2.3.2 [66], a master plan has been formulated specifically for Dakar. However, the 2035 'Dakar Masterplan', although having designated investments, falls short in terms of public engagement and policy implementation and tracking [202].

3.3.3 Dar es Salaam, Tanzania

Dar es Salaam is the former capital and by far the largest city of Tanzania, situated at the coastline with a natural harbour. Majid bin Sayyid, Zanzibar's Sultan, named the city 'Haven of Peace', yet the name has not kept it from forced colonial rule afterwards [208]. The city became the administrative and commercial centre during German rule, which was maintained after the First World War under British trusteeship. Today, the city plays a critical role in Tanzania's economy, serving as the country's largest seaport and a hub for industrial activities for industry, education, and culture. Although it is home to just 8 per cent of Tanzania's total population, it contributes about 70 per cent to the country's GDP. Population growth including rural-urban migration has doubled the urban area in size between 1990 and 2014. The population is predicted to grow from approximately 7 million inhabitants today to about 13.4 million by 2035, making it the first megacity in East Africa [3].

Like in other African cities, colonial city planning has left a large imprint on the urban fabric. The city was racially segregated with the best parts reserved for the European colonisers, followed by the Indian diaspora. The urban space for Africans was moved behind a green belt in densely populated and unserviced areas [81]. Today, more than 60 years after the country's independence, the city morphology still captures a mix of German, British and Arabic architecture, yet, the single-story, detached, six-room Swahili houses prevail.

Unplanned urbanisation and lack of policy implementation have led to scattered city growth and urban sprawl posing a challenge to the authority [81]. The City council has mapped out a master plan to tackle the challenges of rapid urbanisation with a key point of developing satellite towns within the city to reduce traffic congestion and improve services [209]. Currently, about 40% of the urban population lives in informal settlements [47] and less than 10% of the city is connected to the sewage system [210].

3.3.4 Kigali, Rwanda

Kigali, the capital and largest city of Rwanda, has experienced significant population growth in the post-genocide period perpetrated against the Tutsi in 1994. Currently, the city has 1.6 million inhabitants, an incredible increase from about 6,000 inhabitants after its independence in 1962 [211]. The most significant change came about due to the return of hundreds of thousands of Tutsi refugees who had spent years away from Rwanda. Since then, Kigali has undergone an impressive transformation and evolved into a modern and cosmopolitan commercial centre [211].

Kigali's economy is primarily driven by the services sector and has attracted significant foreign investment, and the government has implemented policies that have made it easier to start and run businesses in Kigali. The city is often used as a role model for its green city development and is highlighted as a great example of urban development in Africa [212]. Kigali is known for its innovative urban planning and sustainable development initiatives, as outlined by the city's master plan [213, 211]. This includes the 'green city' initiative, which aims to increase the number of green spaces, promote sustainable transport options, and reduce carbon emissions [213]. Despite significant progress in post-conflict Rwanda and Kigali's portrayal as 'model city', challenges persist. The city's master plan has been criticised for its lack in public consultations and participatory planning, especially missing integration of informal sectors (housing and employment) [212]. The implementation of policies related to the master plan has in some cases led to displacement and increased poverty for urban dwellers of informal settlements [214].

3.4 City boundaries

The city boundaries used for defining the study areas in this thesis are based on official administrative boundaries released by governments in the past five years. Table 3.1 lists the official boundary source and name. All city boundaries, except for Accra, refer to the

city region, which is equivalent to administrative level 1.

City	Boundary name	Source
Accra	Greater Accra Metropolitan Area (GAMA)	Ghana Statistical Services [215]
Dakar	Région de Dakar	National Agency of Statistics and Demography of Senegal (ANSD, Agence Nationale de Statistique et de la Démographie) [216]
Dar es Salaam	Dar es Salaam Region	Tanzania National Bureau of Statistics / UN OCHA ROSA [217]
Kigali	Kigali City	National Institute of Statistics Rwanda [218]

Table 3.1: List of boundary shapefiles used for defining the study cities.

3.5 Satellite images

There are currently more than 2,000 satellites orbiting planet Earth and most of these satellites are capturing images [219, 176]. Generally speaking, high resolution (HR) satellite imagery is defined at a spatial resolution of 1 to 16 meters and very high resolution (VHR) imagery captures scenes < 1 meter [133]. There is a minimum amount of electromagnetic energy that is required for a satellite sensor to capture intensities [220]. Sensors that capture a wider range of the electromagnetic spectrum (e.g. low spectral resolution) have by design, a higher spatial resolution. Conversely, sensors that only capture a small part of the electromagnetic spectrum, have a lower spatial resolution [220]. Satellite images, in contrast to the most common image datasets, share some distinctive features that need to be taken into account when used for computer vision applications.

In this section, I first describe the Maxar dataset that was released in response to COVID-19, which is the data used for the analysis described in Chapter 4 and Chapter 5. I also write about the purchase of additional satellite images, which are used for the analysis outlined in Chapter 6 and Chapter 7, and data lessons learned in the process of analysis.

3.5.1 Maxar satellite images

Maxar basemap

The "basemap" is a satellite image product that has been prepared for analysis by Maxar Technologies, a space technology company. This image covers a specific area of interest and is presented at a resolution of 30 meters per pixel, featuring red, green, and blue (RGB) color bands. The images are in a standard geographic projection (WGS84) and typically have a viewing angle of 0-30 degrees from the vertical (off-nadir angle). These basemaps are processed to ensure color balance, and they are formatted into orthomosaics for ease of use. Throughout the thesis, I will refer to the processed satellite image provided by Maxar as *basemap*. The basemap product can usually be purchased from satellite image vendors such as Maxar, and a few urban basemaps were made available for free as part of Maxar's Open Data initiative.

Maxar Open Data initiative

In May 2020, Maxar released a VHR satellite basemap dataset in response to COVID-19 in GeoTIFF (*Tiff*) format [2]. The dataset includes images of 22 cities in SSA, including Accra in Ghana. The full image list can be found in the Appendix, Table A.1. The satellite images capture different urban areas at a point in time between 2017 to 2020. The spatial resolution of the images is 0.3 meters per pixel and capture three different bands (RGB). The commercial pre-processing of the images included colour-balancing and orthorectification. The satellite basemap of Accra (Figure 3.1) covers an area of $1,312\text{km}^2$ and encompasses most of the GAMA, with only parts of the northwest and northeast not covered.

Even though Maxar's Open Data database contained additional cities of SSA, it turned out that some of the images were not suitable for unsupervised analysis. Unlike the Accra



Figure 3.1: The basemap was acquired from Maxar through the Open Data initiative [2]. The blue border represents the Greater Accra Metropolitan Area (GAMA).

images, the images of other cities, such as Dakar, were stitched together from images captured by different satellite sensors. These image variations that were not visible to the human eye were picked up by subsequent unsupervised analysis. Figure 3.2 shows how the satellite image (right) was stitched together from multiple sensors. The cluster map on the bottom left of the image shows that the algorithm captures these variations and assigns certain parts of the image to separate clusters. Additionally, other cities included parts with lots of cloud coverage and could therefore not be used for the unsupervised analysis. Figure 3.3 shows the example of a cloudy basemap image of Dar es Salaam in Tanzania. A large part of the image is either covered in thick clouds or haze, which was also picked up by the unsupervised analysis.

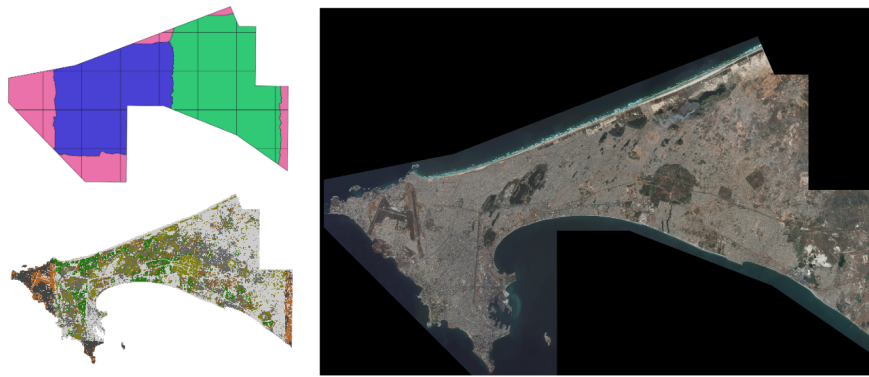


Figure 3.2: Open data satellite image of Dakar acquired from Maxar through the Open Data initiative[2].

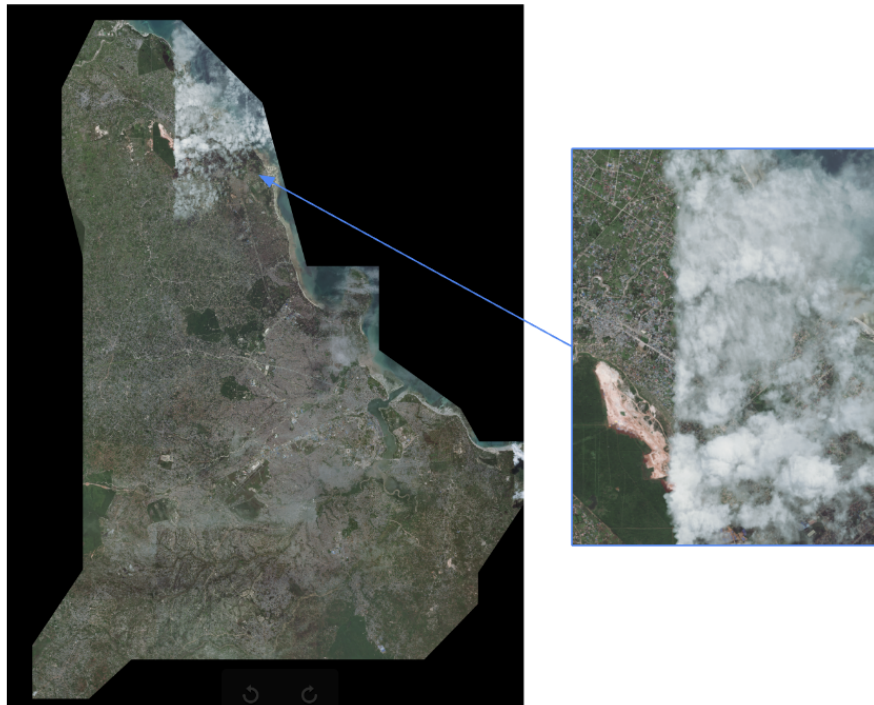


Figure 3.3: Open data satellite image of Dar es Salaam acquired from Maxar through the Open Data initiative[2].

Satellite image purchase

When developing the clustering approach for Accra, I found that the clustering was picking up image variations created by the image stitching as outlined in Section 3.5.1. In order to avoid such issues in the multi-city analysis, I defined the criteria for ordering additional satellite images as follows:

- Very high-resolution (ideally 0.3m/px)
- Low cloud coverage (<3%)
- Satellite images of each city should be captured by the same satellite sensor (ideally also the same sensor/satellite used across all cities)
- 3 bands: RGB (red, blue, green colour channels)
- GeoTiff format
- Orthorectification (pre-processed using a digital elevation model (DEM)), pre-processing level 3
- Recent; from 2019, 2020 or 2021

To make sure the images match the quality of the Accra images obtained through Maxar’s Open Data initiative (as described in Section 3.5.1), and in consideration of the criteria outlined above, I decided to procure the images directly from Maxar. I found that Maxar was the only company that offered full coverage of cities in SSA in very high spatial and temporal resolution. Additionally, Maxar provided leading image processing technology such that the bought images are processed to an analysis-ready level.

Purchased images

I acquired three satellite basemaps from Maxar that matched the criteria outlined in the previous section. The basemaps (Figure 3.4, Figure 3.5 and Figure 3.6) are further described in Table 3.2 and were made available under the internal use license attached in the Appendix Section A.2.

City	Year	Coverage
Dakar	2021/2	570 km^2
Dar es Salaam	2021	1,465 km^2
Kigali	2020	750 km^2

Table 3.2: Satellite image basemaps purchased from Maxar.

The available satellite image for Dar es Salaam (Figure 3.5) did not cover the full city boundary as provided to the vendor. The basemap, however, covers all of the city centre

and most populated areas. Some parts of the south (especially the southeast) are not included in the image dataset. The other city basemaps cover all areas within the city boundary as shown in Figure 3.4 and Figure 3.6.



Figure 3.4: Purchased satellite basemap of Dakar, captured in 2021/2. The blue border represents the Dakar city boundary. Dynamic mosaic ©2022 Maxar Technologies.

City coverage

The images acquired through the Open Data initiative as well as the purchased images, while not fully covering the whole city region boundary, cover all of the metropolitan parts of the city. Accra's city region boundary (Greater Accra Region) covers $3,245\text{km}^2$, which is a very large area compared to the city region of Kigali and Dakar, which cover 750km^2 and 570km^2 respectively. The metropolitan area of the city (GAMA) has an area of $1,431\text{km}^2$, which is already substantially larger than Kigali and Dakar. The parts of the Greater Accra Region that are not included in the image dataset are the rural parts in the east of the city. Similarly, the satellite images of Dar es Salaam cover all of the



Figure 3.5: Purchased satellite basemap of Dar es Salaam, captured in 2021. The blue border represents the Dar es Salaam city boundary. Dynamic mosaic ©2021 Maxar Technologies.

metropolitan areas of the city, and only some rural parts in the south of the city are not included in the image dataset.

3.6 Data for cluster interpretation

3.6.1 Overview

I used various datasets on urban development for evaluating the analysis described in Chapter 4 to Chapter 7. All of these data are either directly visible in a satellite image or closely related to the visible built environment, such as population density. In the



Figure 3.6: Purchased satellite basemap of Kigali, captured in 2020. The blue border represents the Kigali city boundary. Dynamic mosaic ©2020 Maxar Technologies.

following section, I introduce the different datasets used and how they are pre-processed for the analyses.

Table 3.3 gives an overview of all external data sources used for the external cluster interpretation.

3.6.2 Data on the built and natural environment

In order to describe and analyse the results of the clustering analysis of the satellite images described in Chapters 4 to 7, I used environmental data capturing the built and natural environment. The environmental data allows for a detailed analysis of every study city on a small spatial scale. In the following section, I describe each environmental urban characteristic and finally summarise the main patterns of each city. Each urban characteristic is a feature of the urban environment that is either visible or directly related to visible features within the satellite images.

Urban characteristic	Unit	Data type	Source
Building area	m ²	vector	Maxar and Microsoft [8, 9]
Building count	integer	vector	Maxar and Microsoft [8, 9]
Mean building size	m ²	vector	Maxar and Microsoft [8, 9]
Mean building orientation	degrees (°)	vector	Maxar and Microsoft [8, 9]
Sum of major roads	m	vector	OSM [221]
Sum of all roads	m	vector	OSM [221]
Minimum distance from major road	m	vector	OSM [221]
Minimum distance from any road	m	vector	OSM [221]
Mean population density	estimated total number of people per grid-cell	raster (~100m/px)	WorldPop [222]
Mean NDVI	index (-1 to +1)	raster (~30m/px)	Landsat [223]

Table 3.3: Data on population, built and natural environment. NDVI: Normalised Difference vegetation index

Building data

As part of the built environment, buildings play an important role in the health, security and well-being of people. Adequate housing is a fundamental human right, yet major inequalities prevail in all study cities as described in Section 2.2.2. Information on building count and building area are indicators for building quality and living space. For example, informal settlements tend to have a high number of buildings within a small neighbourhood, whereas large buildings surrounded by gardens might be more indicative of an affluent area.

I used two different high-resolution spatial datasets of building footprints of cities in SSA to extract building area and count. I acquired building shapes of Accra that were from

2019 and 2020 and were derived from satellite images by Maxar/ECOPIA [8]. For the other study cities, I downloaded building information from Open Street Map (OSM), which is a collaborative dataset that is free and openly available [221]. These files include buildings mapped by Microsoft and Bing in the year 2021 with high-resolution satellite images by Maxar and Machine Learning (ML) techniques [221, 9]. The building information was published in vector format. I computed the building area and building count per index cell grid defined for the cluster analysis. Figure 3.7 shows the building area (m^2) and building count per index grid ($\sim 75 \times 75$ meters) plotted on the maps of Accra, Dakar, Dar es Salaam and Kigali.

All cities showed a high density of buildings within their metropolitan area. For the cities that are located at the ocean, Accra, Dakar and Dar es Salaam, there was a higher density of buildings along the coast, especially around the harbour areas. In all study cities, there are pockets of high building density that are scattered across the city boundary. In Kigali, the building footprints show the hilly nature of the city and the buildings are more likely to be built on the hillsides and the valleys. All cities also have neighbourhoods and areas with relatively low building area and count, which are usually located inland and at the city boundary.

I further calculated the mean building size and building orientation per building. The orientation of buildings holds significant implications for urban planning, notably influencing factors crucial to building functionality, including lighting and ventilation [224, 225, 226, 227]. It is noteworthy that building orientation can serve as a valuable indicator of historical artefacts and indigenous knowledge, impacting considerations such as vector reduction and solar control. Moreover, building orientation can extend beyond individual structures to encompass broader aspects of urban planning, including the orientation of roads. It is essential to highlight the intricate interplay between building orientation, road orientation and local context. This complexity underscores the potential impact on various facets, ranging from indoor lighting and ventilation to considerations for local wildlife

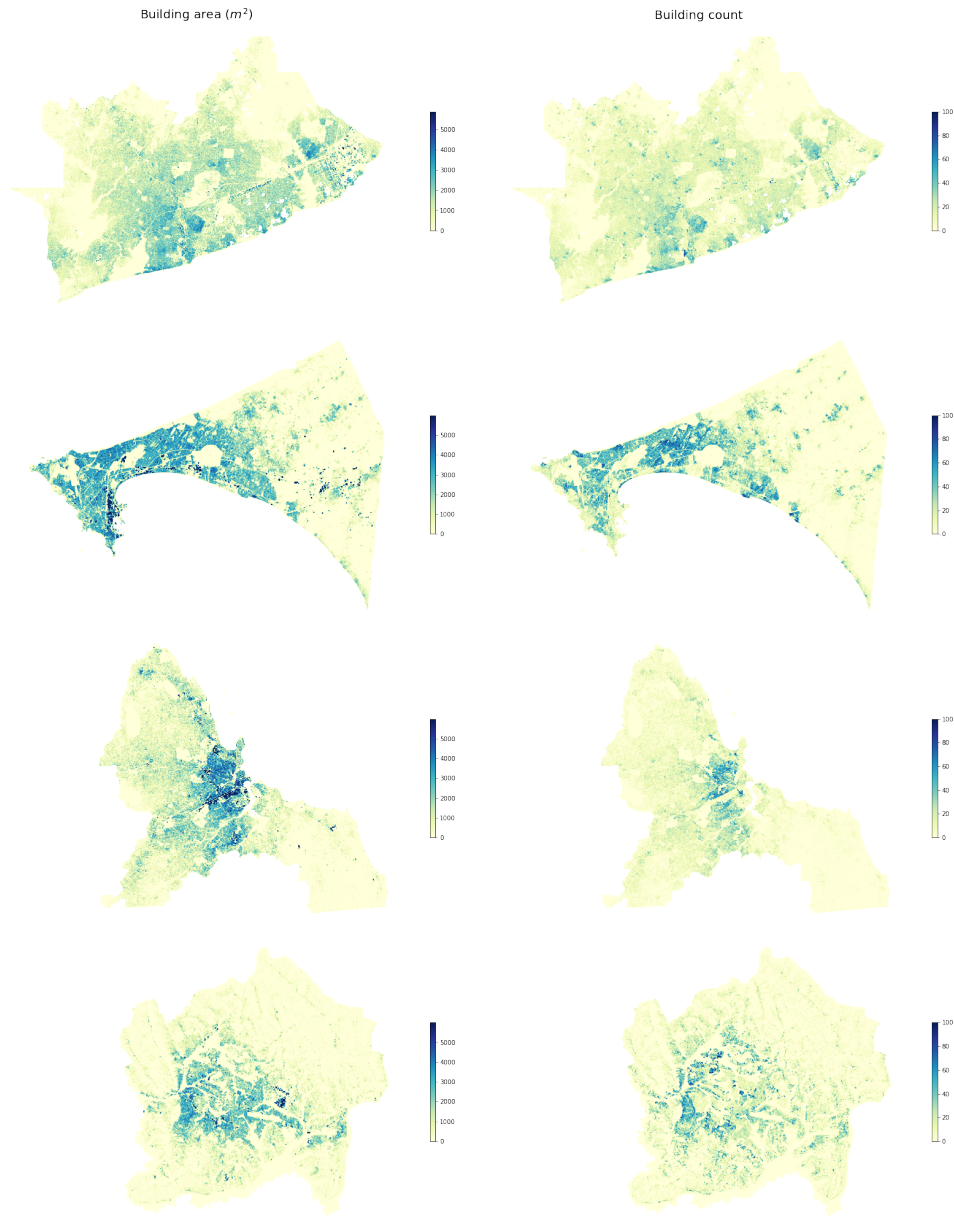


Figure 3.7: Building area (m^2) and building count in all study cities. From top to bottom: Accra, Dakar, Dar es Salaam and Kigali.

[228, 229, 230, 231]. Recognising these inter-dependencies is crucial in comprehending the multifaceted influence of building and road orientation on urban environments.

I calculated the building orientation from the building shapefiles for all cities with the *momepy* package [232]. Building orientation was computed as the deviation of orientation from cardinal directions and defined as an orientation of the longest axis of the bounding rectangle in the range 0-45 degrees.

Figure 3.8 shows the building shapes plotted on top of the satellite image. The buildings are shaded by building orientation deviation from cardinal directions. Black colours indicate high building orientation and white buildings indicate low building orientation.

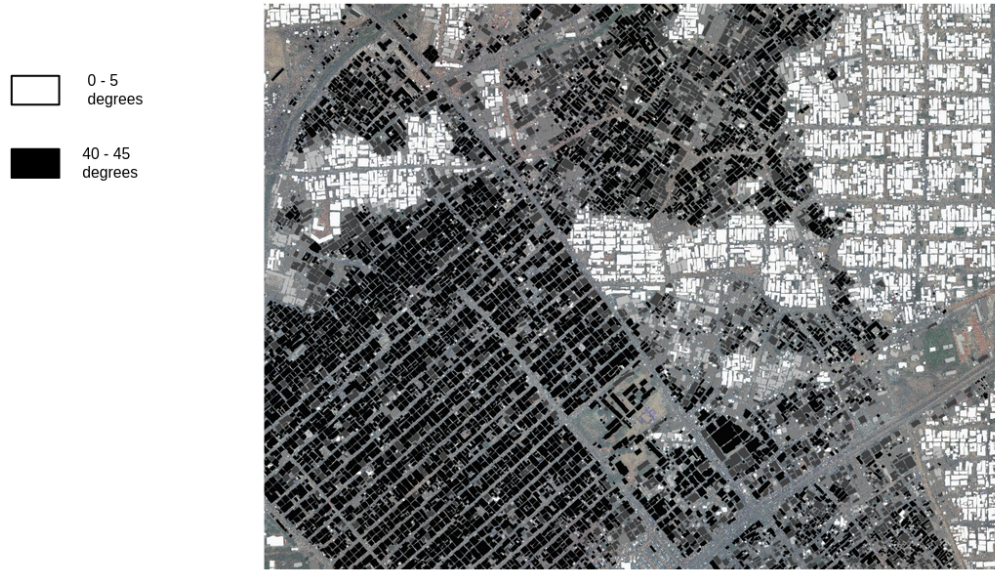


Figure 3.8: Difference in building orientation in degrees ($^{\circ}$) plotted on top of a satellite image of the neighbourhoods *Nima* and *Mamobi* in Accra, Ghana.

Figure 3.9 shows the average building area (m^2) and building orientation (0-45 degrees) per index grid ($\sim 75 \times 75$ meters) plotted on the maps of Accra, Dakar, Dar es Salaam and Kigali.

The average building size reflected the land use types across all cities. In all study cities, the parts that exhibit large average buildings are mainly located in industrial and commercial parts of the city. These are often also connected to harbours or main roads, such as Tema in Accra. In Dakar and Dar es Salaam the industrial and commercial neighbourhoods are both located around the harbour or an arterial road to one of the ports. In Kigali, this is the southern part of *Masoro* that is located next to the trans-African highway and the airport.

The building orientation varied largely between the study cities. Cities with histories of colonial city planning showed distinct patterns, such as in Accra and Dakar. As shown

in Figure 3.9, Dar es Salaam showed a few pockets of neighbourhoods with specific building orientations, however, not as distinct as in Accra and Dakar. The building orientation of buildings in Kigali, however, seemed to be mostly influenced by geography rather than specific city planning. The building orientation across buildings in Kigali looked therefore less structured.

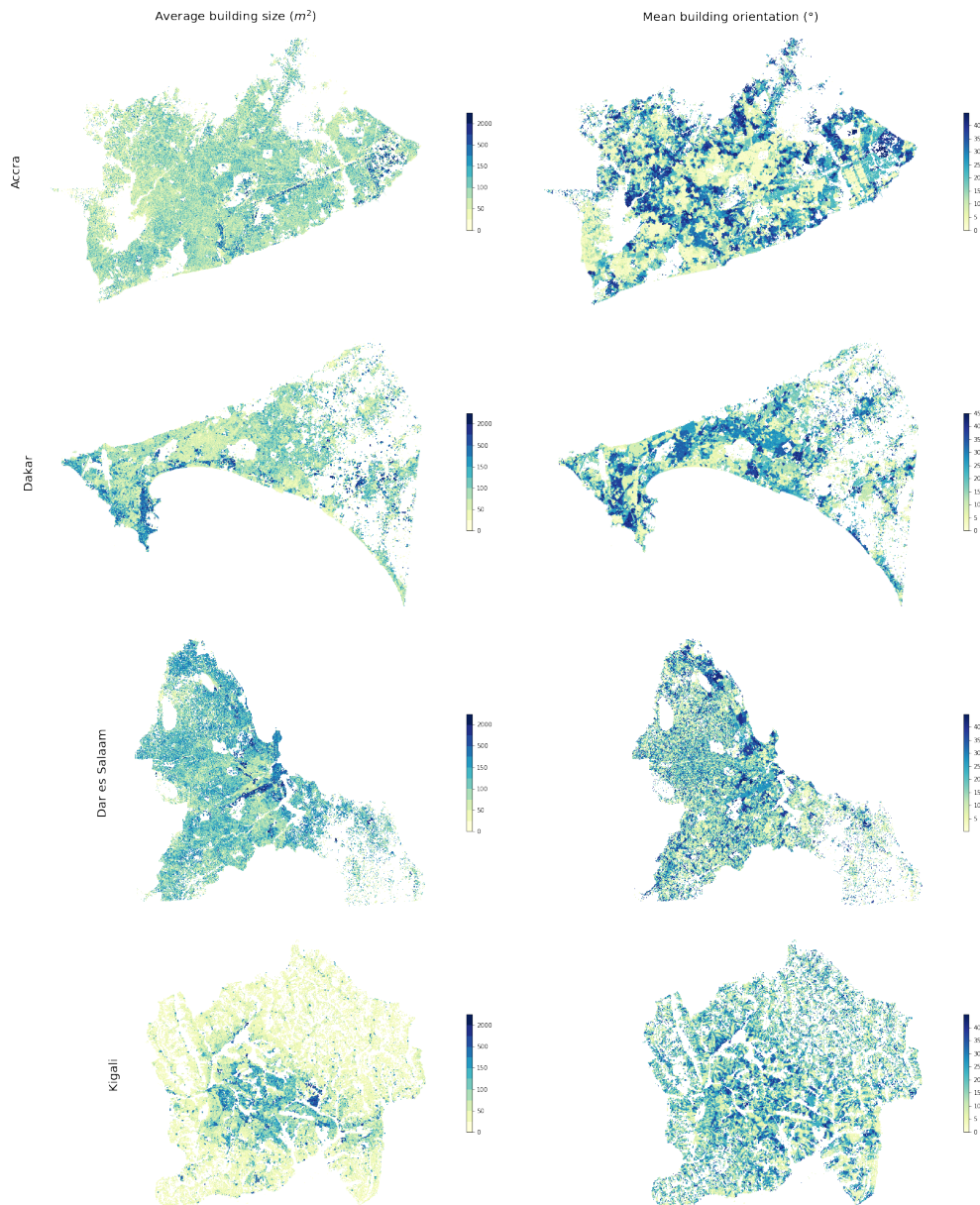


Figure 3.9: Average building size (m^2) and building orientation ($^\circ$) in all study cities. From top to bottom: Accra, Dakar, Dar es Salaam and Kigali.

Road data

Roads and access to roads are vital factors for determining connectivity within a city. The horizontal growth of cities in SSA has left a lot of peri-urban settlers disconnected from services and jobs in the city centre [4]. Additionally, urban spaces and how they are connected can further be indicators of informal settlements or slum areas, which exhibit distinct topological characteristics related to their distance to roads [233]. In order to analyse the clusters in regards to roads and distance to roads, I acquired data on roads and road networks. I downloaded road shapefiles from OSM, as vector shapefiles and at the closest date to image capture. I further classified the roads that were of the OSM category motorway, primary, secondary and tertiary as ‘major roads’. I also removed roads labelled as pedestrian, cycleway, and path from the dataset. Finally, I computed the road length, major road length, distance to all roads and distance to major roads per index cell grid defined for the cluster analysis. Figure 3.10 shows the length of all roads (m) and length of major roads (m) per index grid ($\sim 75 \times 75$ meters) plotted on the maps of Accra, Dakar, Dar es Salaam and Kigali.

I found varying road network patterns across the study cities as shown in Figure 3.10 and Figure 3.11. The road networks were largely influenced by geography, leaving gaps where there are bodies of water, national parks, or, in Kigali’s case, mountains.

In addition, the major road networks across all cities showed expected patterns, where all the roads converge at the central metropolitan area and then spread out can be compared to a spider web or a radial network. The layout is such that it forms a pattern reminiscent of spokes on a wheel, with the hub being the central city area. The roads represent the spokes, spanning from the hub to the city’s outskirts and beyond. The peri-urban areas along the city boundaries were furthest away from the main roads as depicted in Figure 3.11.

Figure 3.11 shows the distance to all roads and major roads per index grid ($\sim 75 \times 75$ meters) plotted on the maps of Accra, Dakar, Dar es Salaam and Kigali. Figure 3.11



Figure 3.10: Length of roads and major roads in all study cities. From top to bottom: Accra, Dakar, Dar es Salaam and Kigali.

shows similar patterns to Figure 3.10. The metropolitan areas in the city centre and next to industrial areas are closer to roads and especially major roads. The peri-urban areas, which are close to the city boundary, are furthest away from major roads and in general not well connected to jobs and services.

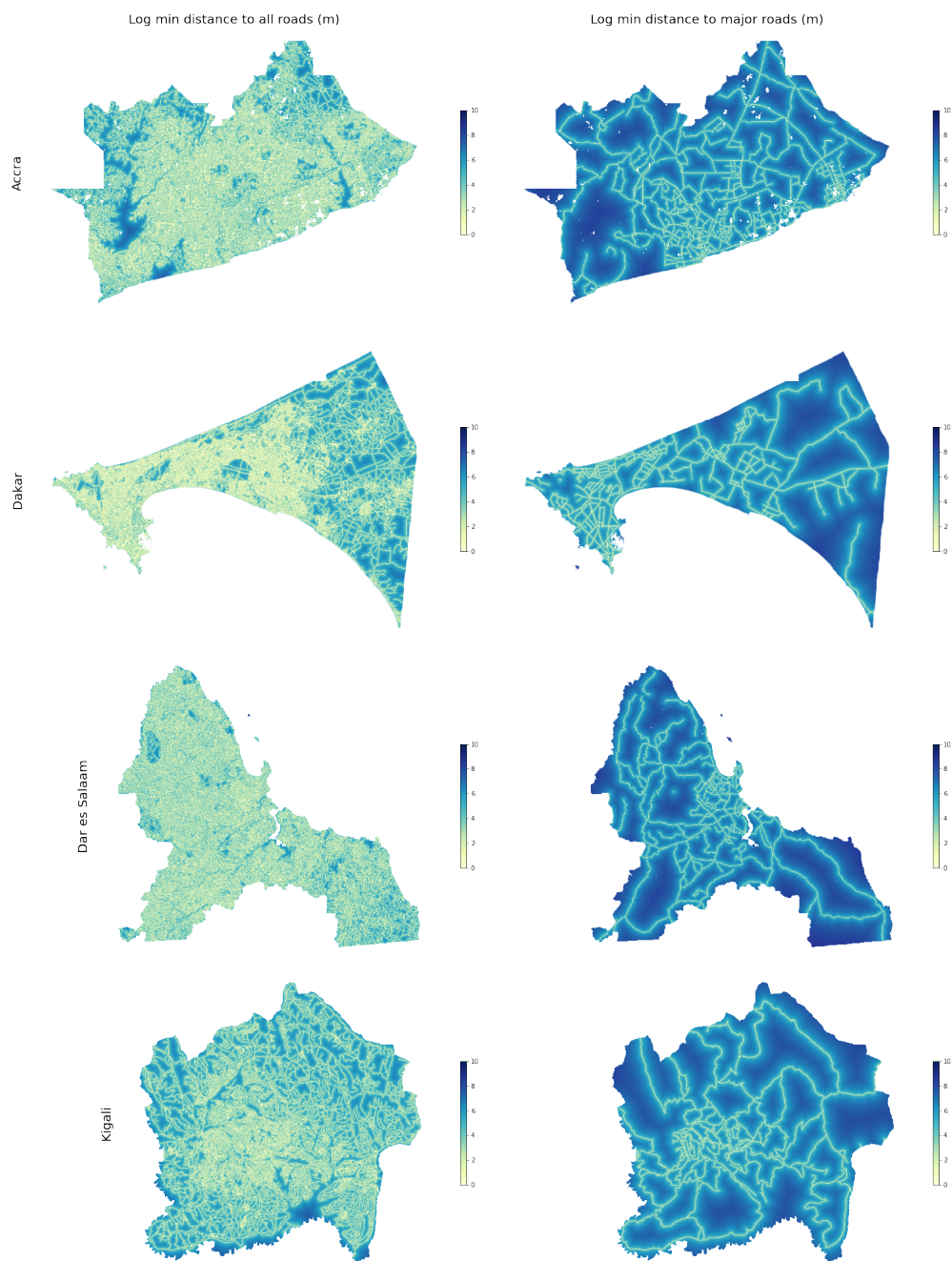


Figure 3.11: Distance to all roads (m) and major roads (m) in all study cities. From top to bottom: Accra, Dakar, Dar es Salaam and Kigali.

Population density

Population density is not directly visible on satellite images, however, has a multitude of visual indicators on remote sensing images, such as buildings, and other features of the built environment. I, therefore, included data on population density to datasets for describing the cities and interpreting the analysis described in Chapters 4 to 7. I acquired the population density raster files through *WorldPop* at a spatial resolution of 100 meters per pixel and are available for all four countries [222]. I chose the population rasters for the closest date to satellite image capture. As part of the data pre-processing, I converted the raster files into vector shapefiles. I finally computed the mean population density per index cell grid defined for the analysis.

The population density patterns across the study cities show, as expected, the highest population density in the metropolitan centre of the city. Accra and Dakar both have relatively large areas with high population density. The city area of Accra and Dakar is also a lot smaller in terms of area compared to the other study cities. Dar es Salaam and Kigali both have large peri-urban areas with patches of agriculture and sparsely populated neighbourhoods.

Normalised Difference Vegetation Index (NDVI)

The NDVI is a widely used index to estimate vegetation levels in remote sensing contexts. As described in Section 2.2.1, blue/green space impact health in multiple ways. Living in the vicinity of green/blue spaces affects the local noise and air pollution as well as the physical and mental health of city dwellers.

The NDVI values range from -1 to 1 and can give a good intuition for whether water, built environment, or vegetation is present. It is defined as follows.

$$\text{NDVI} = \frac{\text{NIR} - \text{RED}}{\text{NIR} + \text{RED}} \quad (1)$$

where NIR represents the reflectance in the near-infrared part of the spectrum, and RED represents the reflectance in the red part of the spectrum captured by a satellite sensor. The NDVI of water is usually below zero, and the index, as its name suggests, is very good at picking up vegetation ($\text{NDVI} > 0.3$), but also the built environment ($\text{NDVI} \sim 0.1$).

I computed the NDVI for cluster interpretation from Landsat 8 images, which can be downloaded from the USGS Earth Explorer [223]. The Landsat images used for calculating the NDVI were downloaded based on the date closest to satellite image capture and a cloud cover of under 5%. As part of the pre-processing, I converted the NDVI raster files into vector shapefiles and overlayed them with the index grid used for cutting the satellite image into smaller tiles to obtain the NDVI per index grid.

Figure 3.12 shows the NDVI across the study cities. Kigali and Dar es Salaam have the highest levels of vegetation. In contrast to Dakar and Accra, which have proportionally larger parts with impervious surfaces and built-up areas. The NDVI further clearly highlights water bodies, such as the *Weija* reservoir in Accra or the *Retba Lake* in Dakar. Additionally, the climate across the study cities varies widely which also influenced the NDVI values. Kigali and Dar es Salaam are located in tropical and humid climate zones, meaning they are naturally a lot greener compared to Accra and especially Dakar. Dakar's climate can be described as tropical arid with lots of sandy and dry land.

Figures 3.13 to 3.16 show the correlation between all urban characteristics. The figures show that for all cities building information is mainly positively correlated with other building information as well as lengths of roads and population density. NDVI is negatively correlated with building area, building count and population density.

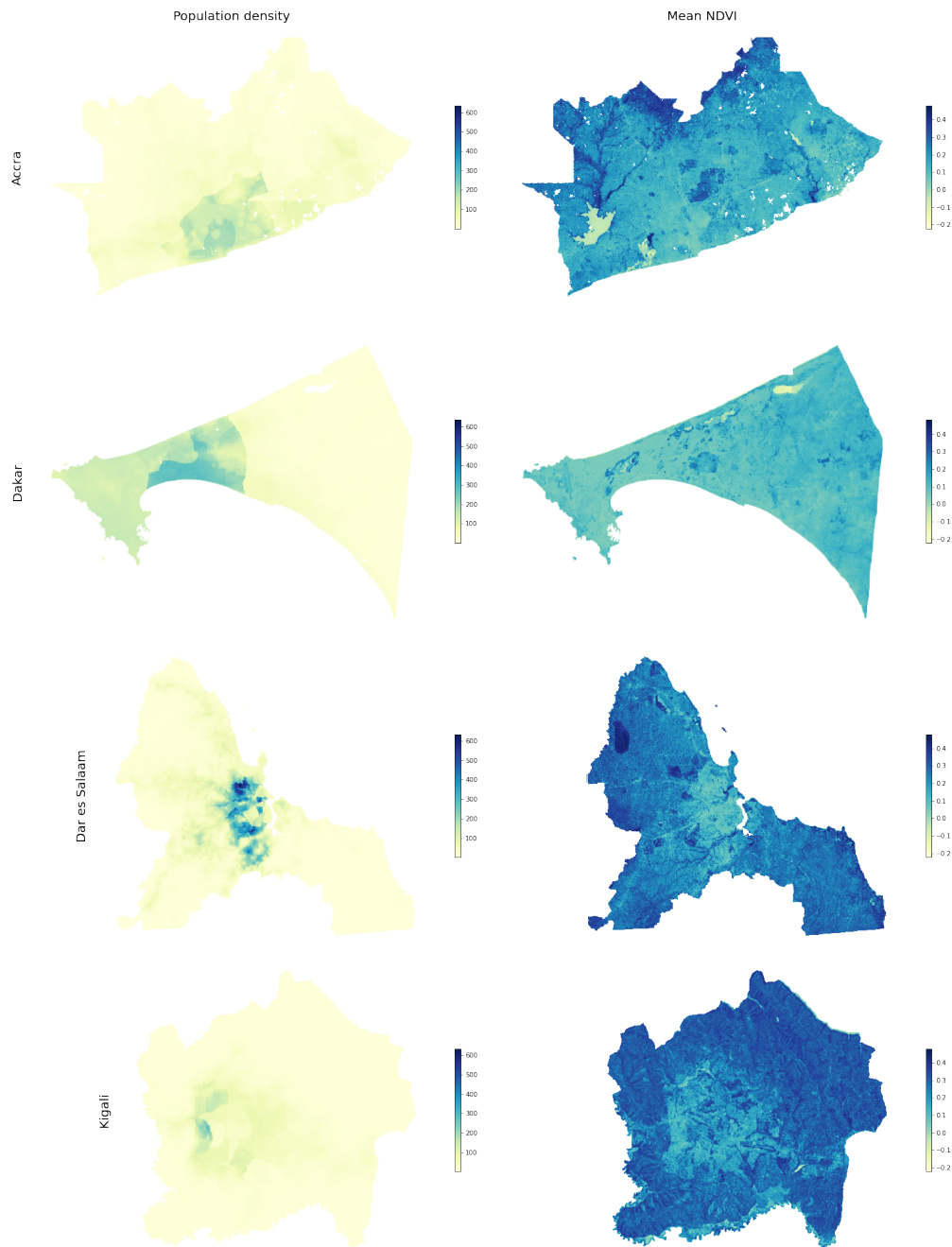


Figure 3.12: Population density and NDVI of all study cities. From top to bottom: Accra, Dakar, Dar es Salaam and Kigali.

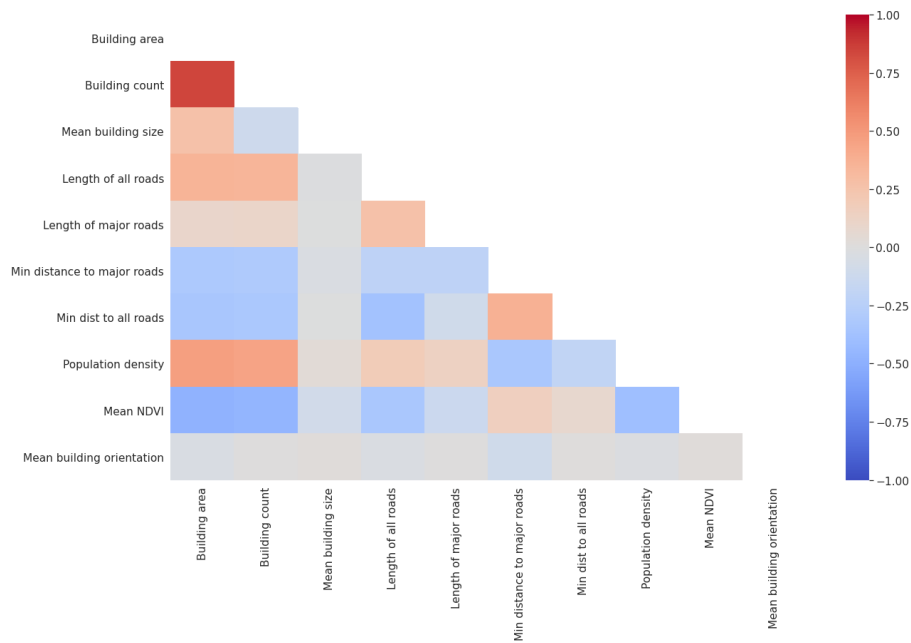


Figure 3.13: Correlation of environmental and demographic variables of Accra. Red colours indicate a positive correlation and blue colours indicate a negative correlation.

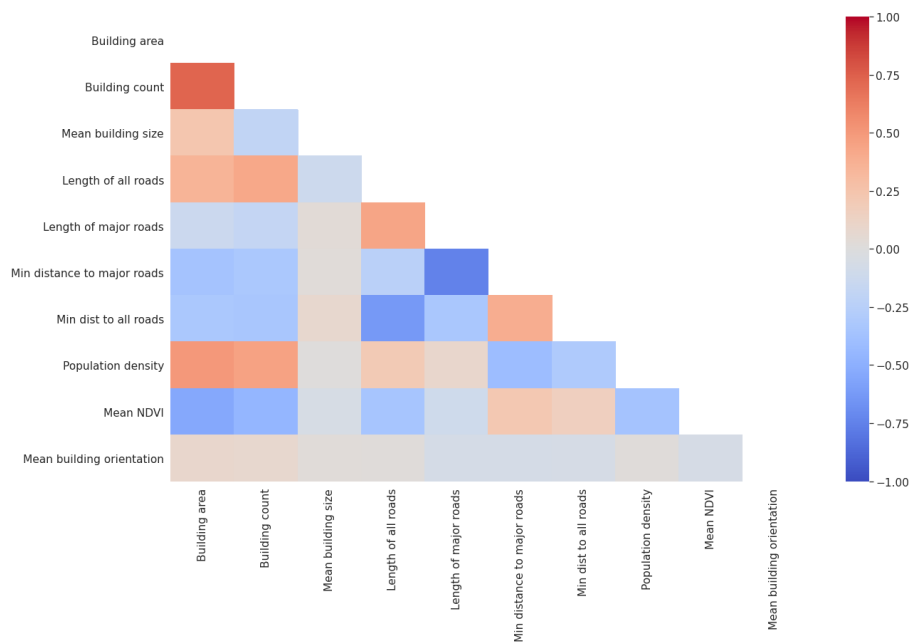


Figure 3.14: Correlation of environmental and demographic variables of Dakar. Red colours indicate a positive correlation and blue colours indicate a negative correlation.

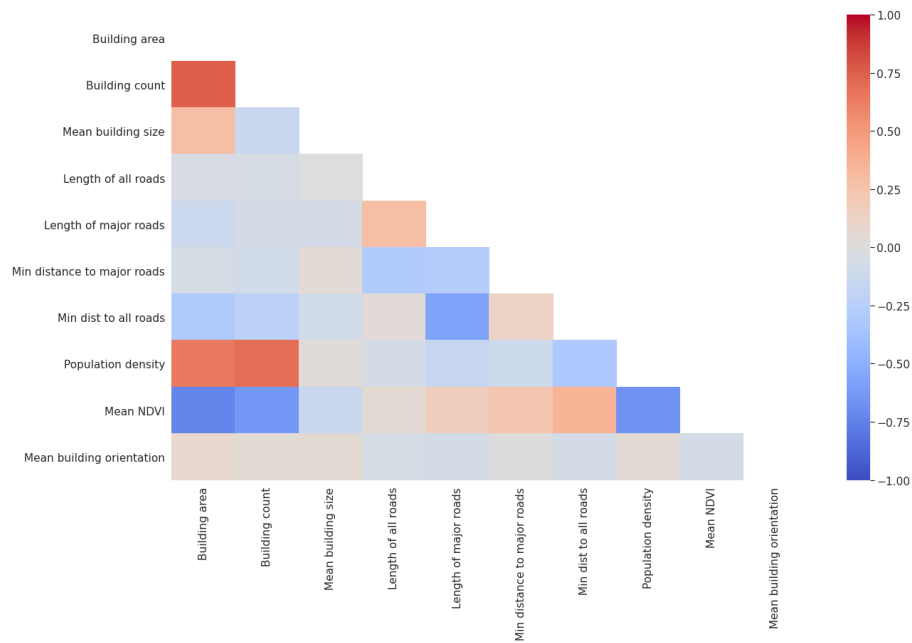


Figure 3.15: Correlation of environmental and demographic variables of Dar es Salaam. Red colours indicate a positive correlation and blue colours indicate a negative correlation.

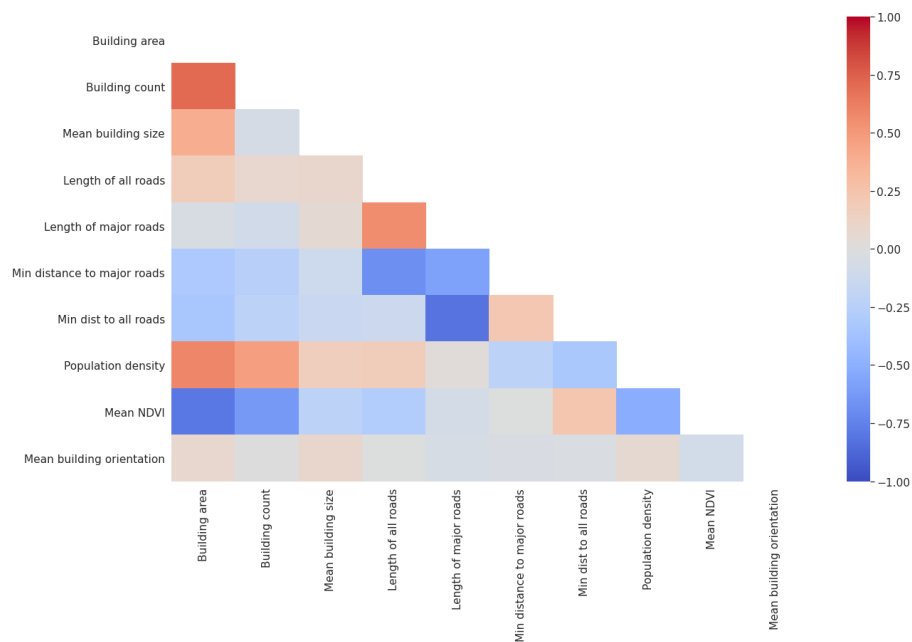


Figure 3.16: Correlation of environmental and demographic variables of Kigali. Red colours indicate a positive correlation and blue colours indicate a negative correlation.

European Space Agency (ESA) land cover data

Before 2020, global land use/land cover maps usually were captured at rather low resolutions (30m - 1000 m). In 2022, ESA released land use/land cover maps at 10 m resolution for 2020 and 2021 [234]. The overall accuracy of the 2021 maps is higher, for that reason, I only use the 2021 land cover maps for the interpretation of the clusters.

The land use/land cover maps are provided free of charge, without the restriction of use, and can be downloaded as raster files through Google Earth Engine App or ESA's own WorldCover viewer [235]. The dataset includes 11 different classes, which include: tree cover, shrubland, grassland, cropland, built-up, bare/sparse vegetation, snow and ice, permanent water bodies, herbaceous wetland, mangroves, moss and lichen. As part of the pre-processing, I converted the land cover raster files into vector shapefiles and overlaid them with the index grid used for cutting the satellite image into smaller tiles to obtain the land cover classes per index grid.

3.6.3 Summary

Four cities in SSA and respective city boundaries were chosen for analysis, namely Accra, Dakar, Dar es Salaam and Kigali. All cities were introduced with short city profiles. Furthermore, satellite images were acquired through an Open Data initiative as well as bought from Maxar Technologies. Data lessons learned about satellite image pre-processing are further outlined in this chapter. Finally, data capturing the built and natural environment, as well as population data and land cover maps were collated for all cities. This data included information on buildings and roads as shapefiles, and information on vegetation through NDVI measures.

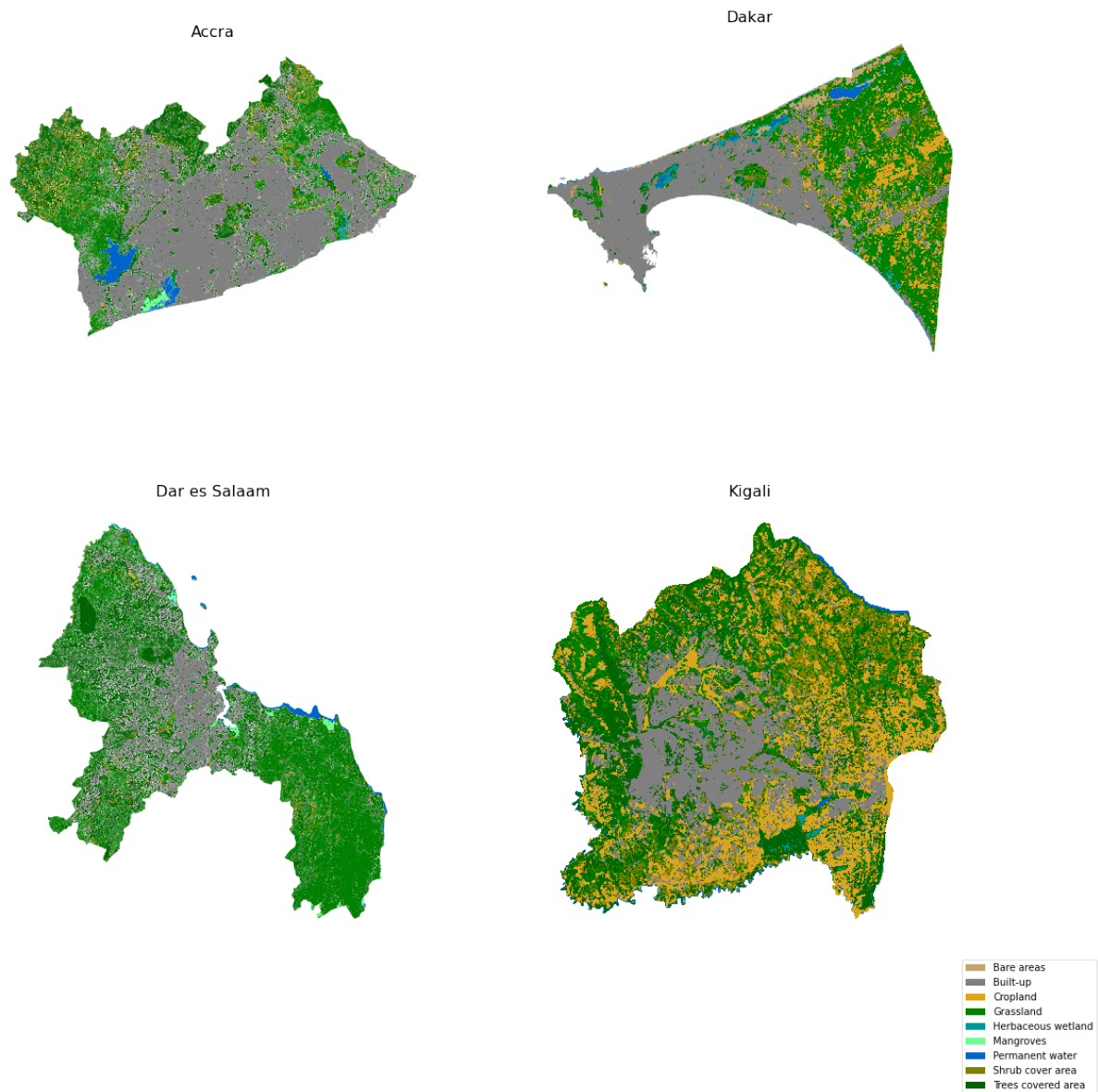


Figure 3.17: ESA land cover map of Accra, Dakar, Dar es Salaam and Kigali. 10m resolution.

Chapter 4

Clustering one city: Accra

This chapter is based on the peer-reviewed publication titled "Phenotyping urban built and natural environments with high-resolution satellite images and unsupervised deep learning" [1], published in The Science of the Total Environment (STOTEN) journal, for which I was first author (DOI: <https://doi.org/10.1016/j.scitotenv.2023.164794>). I hereby give credit to this work without citing the work in every instance in the text.

4.1 Overview and objectives

In this chapter, I describe a combined feature extraction and clustering framework applied to high-resolution satellite images of Accra, Ghana. First, I state the objectives, describe the method and explain how the cluster results are evaluated internally and externally with data on the built and natural environment. Finally, I discuss cluster results in the current policy and health context of the city.

The main objectives of this chapter is to apply unsupervised clustering to high-resolution satellite images of Accra, Ghana and analysing the detected clusters with external demographic and environmental variables. Secondly, it aims to deepen our understanding by

exploring the relationship between these clusters and the external variables introduced in 3. This is achieved through the application of a machine learning classifier, predicting cluster membership and enhancing interpretability with SHapley Additive exPlanations (SHAP) values. Additionally, the chapter contributes by evaluating the clusters internally, analysing cluster assignments in the feature space, and visualising filters and Class Activation Maps (CAMs).

4.2 Methods

4.2.1 Data sources and pre-processing

I used high-resolution satellite image of Accra that covers most of the GAMA area as pictured in Figure 3.1. The spatial resolution of the satellite image amounts to 0.3 meters per pixel and captures three different bands (RGB). The basemap image was captured in 2019 by the WorldView 2 satellite and acquired in 2020 from Maxar Technologies. The pre-processing of the image included colour-balancing and orthorectification as described in Section 3.5.1. Additionally, I used data of the built and natural environment to interpret the clusters. The data and data pre-processing are outlined in Section 3.6.

Tile size

The analysis of high-resolution satellite images has undergone a significant shift in recent years, as mentioned in Section 2.6.3. Previously, the analysis was primarily conducted at the pixel level, considering individual pixels as the basic unit of analysis. However, recent approaches moved towards object or patch-oriented analysis. The grouping of multiple pixels into relatively homogeneous image objects or tiles has become a common practice [176]. This involves identifying and delineating distinct features within the images, such as houses or roads, through object-targeted analysis. By identifying these features, it

becomes possible to study and understand specific objects or regions of interest within the satellite images. Another approach to grouping pixels is by dividing the large satellite images into smaller image tiles or patches. Given the typically large size of satellite images, which can be on the order of 10,000 x 10,000 pixels, tiling or patching is a practical method for further analysis or potential annotation [236]. By breaking down the full scene images into smaller tiles, it becomes easier to process by a CNN and analyse specific sections of the image individually.

There are a number of factors that have to be considered when choosing an appropriate tile size. Dumitru et al. [237] suggest to not only regard typical land cover characteristics of the area but also the capturing device and the resulting pixel size and radiometric image quality. The objective of identifying an appropriate tile size (or range of sizes) is therefore influenced by the definition this research takes of a ‘relatively homogeneous’ image grouping. I defined the following benchmarks to choose a suitable tile size:

- A tile should group an amount of pixels together such that it is able to capture (parts of) common urban features such as buildings and roads.
- A tile should be suitable for CNN analysis with existing network architectures or pre-trained networks on common datasets, such as the ImageNet dataset [238].
- A tile should have a square shape to facilitate subsequent image and deep learning analyses [239, 240].

Furthermore, the authors of the software library GDAL recommend using common image sizes such as 128, 256, 512, or 1024 pixels in width and height [241].

For the main analysis described in this Chapter (Chapter 4), I thus chose the tile size of 256 x 256 pixels, which is commonly used in remotely sensed data with similar resolution. It grants effortless processing and summarises tile information at a neighbourhood level. Chapter 5 further discusses alternative tile sizes and how the varying tile size impacts the analysis.

As shown by Figure 4.1, I clipped the study area into tiles that measure 256 x 256 pixels which represents about 75 x 75 meters on the ground after removing the part of the satellite image that was not covering the geographical city boundary.

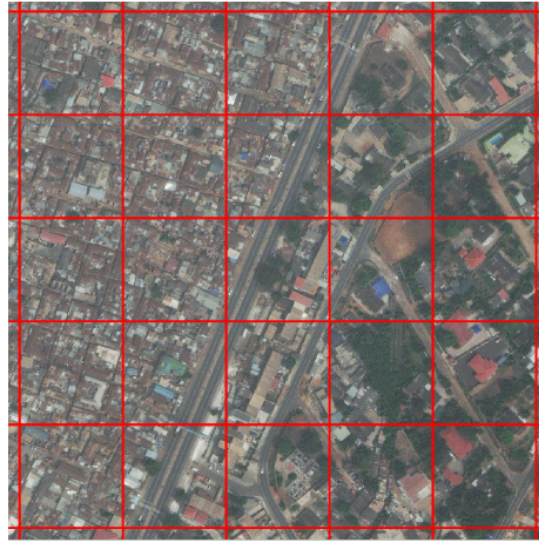


Figure 4.1: The figure shows a fragment of the satellite basemap demonstrating the heterogeneity of the city, with informal settlements to the left of the road and large houses to the right (*Nima* bordering the *Cantonments*). The red squares represent the tile size of 256 x 256 pixels. The sensitivity analysis that led to the selection of this tile size is presented in Chapter 5.

I further removed cloudy or black image tiles. The tiles that captured clouds were identified by computing the median tile colour. Any tiles that had median Red, Green, Blue (RGB) tile colour values above (210, 210, 210) and below (15,15,15) were removed from the dataset, with examples shown in Figure 4.2. The empty and therefore black tiles were all located at the border of the image where the basemap could not be cut into perfect squares. Additional image normalisation was not required as it is part of the satellite basemap processing by Maxar. The 85.4 GB satellite basemap created 321,820 tiles, with 222,240 tiles remaining after the cleaning and pre-processing steps.



Figure 4.2: Examples of tiles that were excluded from the dataset. Cloudy white tile (left) and empty/black tile (right).

Additionally, I used data on the natural and built environment to interpret the clusters, which is described in detail in the Data Chapter 3, Section 3.6.2. Table 3.3 gives an overview of all external data sources used for the external cluster interpretation. To calculate the per tile statistics, I intersected each vector and raster dataset with the 256 x 256 index grid of the satellite image tiles.

4.2.2 Analysis framework

The analysis plan, depicted in Figure 4.3, includes data pre-processing and analysis steps. The image and environmental data are described in Chapter 3. Combined feature extraction and clustering was performed using a method discussed in Section 4.2.3, followed by internal and external analysis of the resulting clusters. External analysis interprets clusters using data not used for clustering. This includes fitting an ensemble-based classifier to predict cluster membership and conducting SHAP analysis to identify defining environmental and demographic characteristics (Section 4.2.5). Internal analysis focuses on feature space analysis (Section 4.2.6), examining relationships among the image representations used in clustering.

Overall, the analysis plan integrates various steps, including data pre-processing, clustering, external analysis for cluster interpretation, ensemble-based classification, SHAP analysis, and internal feature space analysis. These steps collectively provide a comprehensive understanding of the clusters and the factors influencing their formation.

4.2.3 DeepCluster: Combined feature extraction and clustering

I used DeepCluster [242], an end-to-end feature extraction and clustering framework to assign each image tile to a cluster. In the method, a CNN is used as a feature extractor. Each tile is fed into the CNN to create image representations. Here, the tile size does not only relate to the spatial level of analysis (index), but also directly impacts the network's ability to capture intricate patterns and spatial dependencies within the input data, thus influencing the overall effectiveness of the model. The extracted image representations of each 256 x 256 pixel satellite tile are reduced from 4096 to 256 dimensions using principal component analysis and normalised using the Euclidean norm (ℓ_2 -normalisation). The transformed features are then passed to a k-means clustering algorithm (as defined by Equation 4.2). After clustering, each image representation is assigned a cluster membership which is then used as a pseudo-label to update the weights of the CNN classifier (see Equation 4.1). The algorithm iteratively minimises clustering and classification loss, creating new pseudo-labels after each training epoch. The method combines feature extraction and clustering in an end-to-end approach, meaning the cluster formation relies entirely on the data and is not influenced by the availability versus sparsity of labelled data.

In Caron's [242] notation, f_θ is the convolutional network mapping, where θ is the set of corresponding parameters. The vector resulting from applying this mapping to an image is referred to as a feature or image representation. Given a training set $X = x_1, x_2, \dots, x_N$ comprising N images, the aim is to find an optimal parameter θ such that the mapping f_θ generates effective, general-purpose features. Conventionally, these parameters are learned with supervision. That is, each image x_n is paired with a label y_n from $0, 1^k$. This label signifies the image's affiliation with one of the k possible pre-defined classes. A parameterised classifier g_W predicts the appropriate labels based on the features $f_\theta(x_n)$. The parameters W of the classifier and the parameter θ of the mapping are learned jointly.

$$\min_{\theta, W} \frac{1}{N} \sum_{n=1}^N \ell(g_W(f_\theta(x_n)), y_n), \quad (4.1)$$

where ℓ is the multinomial logistic loss, also called negative log-softmax function. The optimisation of this cost function is achieved through the utilisation of mini-batch stochastic gradient descent [243], with the gradient computation being performed via backpropagation [244].

The k-means algorithm takes a set of vectors as input. In this case, it is the features $f_\theta(x_n)$ generated by the CNN. These features are grouped into k distinct clusters based on a geometric criterion. In more specific terms, the model concurrently determines a centroid matrix C of dimensions $d \times k$ and assigns clusters y_n to each image n by addressing the subsequent optimisation problem:

$$\min_{C \in \mathbb{R}^{d \times k}} \frac{1}{N} \sum_{n=1}^N \min_{y_n \in \{0,1\}^k} \|f_\theta(x_n) - Cy_n\|_2^2 \text{ such that } y_n^T \mathbf{1}_k = 1. \quad (4.2)$$

Solving this problem provides a set of optimal assignments $\{y_n^*\}_{n \leq N}$ and a centroid matrix C^* . These assignments are then used as pseudo-labels; the centroid matrix is not made use of [242].

I used the common CNN architecture VGG-16 [240] with batch-normalisation and initialisation with pre-trained ImageNet weights for feature extraction. CNNs trained on the ImageNet dataset are suitable feature extractors that can recognise low-level and mid-level features such as corners and edges which are salient for identifying many features of interest within images and have been shown in a variety of cases to generalise well to other tasks [245, 246]. The images are, therefore, normalised using the mean and standard deviation of the ImageNet dataset before feeding them into the CNN. Most hyperparameters were kept at the same values as in the original DeepCluster paper, except for the number

of clusters (which is discussed in Chapter 5 in more detail), learning rate and the number of epochs, which were set to 0.0001 and 20 respectively based on an initial set of experiments. The algorithm formed the most intuitive and interpretable clusters at a learning rate of 0.0001, compared to a set of alternative learning rates (0.1, 0.01, 0.001, 0.00001 and 0.000001). Learning rates >0.0001 picked up structure in the city-wide satellite image that resulted from how the city-wide image was stitched together (from multiple satellite images) as part of the commercial pre-processing and was unrelated to the content of the image. The training time for around 20 epochs is about 24 hours and training it for another 24 hours (~ 50 epochs) did not change the clusters and their interpretation substantially.

Many cities in the SSA, including Accra, have fragmented spatial structures with pockets of slums neighbouring high-rise business buildings and scattered fringe developments [50]. To avoid smoothing over heterogeneities that can result in missing the full complexity of an urban system, I analysed the tiles independent of their proximity to neighbouring image tiles, i.e. the tiles fed into the neural network thus contain no information about neighbouring tiles or geographic information, e.g., latitude and longitude.

Network weights

As described in Section 4.2.3, I used ImageNet weights to initialise the training of the DeepCluster algorithm. I considered using weights trained specifically on satellite image, however, initial transfer learning approaches did not yield satisfactory results. There is a wide range of satellite images available, however, most of the openly available datasets did not match the high resolution of the images used in this thesis or did not transfer well to urban areas on the African continent. A table listing the openly available and labelled datasets I considered can be found in the Appendix (Table A.3).

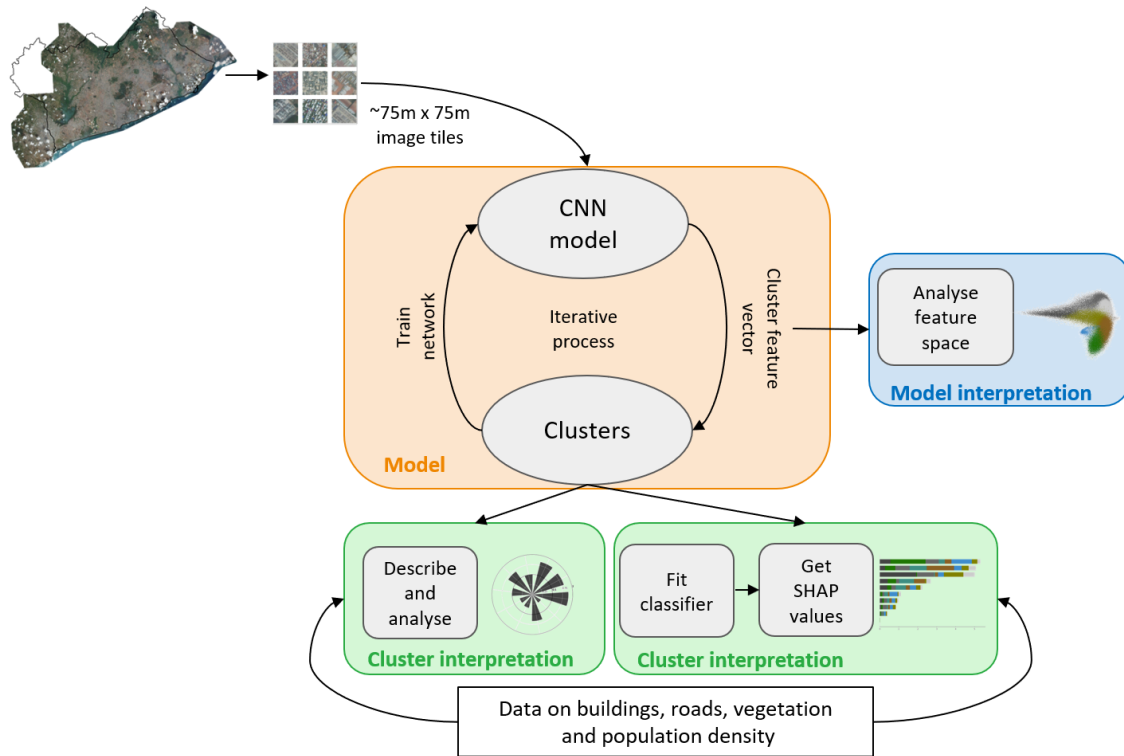


Figure 4.3: In the pre-processing step, the city-wide satellite image was cropped into tiles. Tiles, which were not within the Greater Accra Metropolitan Area (GAMA) or contained clouds were removed. The image tiles were then fed into the end-to-end DeepCluster algorithm which is described in Section 4.2.3. The resultant clusters were described and analysed with external data on buildings, roads, population density, water, and vegetation. After clustering, I fitted a classifier to predict cluster membership to understand which environmental variables were most important for cluster formation, measured by their corresponding SHAP values. To understand how the model learns the image representations, I analysed the clusters in the low-dimensional feature space.

Clustering algorithm

DeepCluster consists of combined feature learning with a CNN and a clustering algorithm. K-means is the baseline clustering algorithm used by Caron et al. [242], and is a popular choice because of its simplicity and scalability. It is an iterative algorithm that uses vector quantisation to partition a dataset into a defined number of clusters [247]. An alternative clustering algorithm, which I have considered using is the Power Iteration Clustering (PIC) algorithm. The PIC is a very scalable graph clustering approach, which finds a low dimensional representation of the dataset and uses truncated power iteration on a pair-wise similarity matrix of the data [248]. The final clusters are then extracted

from the 1D representation. The clustering method is very suitable for large datasets, as it can be parallelised, and also does not require the selection of the number of clusters. However, image representations extracted from CNNs are usually clustered very densely in the feature space and density-based or spectral clustering algorithms tend to either choose very high cluster numbers or a very low number. The PIC algorithm picked over 3,000 clusters when run with DeepCluster and I therefore continued the analysis with the K-means algorithm, which allowed me to choose a cluster number.

4.2.4 Cluster interpretation

Clustering, a machine learning algorithm that does not require any training labels, inherently lacks an absolute ground truth against which to validate its results. The assessment of clustering effectiveness generally falls into two categories: internal and external indices. Internal evaluation indices are based on the intrinsic information in the data and look at the clustering without external information. In a deep neural network context, this includes inspecting the model by looking at its training behaviour and/or creating visualisations to understand how the network is functioning (e.g. deep visualisations or class activation maps (CAMs)) as described in Section 4.2.6. Additionally, external indices allow for the comparison of clusters with independent information and can be used to test hypotheses created based on the clusters identified, as outlined in Section 4.2.5.

4.2.5 External cluster analysis

To interpret the clusters, I used data capturing information on urban form, population density and vegetation described in Chapter 3. I fit an ensemble tree classifier to predict cluster membership and obtain SHAP analysis values, and I further sought input from research colleagues from the *Pathways to Equitable Healthy Cities* Group at Ghana University to add a local perspective to cluster interpretation.

Built and natural environment and demographic characteristics of clusters

I used the data on the built environment, water, vegetation, and population to quantify the characteristics of the image-based clusters and report the median values of each measure for all the tiles that fall in each cluster and visualise them in form of radar charts (Figure 4.6). Each environmental and demographic variable was scaled with a quantile transformer, a non-parametric transformation to map the data to a uniform distribution with values between 0 and 1 (0.5 indicates the median value of a variable across all tiles in the entire image). Tiles with no buildings were included (as zeros) in summary statistics for building count and building area but excluded from the calculation of summary statistics for average building size and orientation so that zero is not used in the denominator. The length of major roads was zero in 93% of all tiles, hence the median of all tiles was zero as was the mean in most clusters.

Built and natural environment and demographic characteristics and predictors of clusters

I used the data on built environment, water, vegetation, and population to quantify the characteristics of the clusters that were formed based on image data alone. I report the median values of each measure for all the tiles that fall in each cluster. Additionally, I used the machine learning classifier XGBoost [249] to quantify which environmental and demographic characteristics, individually and collectively, characterise the image-based clusters. This decision-tree-based method identifies which environmental and demographic variables predict cluster membership. It has the practical advantage of being able to accommodate missing values [250], such as for tiles that have no buildings and therefore no average building size and orientation. To measure which environmental and demographic variables are important for predicting image tiles' membership to different clusters, I used the fitted classifier to generate SHAP values [251]. The SHAP values are summary measures of the importance of each environmental and demographic variable

for each cluster as well as across all clusters, in an additive manner.

I split the dataset (tiles) into 80 per cent training and 20 per cent testing data. I used a stratified approach for splitting the tiles to ensure that all clusters were equally present in the evaluation. The gradient boosting classifier, XGBoost, was fine-tuned with a 5-fold cross-validation method with accuracy as a score. I used the Hyperopt [252] library, which uses Bayesian optimisation for parameter tuning to find the optimal hyperparameters. The hyperparameters selected by the optimisation are listed in the Appendix Section A.4. The final accuracy of the classifier was 0.66, scoring 54% higher than random assignment of a given image tile to a cluster. This classification accuracy score, which is only moderate, reflects that the environmental and demographic variables that were used in the SHAP analysis are only a subset of those that have visual signals. Other visual signals, on which I did not have geocoded data, may include vehicles, building type and material, specific vegetation categories, and types of terrain [253].

Input from researchers with deep knowledge of Accra

In addition to the quantitative analysis, I sought input from research colleagues from the *Pathways to Equitable Healthy Cities* Group at Ghana University for cluster interpretation. To facilitate feedback collection, I created a webpage, where participants could reply directly to a set of specific questions. I was especially interested in gaining insight into a set of questions regarding cluster emergence. These questions were mainly related to geography, human action (e.g. areas of trade), or the history of the city, as well as implications for urban citizens living in a certain area. The website can be accessed under the following link: <https://barbarametzler.github.io/Clusters-of-Accra/>

Figure 4.4 shows a section of the webpage create to get feedback and input from local researchers at Ghana University. The page includes an interactive visualisation of clusters on top of Google Maps to help visualise clusters and their location. I used the replies of

Densely populated areas

Cluster specific questions:

- Are there any differences between the two dark grey clusters?
- What are the implications of living in these areas?

Why do these clusters emerge?

What do they mean for humans living in these areas?

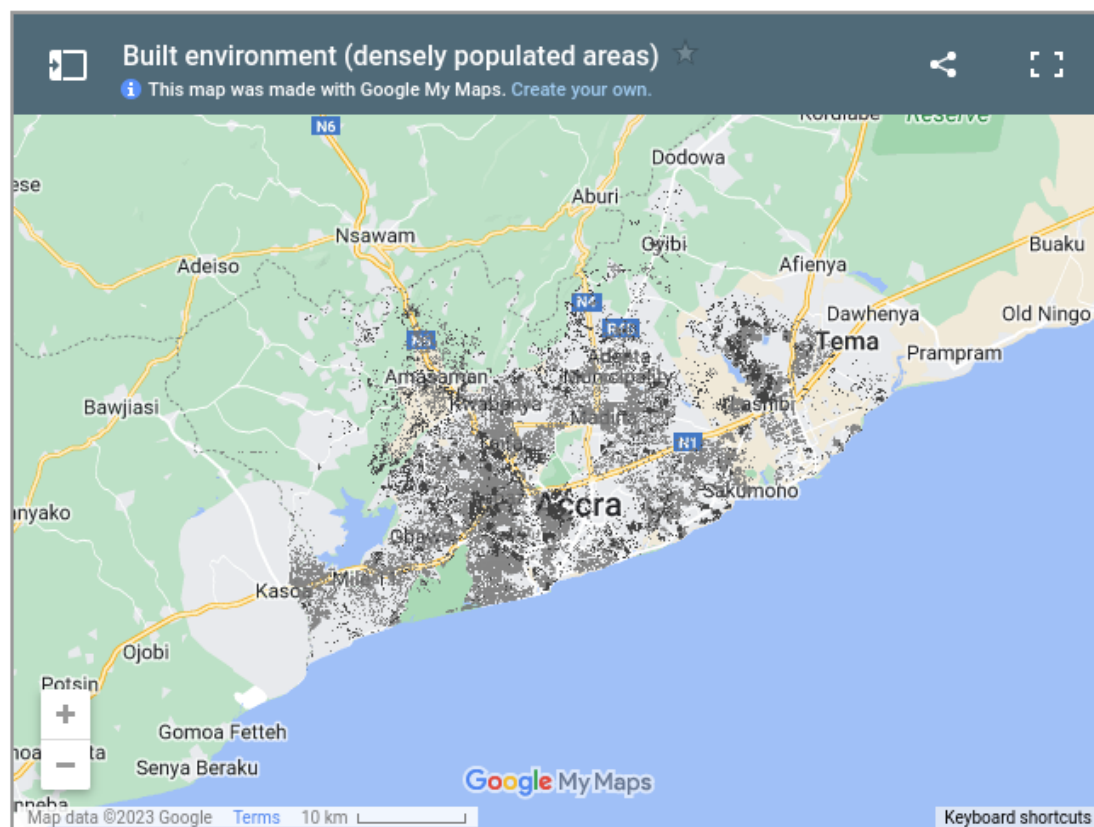


Figure 4.4: Screenshot of a section of the webpage with cluster map and questions. Wepage is accessible via <https://barbarametzler.github.io/Clusters-of-Accra/>

the researchers to add context to the quantitative interpretation of the clusters with data on built and natural environment. Their input added to the interpretation of the clusters and their implications for tracking and enhancing sustainable urban development, which is described in Section 4.4.1.

4.2.6 Internal evaluation of clusters

External evaluation of clustering necessitates additional data to interpret the clusters, whereas internal evaluation, uses the inherent data properties, assesses the clusters independently of any external context. The internal analysis in this chapter, includes the examination of filter visualisations and class activation maps, which illuminates the process by which the network learns image representations. This approach enables a more comprehensive understanding of the data's inherent structure and clustering performance.

Filter visualisations and class activation maps

To understand if the network is behaving as expected, I visualised some of the filter weights of the convolutional layers. Visualising these filters can help to understand if the network has been trained well. Smooth filters, without noisy patterns, usually indicate that the network has been trained properly. Moreover, visualising the weights of the first filter also reveals whether the network is simply capturing colour information or also direction (e.g. edges). Yosinski et al. [254], created a network-based model that synthesises images to maximise the activation of a given filter. Thereby, I was able to interpret the filters of convolutional layers of the network trained on the satellite image dataset by exploring what low-level features might be important for the classification. I used this method, in which I apply a cross-entropy function between the target filter and the other filters in the same layer, to create filter visualisations [242].

Class Activation Maps (CAMs) are a visualisation technique that enables the interpretation of the decision-making process of CNNs. They allow the researchers to see which regions in the input image have activated certain classes or predictions, thus providing insight into what the network might be "seeing" or "focusing on" when making its decision. CAMs generated by performing a weighted sum of the feature maps from the last convolutional layer, where the weights correspond to the importance of each feature for a

specific class output from the final fully-connected layer. This results in a heatmap that highlights the influential regions for the predicted class. Through this, CAMs provide a way to understand the spatial distribution of importance within the input and can be instrumental in explaining the predictions of complex models in computer vision [255, 256]. I used two implementations the CAM and CAM++ to visualise the CAMs of the trained network [255, 256].

Feature space analysis

To understand to what extent DeepCluster learns the intermediate image representations that distinguish clusters, I inspected the cluster membership in the lower-dimensional feature space. For this purpose, I obtained the 256 principal components of the image representations extracted by the second last fully connected layer of the CNN. These 256 principal components together accounted for 99.5% of the variance of the 4096 extracted features which the CNN uses to represent each image. I measured the degree of uniformity of tiles in the feature space that fall within each cluster with the average distance to the cluster centroid (intra-cluster distance).

The average distance to the cluster centroid (intra-cluster distance) for a cluster C is given by:

$$D_{avg}(C) = \frac{1}{|C|} \sum_{x \in C} d(x, c) \quad (4.3)$$

where $|C|$ is the number of points in the cluster, $d(x, c)$ is the distance between a point x and the cluster centroid c .

The degree of uniformity of tiles that fall within each cluster is calculated with the average Euclidean distance to the cluster centroid (intra-cluster distance) in the feature space. This is calculated as *within-cluster sum of squares*, averaged by the number of observations such that clusters of different sizes can be compared.

Additionally, I measured the inter-cluster similarity as the average Euclidean distance between each pair of cluster centroids in the feature space. Essentially, the *between-cluster sum of squares* measures the variation between all clusters and is averaged by the number of clusters. Larger inter-cluster distances generally imply better separation and hence potentially better clustering results, while small values can indicate that clusters are close to each other in the feature space.

The average Euclidean distance between cluster centroids (inter-cluster distance) in the feature space for a cluster C is given by:

$$D_{Euc}(C) = \frac{1}{|C|} \sum_{x \in C} \sqrt{\sum_{i=1}^n (x_i - c_i)^2} \quad (4.4)$$

where $|C|$ is the number of points in the cluster, x_i represents the i -th dimension of point x , and c_i represents the i -th dimension of the cluster centroid c .

I report the results and corresponding visualisation of the feature space in Section 4.3.2.

4.3 Results

4.3.1 External cluster analysis

Clusters of the urban environment

As part of the primary analysis described in this Chapter, I divided GAMA into eight clusters and named each cluster as shown in Figure 4.5. Cluster development over varying cluster numbers is reported in Chapter 5. I further report built (buildings and roads), and natural (vegetation and water) environment and demographic characteristics of the clusters which are shown in Figure 4.6. Some clusters contained a single dominant character-

istic with a strong visual representation. These include tiles with water, dense vegetation, and densely populated areas, especially those with distinctly visible building orientation between 36 and 45 degrees with respect to cardinal directions (as defined and calculated with data described in Section 3.6). The tiles that fell in these clusters had distinct distributions throughout the city, driven both by the regional environment and how the city has developed over time. The *Dark dense vegetation* cluster, which captures forest areas, is mainly located in the periphery of the city although some few patches were also present in the more urban Accra Metropolitan Area (AMA), for example at the University of Ghana campus. The *Water* cluster captures bodies of inland *Water* and is often surrounded by a cluster consisting of vegetation that is less dense and lighter in colour (*Light vegetation*). The two clusters that capture very densely populated areas (*Densely populated areas, >36 degree building orientation* and *Densely populated areas, <36 degree building orientation*) are mostly located in the AMA, and in adjacent metropolises, covering over one quarter of GAMA. Building-related metrics (building count, area, mean size) and population density were high in both clusters, with their distinguishing feature being building orientation with respect to cardinal directions (as defined in Section 3.6) and shown in Figure 4.7.

The *Light vegetation* cluster and *Empty land* cluster captured natural environments that were more heterogeneous than the clusters described above. Tiles falling into the *Light vegetation* cluster varied more in their NDVI than the *Dark dense vegetation* cluster that solely captures dense vegetation in a higher NDVI range (~ 0.3 to 0.5). Tiles in the *Empty land* cluster typically had low population density and were often located next to the vegetation clusters in the northeast and northwest of the study area. Based on visual inspection, these were areas of dry soil, such as gravel, unpaved roads or sandy terrain.

Other clusters were more complex and contained multi-dimensional environmental characteristics rather than any single dominant characteristic. These include the *Buildings*

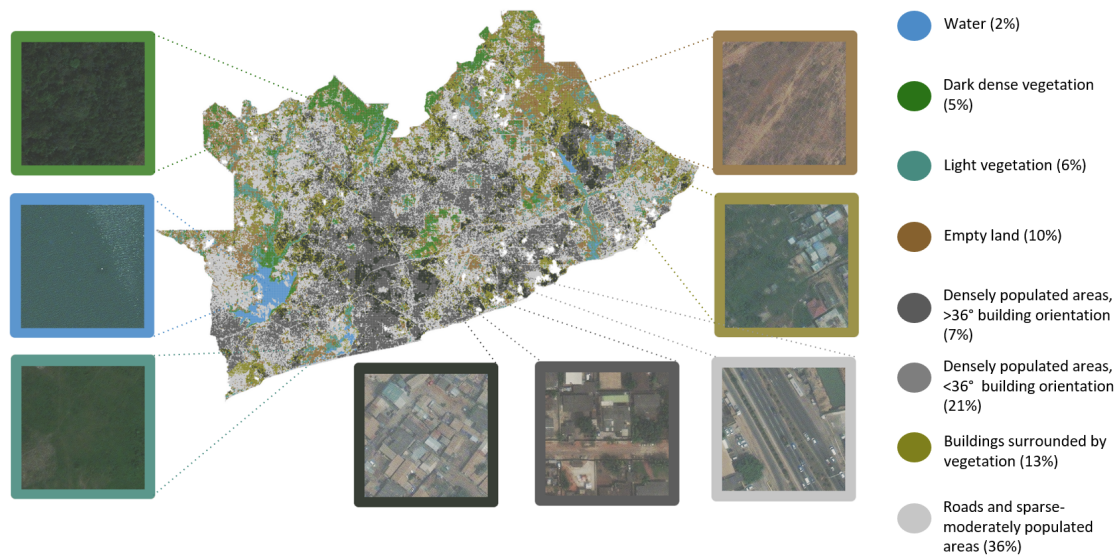


Figure 4.5: The figure shows 222,420 image tiles each assigned to eight clusters. Each cluster is shown in a different colour – the same colour is used for each cluster in subsequent figures. The boxes show examples of the tiles that were assigned to each cluster. The numbers next to each cluster name show the percentage of tiles grouped into the cluster.

surrounded by vegetation and *Roads and sparse-moderately populated areas* clusters, which together covered one half of the GAMA. The *Roads and sparse-moderately populated areas* cluster was spread throughout the city, surrounding and moving out radially from the densely populated areas in the city centre. The *Buildings surrounded by vegetation* cluster consisted of tiles that include buildings with low average building size and that were mostly located in the peri-urban areas.

Prediction of cluster assignment with external variables

To quantitatively evaluate what features of urban form, environment and population are most represented in the visually identified clusters, I trained a machine learning classifier (XGBoost) to predict cluster membership using environmental and demographic variables not used in clustering, as described in Section 4.2.5. Figure 4.8 shows the SHAP values, which are a measure of variable importance for describing cluster assignment. A higher SHAP score indicates larger relevance of a certain variable on predicting cluster assignment. The SHAP values show that NDVI was an important external predictor for



Figure 4.6: Caption next page.

Figure 4.6: The radar charts depict the built and natural environment and demographic characteristics of each cluster. Each environmental and demographic variable is scaled with a quantile transformer, a non-parametric transformation to map the data to a uniform distribution with values between 0 and 1 (0.5 indicates the median value of a variable across all tiles in the entire image). The colours of each chart correspond to the cluster map in Figure 4.5. Tiles with no buildings were included (as zeros) in summary statistics for building count and building area but excluded from the calculation of summary statistics for average building size and orientation so that zero is not used in the denominator. The length of major roads was zero in 93% of all tiles, hence the median of all tiles was zero as was the mean in most clusters.

NDVI: Normalised Difference Vegetation Index.

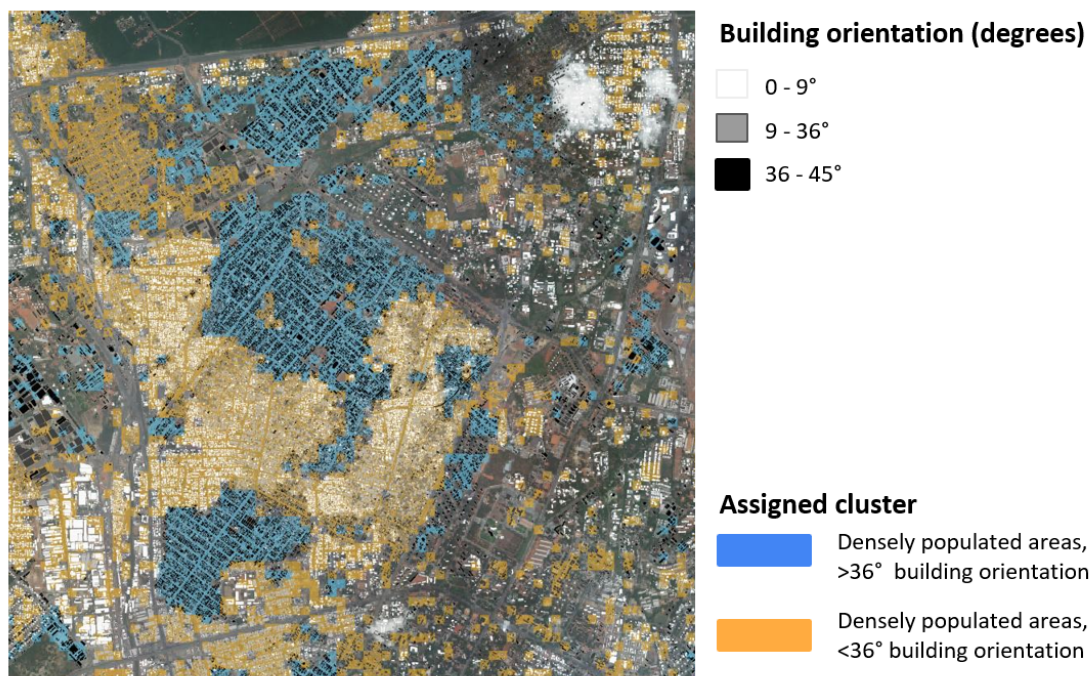


Figure 4.7: The figure shows a section of the satellite image that captures the *Nima* and *Mamobi* neighbourhoods in the centre of Accra. The building shapes are plotted on top of the satellite image and are shaded by building orientation from cardinal directions. The two clusters that capture densely populated areas are shaded in orange and blue. The clusters capture a visible distinction between how the buildings are arranged within the neighbourhood.

image-driven cluster membership, especially through its substantial role for identifying the *Water* (which had very low NDVI) and *Dark dense vegetation* (very high NDVI) clusters. At the same time, other variables helped predict cluster assignment beyond the role played by NDVI. Specifically, NDVI was followed by mean building area, whose most

salient role was in predicting *Empty land* and the two vegetation clusters (which all had very low building area), and building orientation, whose importance was driven most by its ability to predict *Densely populated areas*, *>36 degree building orientation*. Tiles belonging to the heterogeneous clusters of *Roads and sparse-moderately populated areas* and *Buildings surrounded by vegetation* were not predicted by any single environmental variable but rather had contributions from multiple ones. Distance and length of major roads were the least relevant variables in predicting cluster memberships, likely because their role was already captured in the same metrics for all roads. Population density and mean building size were ranked comparatively low as well, possibly because the information on building size was already captured by building area and count, and population density was moderately correlated with building-related metrics (Appendix Figure 3.13). This happens because the SHAP importance score is partitioned such that correlated features will not rank as highly as when their impacts were considered in isolation.

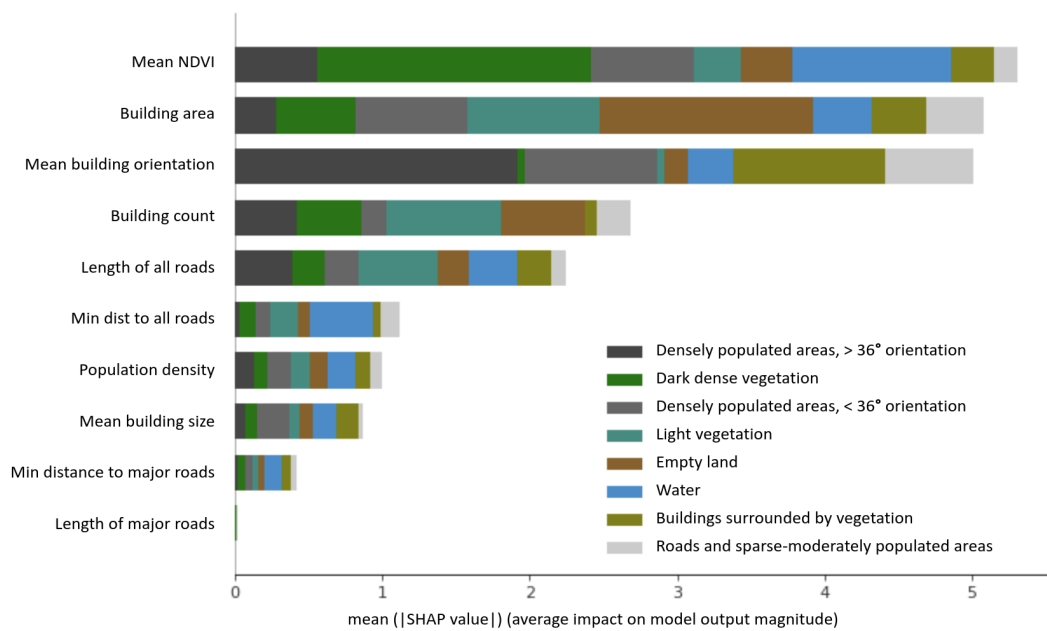


Figure 4.8: The figure shows SHapley Additive exPlanations (SHAP) [251] values, obtained by fitting XGBoost [249] classifier to predict cluster membership by environmental and demographic variables. The SHAP value for each variable indicates its predictive power for assignment to various clusters, and hence identify the measures of urban form (buildings and roads), environment (water and vegetation), and population that externally differentiate clusters that were generated based on images alone. The mean SHAP values from the XGBoost classifier were calculated for each environmental and demographic variable as described in Methods. The total length of each bar, which is the mean absolute SHAP value, represents the overall importance of each variable for predicting cluster membership, and the different colours represent the importance for assignment to each cluster.

Input from researchers with deep knowledge of Accra

The responses from researchers at Ghana University and the *Pathways to Equitable Healthy Cities* collaboration revealed that the city’s urban clusters and natural environments are largely influenced by human activities and historical planning. The densely populated regions, such as those with a $>36^\circ$ building orientation, were associated with high levels of informal business activities, increased access to transportation, and a heightened sense of social cohesion by the researchers. Further replies implied that the areas, indicative of urban sprawl, were heavily influenced by trade, transport, and affordable

housing stocks. Historical planning, particularly remnants of colonial city planning, also plays a significant role, with cities like Accra maintaining their original radial road network from the colonial period. For residents, these settings offer economic opportunities but also face challenges related to noise, pollution, and petty crime, as per another researcher. Natural environments, such as the *Achimota Forest* and community parks, have been preserved due to planning efforts. While there are no mining activities in GAMA, sand-winning and quarry activities do occur. Further responses indicated that the natural vegetation in these areas provides critical protection against run-offs during the rainy season, although recent encroachments for development have led to flooding issues in some areas.

4.3.2 Internal evaluation of clusters

To understand to what extent the CNN learned meaningful image representations used for clustering, I visualised network filters and CAMs, as well as plotted cluster assignment and the image tiles for each cluster in the feature space as described in Section 4.2.6.

Filter visualisations and class activation maps

Figure 4.9 shows examples of the filter weights of the first layer, which usually pick up information on different colours and directions. These direction and colour filters then get combined into a basic grid and spot textures deeper into the network. Finally, these textures gradually get incorporated into increasingly complex patterns taking more spatial information into account. Top activated filters refer to the filters that produce the highest activation values when applied to the satellite images. In other words, these are the filters that recognise patterns which are most present or most strongly expressed in the input data. These activated filters provide a good intuition about what features or patterns the CNN is "looking at" or "picking up on" when processing a given input. Therefore,

analysing the top activated filters can give us insights about how the CNN is making its predictions or classifications, helping us understand and interpret its internal workings.

The analysis of the top activated filters revealed expected patterns, such as parallel lines and edges and combinations of such, so that I could conclude that the CNN did not only pick up colour values but also patterns and contexts. For example, one filter of the fifth convolutional layer (Conv5) was really good at picking up the waves of water as depicted by Figure 4.9. Similarly, another filter highlighted lots of parked cars or trucks that formed a specific pattern. Figure 4.10 illustrates which parts of certain satellite tiles (on the far left) contributed the most to cluster assignment. Here, two main patterns emerged. First, the built environment such as roads and buildings was important for learning features as shown by the images. Furthermore, objects such as cars also highly impacted feature learning. For tiles that capture homogeneous landscapes, such as water or dense vegetation, the CAMs revealed activation across the whole image.

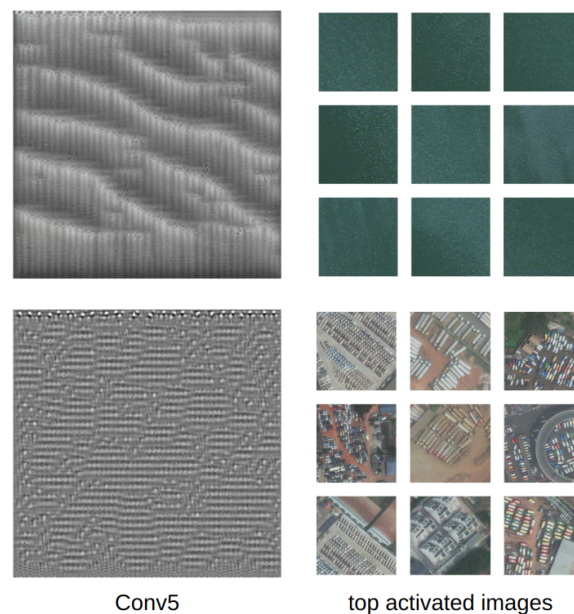


Figure 4.9: Top 9 activated images and filter extracted from the 5th convolutional layer (Conv5).

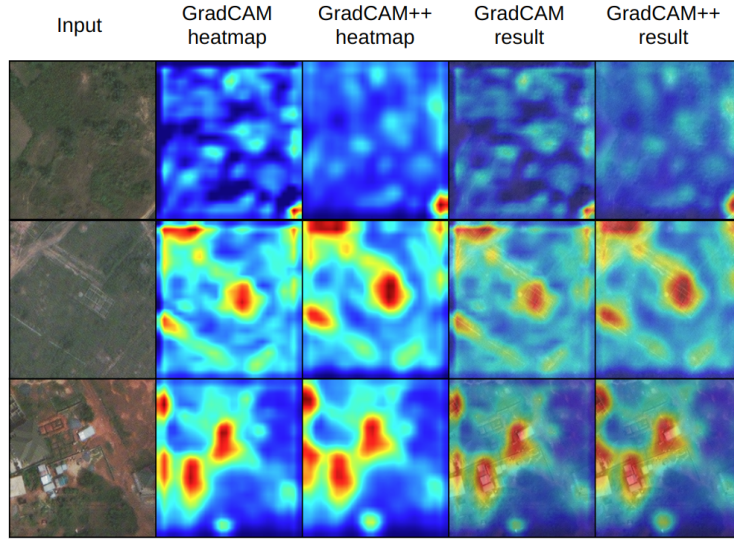


Figure 4.10: Class activation maps (GradCAM and GradCAM++) of DeepCluster trained on satellite images of Accra.

Cluster formation and variability in the feature space

To understand to what extent the network learns meaningful image representations used for clustering, I visualised cluster assignment and the image tiles for each cluster in the feature space (Figure 4.11) and measured the intra-cluster and inter-cluster distance (Figure 4.12), as described in Section 4.2.6. Tiles that fell into *Water* and *Dark dense vegetation* had small intra-cluster distance which indicates they were very homogeneous, and those in *Densely populated areas*, *>36 degree building orientation* and *Buildings surrounded by vegetation* cluster were highly variable (Figure 4.12); other clusters fell between the two groups in terms of their intra-cluster similarity versus variability. I also measured the inter-cluster similarity by calculating the average distance between each pair of cluster centroids in the feature space. The clusters that captured the natural environment (*Water*, *Dark dense vegetation*, *Light vegetation*, and *Empty land*) were close to one another in the feature space. The *Densely populated areas*, *>36 degree building orientation* and *Densely populated areas*, *<36 degree building orientation* clusters were furthest apart from the water and vegetation clusters, and the *Buildings surrounded by vegetation* and *Roads and sparse-moderately populated areas* clusters were intermediate distance to these groups.

The clusters that contained multi-dimensional environmental characteristics, *Buildings surrounded by vegetation* and *Roads and sparse-moderately populated areas*, were also close to each other in the feature space.

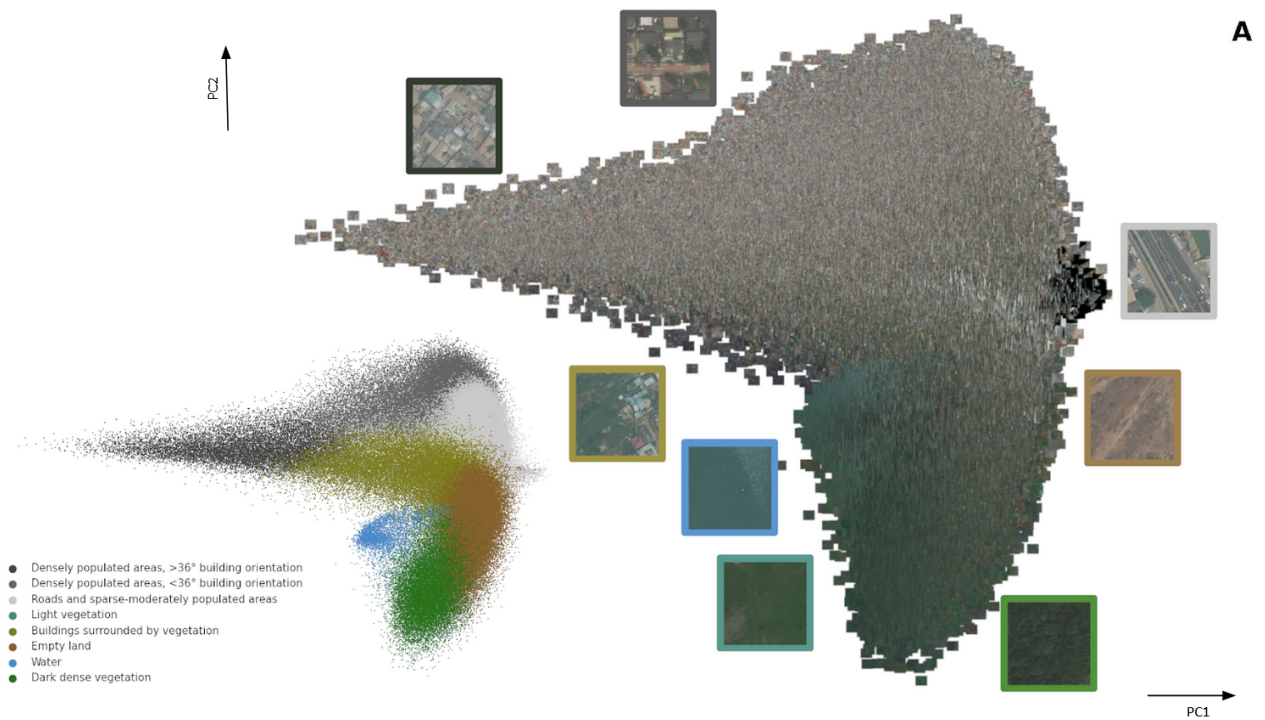


Figure 4.11: The figure shows the 222,420 image tiles plotted on the first and second principal components (PCs) of 4,096 features extracted from the second to last fully connected layer (of the CNN; these PCs account for 40% and 17% of the overall variance, respectively). Each tile is placed at the corresponding value of its features' PC1 and PC2. Each tile is also shown as a point, coloured in the same colours the cluster map in Figure 4.5. This presentation allows visualising image tiles, and their cluster assignment, in relation to extracted features, and can provide an intuition for how the network learns image representations in the process of cluster assignment.

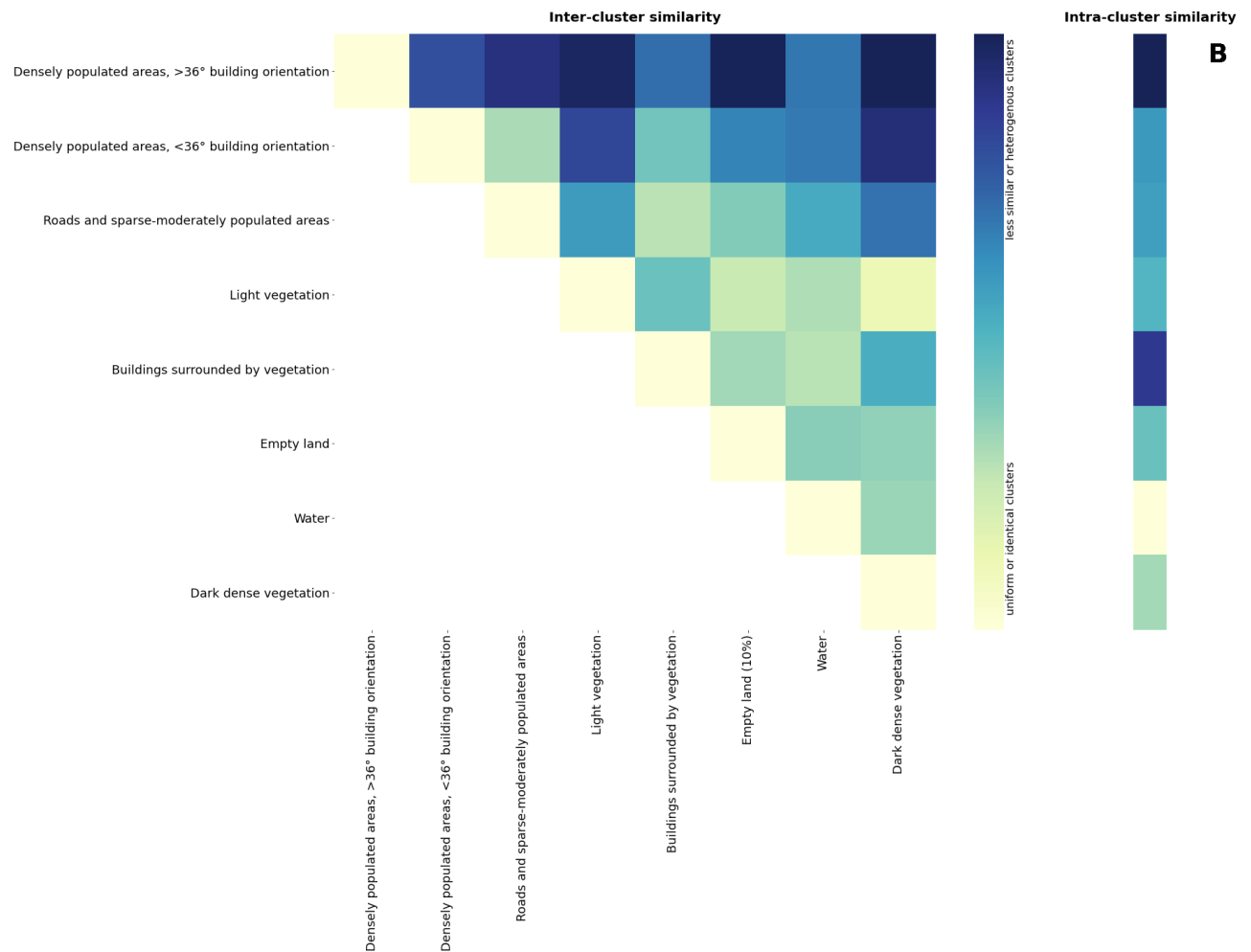


Figure 4.12: The figure shows average two measures of intra- and inter-cluster similarity. The intra-cluster similarity is calculated as average distance of tiles to cluster centroid and is a measure of within cluster uniformity (smaller distances) versus heterogeneity), and (larger distances). The distance between centroids (of clusters is a measure of similarity across clusters). The smaller the distance between the clusters, the more similar they are, and vice versa. For computational efficiency, within- and between-cluster distances were calculated using the first two PCs.

4.4 Discussion

Cities are complex dynamic systems whose built and natural environments, including buildings, roads and vegetation, are shaped through an interplay of local geography and human activity. These environments in turn affect where people live and conduct their

activities, how they commute among these places, and their impacts on health and well-being. My analysis showed that the application of unsupervised clustering can capture single- and multi-feature urban environments and hence offer a novel way of coherently and comprehensively characterising and tracking urban environmental change, especially in settings where labelled data are limited.

4.4.1 Implications for tracking and enhancing sustainable urban development

My results show that the image-based clusters present interpretable insights into neighbourhoods. The proposed framework can be used to track changes in the built and natural environment at a fine spatial scale and in near-real-time to inform urban planning and services. The clusters capturing vegetation and water, which have distinct visual features, are influenced by regional geography and whether and how it is preserved or modified. The *Dark dense vegetation* cluster captures the hills and valleys in the north of the GAMA, where human activities have so far been relatively minimal. These forests used to cover even larger parts of north and northwest regions of the GAMA, but land clearing for urban growth and agriculture, charcoal making and firewood collection has substantially reduced the extent of trees [257, 258]. The *Water* cluster covers waterways and other water bodies and is surrounded by *Light vegetation* cluster, which lies closer to human settlements than the *Dark dense vegetation* cluster. In addition to riverine vegetation, the *Light vegetation* cluster captures a wetland that separates the AMA from the adjacent metropolis of *Tema* in the east, as well as protected areas around a reservoir. These areas face threats of urban encroachment, and their conservation is necessary for the preservation of the region's biodiversity and to protect the city from flooding [201, 259, 260, 261, 262]. In particular, large parts of Accra's informal settlements (which fall in the two densely populated clusters) are located next to riparian areas and are exposed to a risk of flooding [263, 264], which is predicted to increase due to global climate change [260]. The *Empty land* cluster

captures open unvegetated land, such as sandy or bare soil with a few shrubs or unfinished buildings and unpaved paths. *Empty land* that has recently lost vegetation cover may be a setting for imminent road and building construction and could indicate the beginning of city sprawl [265]. These changes, and how they change the city's environment, can be readily monitored with sequential satellite images and my clustering approach.

The two densely populated areas (*Densely populated, >36 and <36 degree building orientation*) are mostly located in the dense core of the city and adjacent metropolis, a feature that is seen in many cities in Africa [266]. These densely populated areas tend to be poorer than other parts of the city but are well-connected to urban transport and trade hubs, which makes them a setting for informal and formal business and trade activities. The dense population and the commercial activities create more social cohesion but also make these areas noisier and more polluted [267, 53, 268]. Despite currently having no vertical layering, both clusters are dense in terms of building footprint, as is the case in informal settlements in other cities in the developing world [51, 269]. There is a trend by private and public-private developers towards larger and higher buildings in these highly accessible locations, which contrasts with earlier urban sprawl in Accra and other major African cities [4, 270]. Vertical densification can increase economic productivity, while also displacing their current residents unless accompanied with appropriate housing in the same or nearby locations as a part of redevelopment. It will also likely change the visual characteristics of these areas and hence can be measured and monitored through the approach that I presented. The main difference between the two densely populated clusters is the visually distinct building orientation, which impacts thermal comfort. This feature which will be increasingly relevant as extreme weather events, especially high temperatures, become more frequent with changing climate [224, 225, 227, 226].

The remaining clusters capture a more complex and heterogeneous mix of land cover and land use. The areas captured by the *Buildings surrounded by vegetation* cluster are

either farmland at the fringe of the metropolitan area or wealthier neighbourhoods with freestanding houses surrounded by gardens, often as part of gated communities. These areas have lower pollution [267, 53, 268] and benefit from proximity to green space. The parts of the cluster that are at the fringe of the city are expected to grow further into the surrounding natural environments [271], driven by population growth, cheaper housing stock compared to the city centre, and poorly controlled private land development due to weak enforcement of urban planning and development rules [201, 272]. Limiting this sprawl requires a combination of land tenure reforms, and introduction/enforcement of urban planning and zoning regulations, so that outward growth and sprawl are balanced with (vertical) densification of already-built areas as described earlier [201, 273]. Finally, the *Roads and sparse-moderately populated areas* cluster represents the combination of the city's low-medium density residential and commercial settlements and its road network. Road capacity in GAMA is inadequate for the increasing number of vehicles, and the peri-urban areas are underserved [274], as evidenced by the relative underrepresentation of *Roads and sparse-moderately populated areas* cluster in the peri-urban areas, especially in the northeast of GAMA. The combination of this cluster, the two densely populated clusters and the *Buildings surrounded by vegetation* cluster, captures most of the paved and unpaved roads, and provides a good representation of the connectivity and accessibility of the city. The arrangement of these clusters can help to identify areas that are poorly connected, and reveal options for improving their connectivity including walkable and bikeable areas, those that can be connected to central Accra with radial rapid rail or bus transport systems, and those that may require additional roads [50, 275, 276].

Application and extension to other cities and multiple time points

The unsupervised approach can be applied to multiple cities to reveal similarities and differences in the character of natural and built environments. I discuss specifics and approaches to how the analysis can be moved from one city to multiple cities in Chapter

6 and Chapter 7. Similarly, the framework can be used for longitudinal analysis of satellite images taken at different points in time in order to track urban change based on how each phenotype expands or replaces others.

4.5 Summary

An unsupervised feature extraction and clustering algorithm was applied to high-resolution satellite images and evaluated with external data on the built and natural environment and population which was not used for the clustering. The external analysis included cluster interpretation with summary statistics and fitting a classifier with the external data to predict cluster membership. The fitted model was used for a subsequent SHAP analysis to quantify the environmental and demographic characteristics defining the clusters. The internal analysis was compromised of visualising CNN filters and CAMs, as well as the examining of the clusters in the feature space.

Chapter 5

Sensitivity analyses: Impact of scale of analysis and cluster number on cluster assignment

This chapter is based on the peer-reviewed publication titled "Phenotyping urban built and natural environments with high-resolution satellite images and unsupervised deep learning" [1], published in The Science of the Total Environment (STOTEN) journal, for which I was first author (DOI: <https://doi.org/10.1016/j.scitotenv.2023.164794>). I hereby give credit to this work without citing the work in every instance in the text.

5.1 Overview and objectives

In this chapter, I describe the sensitivity analysis of the feature extraction and clustering framework applied to high-resolution satellite images of Accra, Ghana as described in Chapter 4.

The main objectives of this chapter is to test the sensitivity to methodological choices applied to the main analysis described in Chapter 4. These methodological choices include examining cluster changes based on tile size, cluster number and additional model hyperparameters.

5.2 Methods

5.2.1 Data sources and pre-processing

I used high-resolution satellite image of Accra that covers most of the GAMA area as pictured in Figure 3.1 and described in Chapter 3. The spatial resolution of the satellite image amounts to 0.3 meters per pixel and captures three different bands (RGB). I further removed tiles that were not part of the GAMA as well as empty tiles and tiles capturing clouds as described in Section 4.2.1.

Tile size

As part of the pre-processing, I cropped the satellite image to cover only the GAMA and cut the satellite basemap into tiles that vary in size. These tiles, sized 128 x 128, 384 x 384, and 512 x 512 pixels, correspond to about 38 x 38 meters, 115 x 115 meters, and 154 x 154 meters on the ground, respectively. While the primary analysis uses 256 x 256 pixel tiles (as described in Chapter 4), for this analysis, I used both smaller (128 x 128, 50% smaller) and larger (384 x 384, 50% larger and 512 x 512, 100% larger) tiles to track changes in cluster membership and characteristics. Figure 5.1 presents these different tile sizes plotted on a segment of a satellite image.

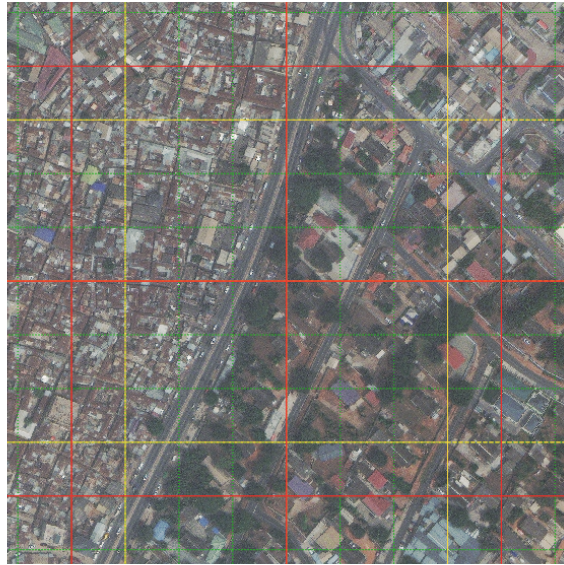


Figure 5.1: Segment of the satellite basemap with a 128 x 128 pixel (green), 384 x 384 pixel (yellow) and 512 x 512 pixel (red) pixel grid indicating the tile sizing of the sensitivity analysis.

Shifting the index grid

Apart from dividing the satellite basemap into different tile sizes, I generated an additional dataset by offsetting the index grid used in the main analysis (described in Chapter 4) by 50%, which corresponds to a shift of 128 pixels to the east and north. The dimensions of the tiles were retained at the same size as those used in the main analysis detailed in Chapter 4. The tile size roughly equates to an area of 75 x 75 meters on the ground.

5.2.2 Methods of sensitivity analyses

I used the DeepCluster method as described in Section 4.2.3. First, I analysed the robustness of the cluster change based on the scale of the analysis (i.e. tile size and zone). Second, I examined cluster development with changing number of clusters and finally, described the influence of hyperparameter k on feature learning.

Sensitivity to image tile dimensions

I investigated how the spatial scale of analysis impacted the cluster formation. I used the same set-up for the DeepCluster algorithm as described in Section 4.2.3, except feeding a different tile size into the CNN. The learning rate and number of clusters (K) were set to 0.00001 and 8, respectively. DeepCluster was run for 20 epochs with 3 RTX6000 GPUs, 72GB memory and a run time of approximately 24 hours.

To interpret the clusters, I used data capturing information on urban form, population density and vegetation described in Chapter 3. The external data was scaled with a quantile transformer and I reported the median values of each measure for all the tiles that fall into each cluster and visualised the results in the form of radar charts. Additionally, I generated co-occurrence plots to compare cluster membership across approaches with different tiles. For this purpose, I overlaid the two index grids for comparison and assign cluster memberships using the smaller index tiles. The results are reported in Section 5.3.1.

Sensitivity to zoning

I further examined how zoning impacted the cluster formation. For that reason, I repeated the DeepCluster analysis, as described in Section 5.2.1, with the shifted grid. I applied the DeepCluster algorithm in a manner consistent with the approach outlined in Section 4.2.3, except using the shifted image tiles. I set the learning rate to 0.00001 and the number of clusters (K) to 8. The algorithm was run for 20 epochs using a single RTX6000 GPU with 72GB of memory and the training and clustering were completed in approximately 24 hours.

Sensitivity to the number of clusters

In Chapter 4, I presented eight clusters of urban environment, each with its own phenotypic characteristics. This number was chosen based on visual inspection of results, and an initial set of experiments where I tested the separation of clusters and the level of detail required for an intuitive classification of the urban environment. To further understand how the choice of cluster number influences the separation of the city into clusters and the character of the resultant clusters, I modified the number of clusters, K , in the DeepCluster analysis and report how the clusters change from $K=2$ to $K=12$ in 5.3.3. I thus used the same setup as described in Section 4.2.3, except for changing the number of clusters (K). All other hyperparameters remained unchanged.

Influence of hyperparameter k on feature learning

DeepCluster iteratively groups the image representations deep in the network with a standard clustering algorithm, k-means, and uses the subsequent assignments as supervision or pseudo-labels to update the weights of the network as part of a classification task. The choice of the hyperparameter k in the k-means clustering part of the algorithm is distinct from the final number of clusters (K) in the data; rather, k influences how the algorithm learns distinctive image representations. It may be the case that a large hyperparameter k is better suited for feature learning even if for interpretation one prefers a smaller number of clusters [277, 260]. To examine the role of the number of clusters in the k-means algorithm on learning image representations, I carried out the analysis as a two-step approach: in the first step, the CNN is used to create deep image representations (features) with specific values of k , and subsequent step of clustering into K clusters. To investigate whether a larger k helps to learn more discerning features, I created a set of different deep features (DF) with the three choices of k ($k \in 8, 50, 100$), namely DF_{k8} , DF_{k50} and DF_{k100} . I then clustered the intermediate DF with a k-means ($K=8$) (i.e. as in the analysis described in Chapter 4) to compare the cluster memberships.

Cluster number and hyperparameter k

To find the optimal number of clusters for each set of deep features, I computed the distortion score and the Calinski-Harabasz (CH) from $K=2$ to $K=15$ with all sets of DF (DF_{k8} , DF_{k50} and DF_{k100}). The distortion score is calculated by summing the squared distances from each point to its assigned centre, which is also the intra-cluster distance defined in Equation 4.3. The CH score is a measure that summarises cluster compactness and spatial separation. It is defined as

$$CH(k) = \frac{B(k)/W(k)}{(n - k)/(k - 1)}, \quad (5.1)$$

where n is the number of data points, k is the number of clusters, $W(k)$ is the within-cluster variation, and $B(k)$ is the between cluster variation. It is therefore a measure that combines inter- and intra-cluster variation as described in Section 4.2.6.

Additionally to the CH score, I also computed the distortion or the sum of squared distances within clusters and plotted them against the number of clusters. This plot is often referred to as ‘elbow’ chart. The ‘elbow’ or bend in the plot indicates the point where adding more clusters does not significantly improve the model’s performance. I used the plots to observe how the clusters change based on a different hyperparameter k and cluster number.

5.3 Results

5.3.1 Sensitivity to image tile dimensions

I investigated cluster stability with varying tile size, specifically 128 x 128 (i.e. 50% smaller than the main analysis in side length), 384 x 384 (50% larger), and 512 x 512 (100% larger) pixels, as described in Section 5.2.2. Figure 5.2 shows the cluster assign-

ments plotted on a map of GAMA. The analysis showed that similar prominent features of the natural and built environment can be picked up by all different tile sizes and are subsequently clustered together. These dominant characteristics such as vegetation, water, and densely populated areas were clustered largely independently of tile size (Figure 5.3). The resultant clusters were stable in their (external) characteristics and in terms of areas they cover, supported by the co-occurrence plots that compare cluster membership between the different tile sizes, as shown in Figure 5.4. Tiles capturing multi-dimensional environmental characteristics in the main analysis, such as the *Buildings surrounded by vegetation* and *Roads and sparse-moderately populated areas* clusters, varied the most depending on the tile size and areas of the image they are assigned to.

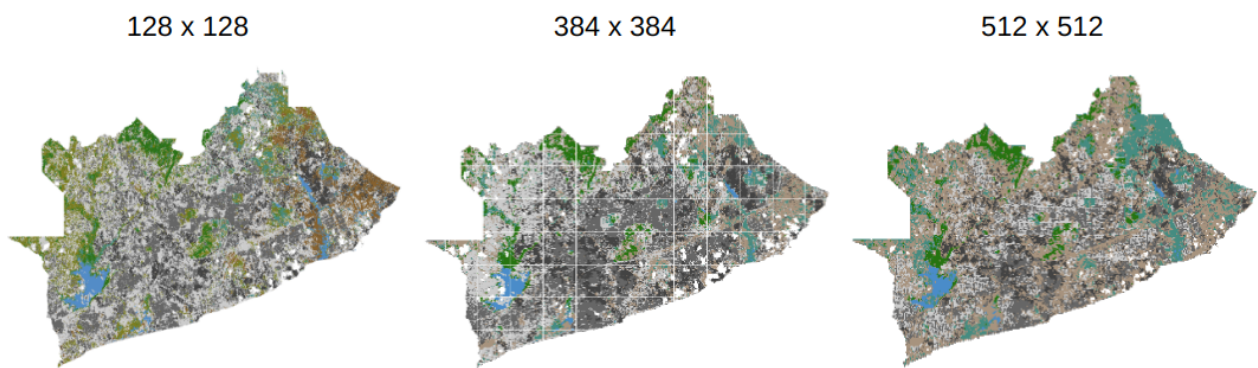


Figure 5.2: Cluster results of varying tile size plotted on a map of GAMA.

The sensitivity analysis demonstrates that similar prominent features of the natural and built environment were picked up by all different tile sizes and are subsequently clustered together. The main characteristics of the city, such as vegetation and features of the built environment, were all picked up by different tile sizes. The 128 x 128 pixels tiles did not capture contexts as well and larger tile sizes (384 x 384 pixels, 512 x 512 pixels). These larger tiles were by their nature more likely to capture heterogeneous landscapes. Therefore, the clusters that include higher heterogeneity of features in the analysis with 256 x 256 pixel tiles, such as *Buildings surrounded by green*) and *Buildings surrounded by vegetation*, were the ones that vary the most depending on the tile size.

Smaller tiles (128 x 128 pixels) captured more homogeneous landscapes and less mixed

environments than the original and larger tile sizes, such as *Buildings surrounded by vegetation*. In contrast, larger tiles (384 x 384 pixels, 512 x 512 pixels) were more likely to capture mixed environments, such as sparse-moderately populated areas and buildings and roads together with vegetation. These two clusters make up 49% and 54% of the total tiles for the analysis with 384 x 384 pixels and 512 x 512 pixels, respectively, compared to 39% in the 128 x 128 pixel analysis.

5.3.2 Sensitivity to zoning

As outlined in Section 5.2.2, I investigated how the cluster assignment changed by using a shifted dataset. Figure 5.5 shows the standardised averages of clusters between the main analysis as described in Chapter 4 and the DeepCluster results of the shifted grid. The analysis revealed that tiles with characteristics that were very similar in both approaches were clustered together. Both approaches had two densely populated clusters with main differentiation in building orientation. Additionally, both approaches picked up very similar clusters of the natural environment. These included *Water* clusters (cluster 6 (left) and cluster 4 (right)) and a *Dark dense vegetation* cluster (cluster 7 (left) and cluster 2 (right)) as shown in Figure 5.5. The main differences were seen in clusters that captured mixed environments, such as the *Buildings surrounded by vegetation* cluster that varied mostly in median building orientation and population density.

5.3.3 Sensitivity to number of clusters

As described in Section 5.2.2, I analysed how the clusters change from setting $K=2$ to $K=12$ (Figure 5.6). Initially ($K=2$), the tiles were separated into a cluster containing two very homogeneous natural environments (*Water* and *Dark dense vegetation*) and a cluster that captured all the other tiles. At $K=4$, that mixed cluster split into clusters

with distinct visual character including populated areas and *Empty land* and *Light vegetation*. The natural environment cluster that formed at $K=2$ further split into dark green vegetation and *Water* tiles at $K=6$. At $K=6$, the algorithm also grouped densely populated areas with high building orientation together, a cluster that stays constant until $K=12$. Clusters capturing mixed environments such as *Roads and sparse-moderately populated areas* and *Buildings surrounded by vegetation* appeared later at $K=8$. By $K=10$ and $K=12$, very particular clusters appear, such as a cluster that captures populated areas with very low building orientation (<10 degrees), a cluster capturing riparian areas, and a cluster that contained edges of clouds. Comparing the cluster development with the cluster homogeneity in the extracted feature space (Figure 4.11), the most separable clusters (dark and *Light vegetation* and *Water*) were also the clusters which emerged to be the most internally uniform as shown by the feature plot (Figure 4.12).

Influence of hyperparameter k on feature learning

To investigate how the hyperparameter k influences the learning of image representations, I compared the cluster results of three sets of deep features (DF) that were learnt with different choices of hyperparameter k ($k \in 8, 50, 100$), namely DF_{k8} , DF_{k50} , and DF_{k100} , as described in Section 5.2.2. The visual interpretation through the radar plots of the three sets of clustered deep features (Figure 5.6) and the co-occurrence of cluster assignments (Figure 5.7) showed that the tiles were mostly grouped in similar ways for different values of k and had consistent environmental and demographic characteristics. Each choice of k included a *Water*, *Dark dense vegetation*, *Light vegetation*, and *Densely populated areas with >36 degree building orientation* and *<36 degree building orientation* cluster. The main changes between cluster results were in the *Empty land* cluster and clusters with the mixed environments, the *Buildings surrounded by vegetation* and the *Roads and sparse-moderately populated areas* clusters, which differed in building area and count, as well as building orientation. The building area and count of the clusters formed from

the DF_{k50} and DF_{k100} were slightly higher than the cluster results from DF_{k8} for the *Roads and sparse-moderately populated areas* cluster, but slightly lower for the *Buildings surrounded by vegetation* cluster. This change arose mainly due to the fact that more tiles were assigned to the *Roads and sparse-moderately populated areas* and *Buildings surrounded by vegetation* clusters than in the main analysis, with 60% of tiles being assigned to these two clusters in the DF_{k50} analysis and 69% in the DF_{k100} analysis, compared to 50% in the main (DF_{k8}) analysis. The additional tiles for the *Buildings surrounded by vegetation* mainly came from the *Empty land* cluster in the main analysis (Figure 5.7). Furthermore, the remaining part of the *Empty land* cluster created with DF_{k50} and DF_{k100} also contained some vegetation compared to the clusters formed in the main (DF_{k8}) analysis.

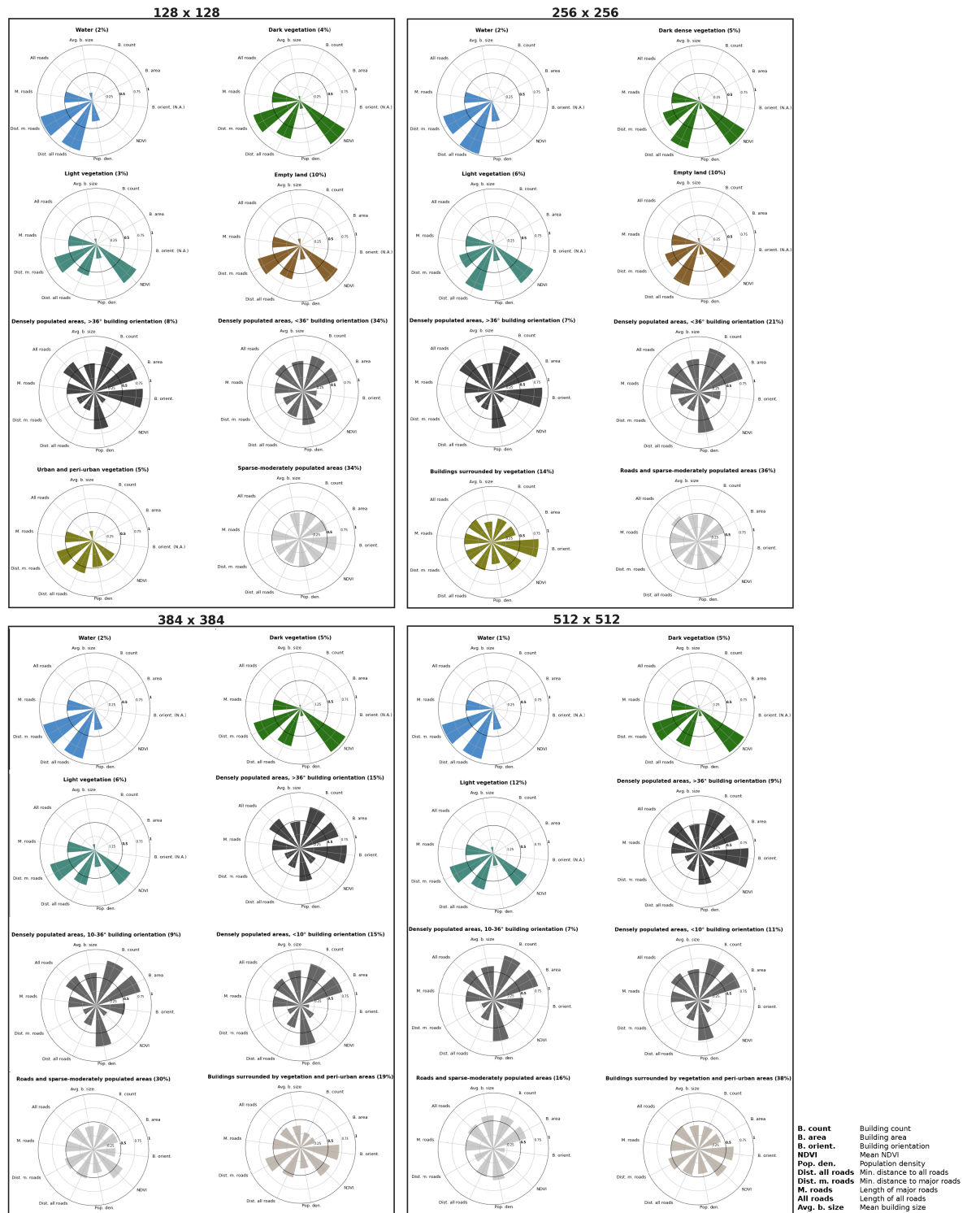


Figure 5.3: The radar charts show the built and natural environment and demographic characteristics of each cluster per analysis with tile sizes: 128 x 128 pixels, 256 x 256 pixels (main analysis), 384 x 384 pixels, and 512 x 512 pixels. Each environmental and demographic variable is scaled with a quantile transformer, a non-parametric transformation to map the data to a uniform distribution with values between 0 and 1 (0.5 indicating the median value of that certain feature across all tiles in the entire image). Tiles with no buildings were included in summary statistics for building count and building area (as zeros) but excluded from calculation of summary statistics for average building size and orientation so that zero is not used in the denominator.

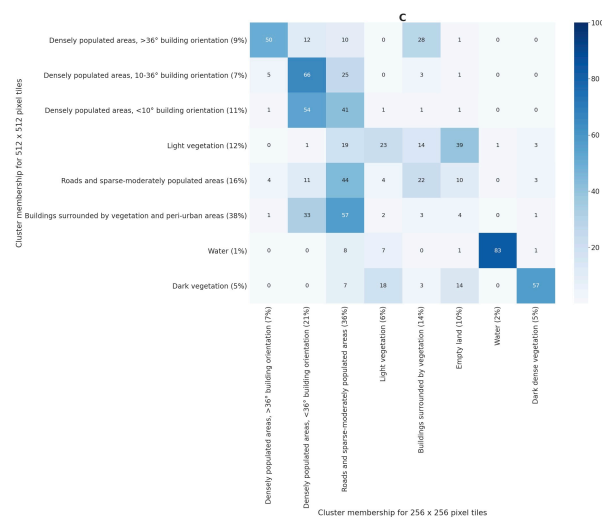
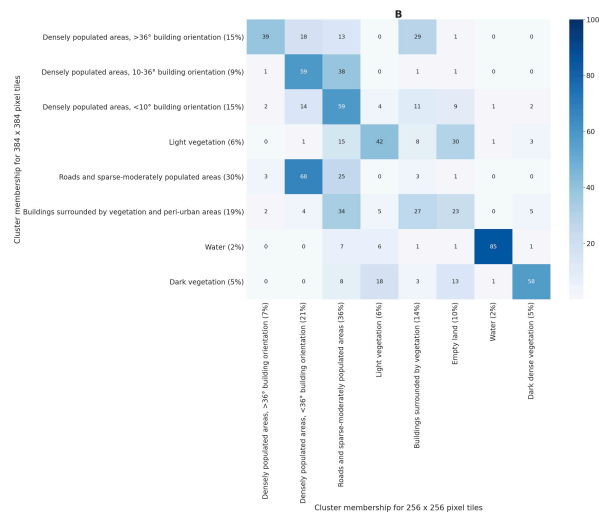
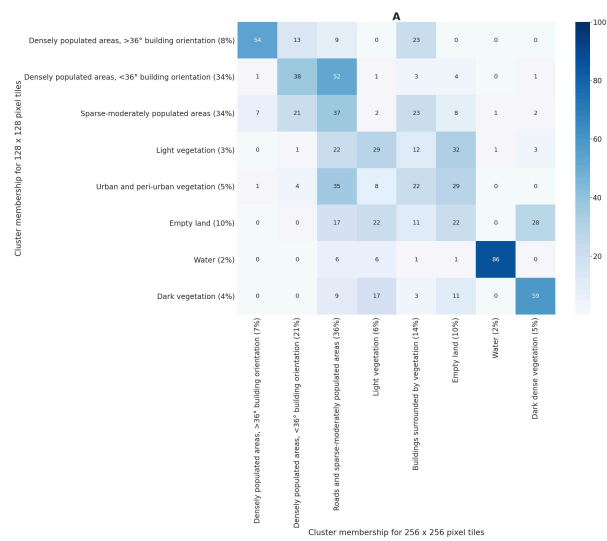


Figure 5.4: The plots show the co-occurrence of assignment to specific clusters, comparing the 128 x 128 pixels analysis (A) with the main analysis, 384 x 384 pixels analysis (B) with the main analysis, and 512 x 512 pixels analysis (C) with the main analysis. The co-occurrence is calculated by spatially joining the 256 x 256 pixel grid with either (A), (B) or (C) and computing a frequency table of cluster membership. The values are presented as, and coloured by, percentages of the main analysis.

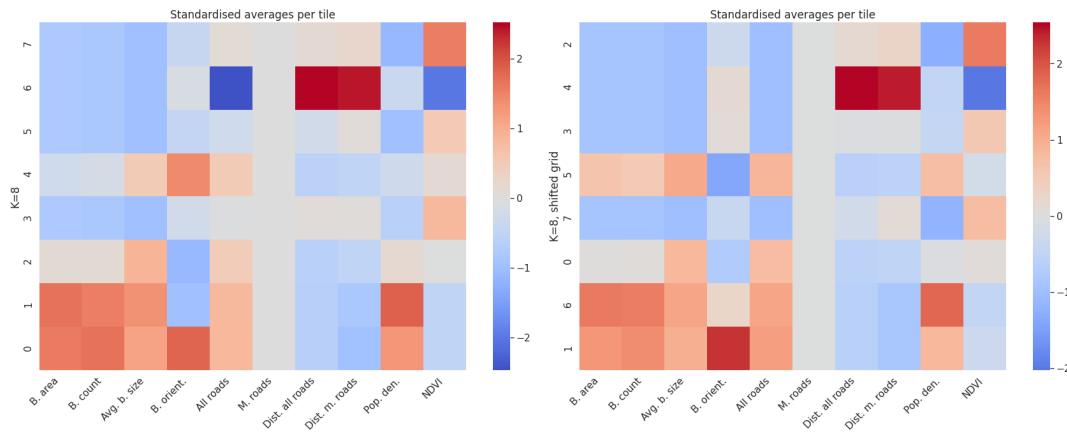


Figure 5.5: Standardised averages of clusters with $K=8$. Main analysis (left) and analysis with shifted grid (right).

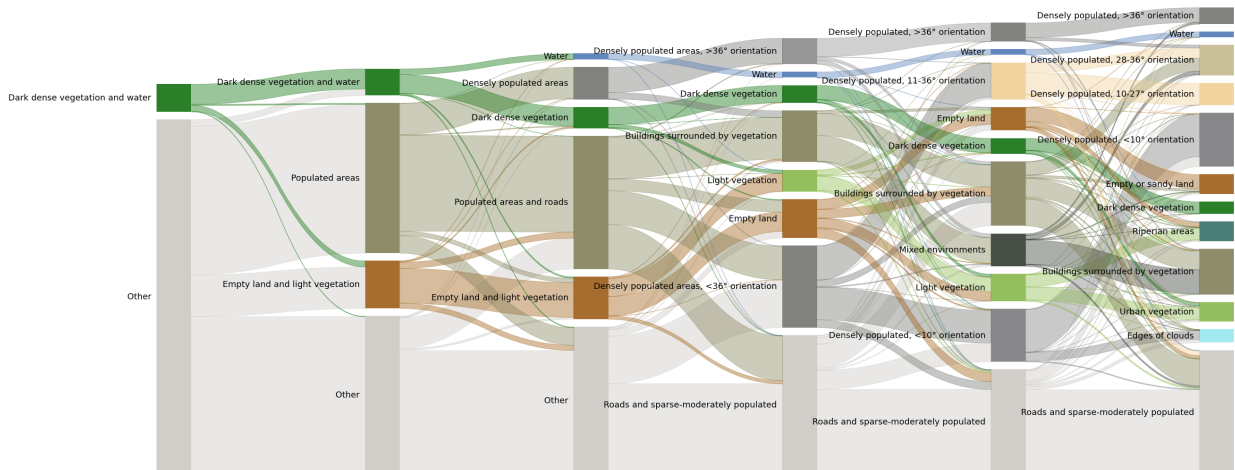


Figure 5.6: The Sankey plot shows the sizes of clusters when cluster number (K) varied from 2 to 12. The flow of data indicates how clusters partition and merge based on repeating the main analysis (as described in Chapter 4) with varying K . The naming of the clusters for each K is based on a similar process as in the main analysis, and based on the environmental and demographic characteristics of each cluster.

Figure 5.6: Built and natural environment and demographic characteristics of the clustered deep features (DF_{k8} , DF_{k50} and DF_{k100}) are shown as radar charts similar to those in Figure 3. Each environmental and demographic variable is scaled with a quantile transformer, a non-parametric transformation to map the data to a uniform distribution with values between 0 and 1 (0.5 indicating the median value of that certain variable across all tiles in the entire image). Tiles with no buildings were included in summary statistics for building count and building area (as zeros) but excluded from calculation of summary statistics for average building size and orientation so that zero is not used in the denominator.

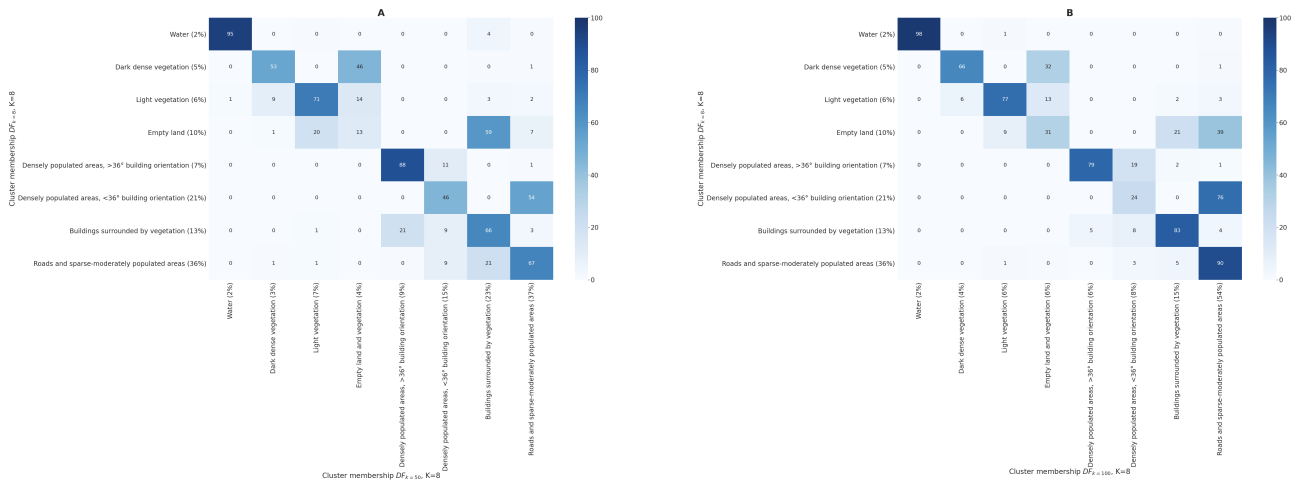
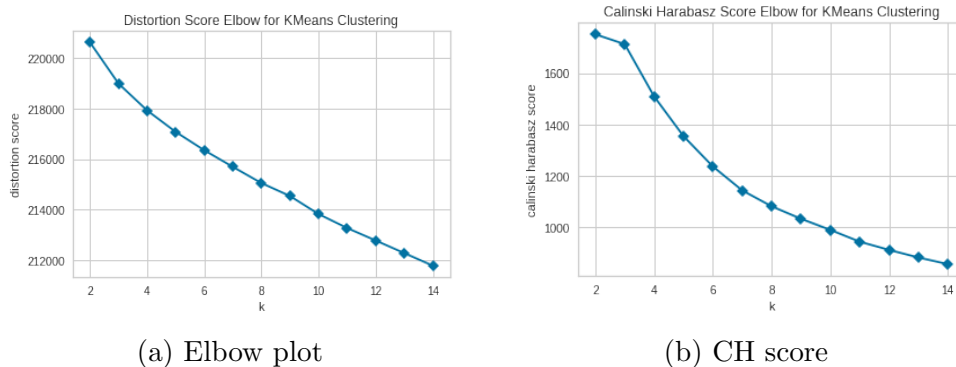
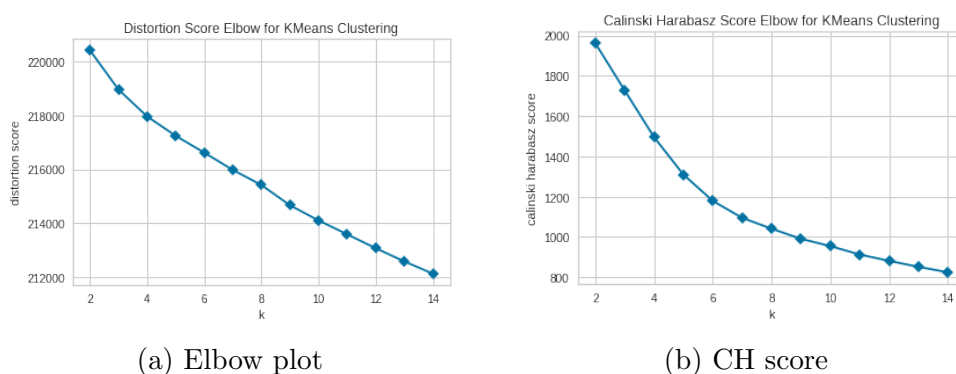


Figure 5.7: The plots show the co-occurrence of cluster assignments, comparing the cluster membership of the main analysis (DF_{k8}) with the DF_{k50} analysis (A) and the main analysis (DF_{k8}) with the DF_{k100} analysis (B). The co-occurrence is calculated by computing a frequency table of cluster membership. The values are presented as, and coloured by, percentages of the main analysis.

Cluster number and deep features

To observe how changing the hyperparameter k together with the number of clusters changes the results, I calculated the distortion score (also known as ‘elbow’ chart) for $K=2$ to $K=15$ for each set of deep features. These deep features are image representations learned with a specific hyperparameter k as described in Section 5.2.2. All elbow plots (Figure 5.8 - Figure 5.10) showed that there was no clear ‘elbow’ in all three plots but rather the sum of squared distances steadily decreasing. Similarly, the CH score is decreasing in a similar matter, however, the curve exhibits a subtle tapering at around

Figure 5.8: Elbow plot and CH score from $K=2$ to $K=15$ of DF_{k8} Figure 5.9: Elbow plot and CH score from $K=2$ to $K=15$ of DF_{k50}

$K=8$.

5.4 Discussion

Cities, as complex and dynamic systems, feature an interplay of human activity and local geography, impacting everything from living and activity locations to commuting patterns and health outcomes. My research has revealed the potential of unsupervised clustering to capture these intricate urban environments, particularly when labelled data is scarce. Yet, when working with a single large satellite image that requires pre-processing for analysis, certain considerations arise regarding how the pre-processing steps can impact cluster assignments. These arise from artificially imposed zones and tile sizes on continuous geographic phenomena, which can influence the results of spatial statistical analysis.

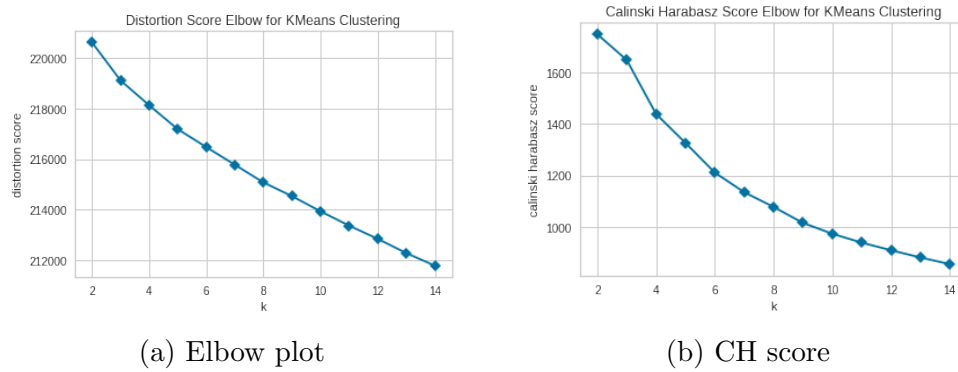


Figure 5.10: Elbow plot and CH score from K=2 to K=15 of DF_{k100}

5.4.1 Finding a suitable scale of analysis

Using a basemap satellite image, which has been stitched together from multiple satellite images, gives the researcher a lot of freedom in terms of scale and zone of analysis. While it allows for a lot of flexibility, it also opens up a set of new questions related to zoning and tile size. In this thesis, I analysed the sensitivity to zone and scale through changing the tile size and tile zoning before repeating the analysis. In geographical and spatial statistical analyses, scale and zoning are two fundamental factors that can greatly influence the results. In this context, the dependence of statistical results on spatial partitioning of the data is also known as the Modifiable Areal Unit Problem (MAUP) [278]. The MAUP consists of two related problems. Firstly the problem of scale, which arises when the same set of data is aggregated at different spatial levels (e.g., zip codes, counties, states). The results of the analysis can vary widely depending on the level of aggregation, which can lead to very different interpretations. Secondly, the zoning problem, which refers to the issue that even if the scale is constant, the way you divide the study area into units (how you draw the boundaries) can also influence the results. The conducted sensitivity analyses relate to both aspects of the MAUP. I addressed the problem of scale by running the analysis with different tile sizes that were larger and smaller than the original analysis described in Chapter 4. I further approach the problem of zoning by shifting the index grid by half of the analysis index, i.e. 128 pixels, to the east and north.

The sensitivity analysis with changing tile size with results reported in Section 5.3.1 was supported by He et al. [279]. The authors describe that when satellite classifications are aggregated by majority class the majority classes become larger while smaller classes are reduced (or disappear). This is also the case for the larger tile sizes, where for example, the two mixed environment clusters (*Buildings surrounded by vegetation* and *Sparse-moderately populated areas*) become larger with increasing tile size.

On the contrary, in the field of deep learning for remote sensing as well as medical imaging [280, 281, 282], tiling methods have often been viewed as solutions to overcome memory limitations rather than strategies to enhance the accuracy of the algorithms. This suggests that the main purpose of employing tiling methods is to offset memory constraints rather than to boost the predictive capability of the algorithms. Reina et al. [280] argue that larger degrees of image context, such as using larger tile sizes improve model performance while training the neural network and are less sensitive to hyperparameters during an inference task. However, this might be more applicable to segmentation or object detection tasks, rather than unsupervised analysis.

Lastly, expanding on the analogy drawn from plant ecology, the concept of beta diversity can provide valuable insights into the impact of varying tile sizes in remote sensing analyses. In ecology, beta diversity measures the variation in species composition between different ecosystems or habitats [283], reflecting the turnover of species as one moves from one location to another. Similarly, in the context of image analysis with changing tile sizes, the turnover of features or classes across varying tiles can be considered as a form of beta diversity.

In any case, DeepCluster clustered tiles with similar characteristics together, despite different tile sizing. In the analysis with the larger tile sizes, the mixed clusters became the prominent clusters and thus become more difficult to interpret compared to the more homogeneous clusters such as the clusters capturing water or densely populated areas. Smaller tile sizes (128 x 128 pixels) did not capture mixed environments that well and most clusters either captured the built environment or the natural environment and only

one cluster captured a mix of the two.

5.4.2 Finding a suitable K

Geodemographic and environmental characteristics' classifications aim to summarise the distinct socio-environmental traits of small area zonal geography, as demonstrated in previous studies by Ojo, Skupin, and Singleton [284, 285, 286]. Drawing from these works, my research sought to depict the variations in urban form at a small geographical level, however, using imagery data, as opposed to their reliance on survey and census data. Similar to geodemographic applications where data have high dimensionality, choosing the ideal cluster number, K , for classifying the urban landscape is a subjective process. As Singleton and Skupin have noted, the choice of K often relies more on qualitative factors, such as the desired granularity of output, usefulness to end users and comparability with other classifications, and is informed by a range of measures like cluster frequencies or mapped outputs.

In the context of this thesis, a suitable K should yield clusters that demonstrate differences in measures of urban form, as interpreted with external data, and display visually discernible differences. Yet, determining an exact K is challenging; any number greater than 12 tended to be difficult to interpret for visual interpretation. Analyses with varying numbers of K and different hyperparameter k for feature learning showed similar trends. The features were clustered densely together in the feature space and the commonly used elbow plot did not result in a clear 'elbow' and thus a preferable number of clusters. The CH score favoured very low values for the number of clusters in all cases, however, did seem to flatten at around $K=8$. The Sankey analysis further revealed two distinct super-clusters (one capturing the built environment and one the natural environments) which were further divided into sub-clusters. A lower number of clusters (K) can thus aid in distinguishing the built and natural environment, and a higher number of clusters can

highlight more specialised clusters that capture mixed environments. Generally, choosing the optimal K value is a multifaceted process, with no one-size-fits-all solution. It is invariably influenced by the specifics of the application and the level of detail desired in the classification output.

5.5 Summary

The sensitivity of the feature extraction and clustering algorithm was probed by applying the framework to high-resolution satellite images of various sizes and different zoning compared to the main analysis described in Chapter 4. The resulting clusters were evaluated with external data on the built and natural environment and population which was not used for the clustering. The external analysis included cluster interpretation with summary statistics and comparing results between parameter changes. These changes entailed repeating the analysis with varying tile size, shifted tile grid, and varying cluster number (K). Finally, alternative values for hyperparameter k were also tested. The sensitivity analysis showed that clusters that captured homogeneous landscapes were more robust compared to clusters that captured mixed environments.

Chapter 6

Clustering multiple cities

6.1 Overview and objectives

In this chapter, I explain how I extended the cluster analysis described in Chapter 4 and Chapter 5 to further cities in SSA. First, I state the objectives, and describe the data and methods including joint, separate and hybrid multi-city clustering experiments. Finally, I report the results of the experiments and evaluate the cluster results of the various experiments.

The primary objective of this chapter is to cluster high-resolution satellite images of four cities in Sub-Saharan Africa (SSA), specifically Accra (Ghana), Dakar (Senegal), Dar es Salaam (Tanzania), and Kigali (Rwanda). To tackle differences in urban characteristics across these cities, I investigate an array of clustering approaches. These range from joint clustering to separate and hybrid clustering approaches. These clusters are further assessed using demographic and environmental data, allowing for a comparison of cluster membership across the different approaches. Ultimately, the most effective approach is identified through visual interpretation and external analysis, laying the groundwork for the comprehensive examination and analysis detailed in Chapter 7.

6.2 Methods

6.2.1 Data sources and pre-processing

I applied and tested the unsupervised clustering approach as described in Chapter 4 and Chapter 5 with high-resolution satellite images of four study cities in SSA, namely Accra (Ghana), Dakar (Senegal), Dar es Salaam (Tanzania), and Kigali (Rwanda). The satellite raster data was acquired from Maxar in GeoTIFF format as described in Section 3.5.1. The spatial resolution of the images amounts to 0.3 meters per pixel and captures three different bands (RGB). Colour balancing and orthorectification were applied to the images as part of commercial pre-processing.

I cut the satellite images of all four cities into 256 x 256 pixel tiles, which is equivalent to about 75m x 75m on the ground. This tile size contains multiple features of urban form, e.g., houses and roads, such that it captures objects with their urban context and surroundings. Table 6.1 lists the number of tiles used for the analysis and further details about coverage and year of satellite image capture.

City	Year	Coverage	Number of tiles
Accra	2019	1,312 km^2	222,420
Dakar	2021/2	570 km^2	96,627
Dar es Salaam	2021	1,465 km^2	248,375
Kigali	2020	750 km^2	127,082

Table 6.1: Satellite images used for the cluster analysis of four cities.

Additionally, I used four different sets of data on urban characteristics in each city to interpret the image-based clusters. The vector and raster datasets capturing the urban characteristics are described in Section 3.6. To calculate the mean per-tile statistics of the urban characteristics, I intersected each vector and raster dataset with the 256 x 256 pixel index grid of the satellite image tiles for each city.

6.2.2 Inference on Accra model

Before running a new model on the satellite images of the four cities, I tested how well the Accra model (approach 0) described in Chapter 4 generalises to unseen satellite images of other cities. The DeepCluster model, outlined in Section 4.2.3, uses cluster memberships as labels to retrain a CNN and consequently learns image representations of the satellite tiles. I saved the network trained on the images of Accra and fed unseen images into the network for image inference without retraining the network. The satellite images of Dakar were pre-processed in the same way as the images of Accra, meaning they were normalised based on the mean and standard deviation of the ImageNet dataset, since these weights were used for initialising the training of the network.

The results proved unsatisfactory. One notable issue was the model's tendency to classify a significant number of satellite tiles as part of the *Water* cluster. As illustrated in Figure 6.1, this problem became apparent when inferring cluster assignment for satellite image tiles of Dakar using the Accra model described in Chapter 4. This could be because the *Water* cluster is relatively close to the densely populated clusters in the feature space as shown in Fig. 4.12. The NDVI of water is also closer to the NDVI that captures the built environment than the NDVI of the natural environment. Even though it is not clear why the urban landscape was misclassified, this outcome indicated that the model's ability to generalise effectively across different data distributions was limited.

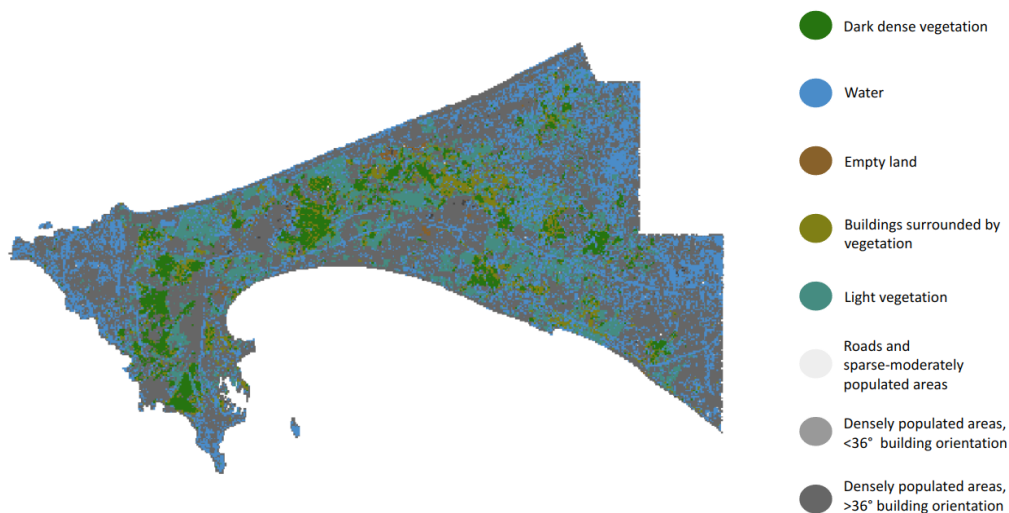


Figure 6.1: Transferring the Accra model to satellite images of Dakar.

6.2.3 Model considerations for the multi-city analysis

Cities in SSA share similarities, such as having a high share of compact, low-rise buildings; displaying indications of city sprawl that show stark internal contrasts, such as business developments next to informal settlements [4]. While these cities are similar in certain ways, they also differ based on geography and history. Different climates and geographies (like mountains or access to the sea) also shape a city’s development as described in Chapter 3. The aim of this chapter is to create interpretable and actionable land use classes with high-resolution satellite images of multiple cities. Moving the analysis from one city (as described in Chapter 4 and Chapter 5) to multiple cities opens up a set of methodological questions.

I compare three cluster approaches for the feature learning and clustering of satellite images, as depicted in Figure 6.2. As described in Chapter 4, the DeepCluster framework includes feature learning and clustering. Both steps can be included in a single pipeline or taken apart. I took advantage of this possibility to move along the city-specific and joint spectrum to cluster the images of multiple cities. The three different approaches are as follows:

- Approach A: **Joint one-stage model**: Joint feature learning and clustering of all cities together.
- Approach B: **Separate one-stage models**: Separate learning of features and clustering of each city one at a time.
- Approach C: **Hybrid two-stage model**: Joint feature learning across all cities and separate clustering by city.

In Approaches A and B, feature learning and clustering are performed in one step, using the same framework in Section 4.2.3. Approach C includes two steps, where representations of the satellite images of all the cities are learned together in the network and subsequently clustered by city. Different approaches implicitly have different assumptions about urban features and clusters. For Approach A, I assume that the cities share similar features, such that we find clusters that are shared between cities and hence the clustering is done together. When the analysis is done for each of the cities separately, as in Approach B, I assume that the inter-city variations are very high, and therefore, there is no benefit in sharing information about cities during feature learning and clustering. Finally, the hybrid two-stage approach, Approach C, assumes that inter-city similarities are useful for feature learning, but clusters cities separately to avoid missing any city-specific intricacies that would be lost if the cities were clustered together due to different distributions of urban land use. This last approach shares some similarities with transfer learning, where knowledge gained from training on one task is used to enhance the learning and performance of a different but related task.

In order to test the different assumptions, and to find out if the different cluster approaches give intuitive and meaningful clusters, I developed three sets of experiments (Approach A, B, and C), as depicted in Figure 6.2.

Approach A: Joint one-stage model

For Approach A, feature learning and clustering are done jointly in an end-to-end fash-

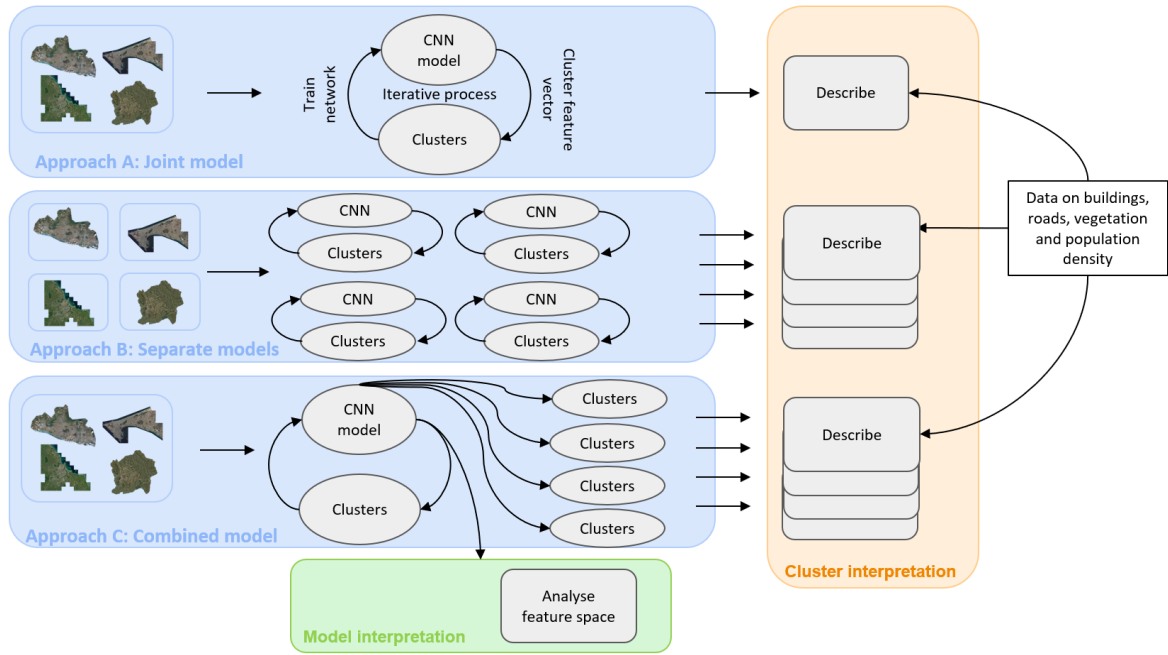


Figure 6.2: In Approach A, the deep features are learned and clustered in one step for all cities together. In Approach B, all deep features are learned and clustered together for each city separately. Finally, in Approach C, the features are learned together, however, clustered individually per city. This approach allows for cluster comparisons in the feature space.

ion. A single model is trained for all cities. The hyperparameter k that is used in the DeepCluster approach is set to the final cluster number K , as described in Section 6.2.5. I investigated a set of different values for the number of clusters K and report the results of $K=8$ and $K=16$ in Section 6.3.1.

Approach B: Separate one-stage models

For Approach B, the features are learned and clustered in a single end-to-end pipeline for each city. Figure 6.2 illustrates the workflow. The satellite images of one city are fed into the hybrid feature learning and clustering pipeline. Here, the hyperparameter k , which is used to learn the features in the DeepCluster approach is set to the final cluster number K , as described in Section 6.2.5. For each city, a separate model is trained and the results are reported in Section 6.3.1.

Approach C: Hybrid two-stage model

Finally, I propose a two-step approach that learns and clusters the image representations in separate steps. In the first step, the number of clusters to learn the features, k , was set to $k = 1000$ in the DeepCluster approach to allow learning of diverse features and clusters across cities. Subsequently, the extracted features are PC-reduced and ℓ^2 normalised before clustering them with the final number of clusters, K , separately for each city.

6.2.4 Cluster evaluation

External cluster evaluation

I used the data on the built environment, water, vegetation, and population to quantify the characteristics of the image-based clusters. The external data sources are described in detail in Chapter 3. Furthermore, I computed the deviation of each urban characteristic per tile and cluster, relative to the median of its assigned cluster, a metric I refer to as the Median Absolute Deviation (MedAD). The MedAD measure is robust to outliers and is delineated as follows:

$$MedAD = \text{median}(|y - \hat{y}|)$$

where y represents the actual feature characteristic for each tile, and $\hat{y}$ represents median feature characteristic for each cluster.

The overall sum of all *MedAD* scores across all urban characteristics, the Median Absolute Deviation (MAD) score, is defined as follows:

$$MAD = \sum \text{median}(|y - \hat{y}|)$$

I use the *MAD* score to summarise variability across urban characteristics for different cluster approaches. A sign that the algorithm effectively differentiates between natural

and built environments would be its ability to assign tiles with similar urban characteristics to the same cluster. This metric thus aims to encapsulate the homogeneity of urban characteristics within a single cluster. A cluster with highly variable urban characteristics would have a higher MAD than a cluster whose tiles are closer to the overall median of a certain urban characteristic (e.g., NDVI or building count) specific to that cluster. Concurrently, I use the MAD score to compare the cluster assignments of approaches A, B, and C, per city and across cities in Table 6.5.

In addition to the MAD score, I used Sankey diagrams to analyse the cluster changes between approaches. To complement the visual analysis through the Sankey diagrams (as reported in Section 6.3.1), I used the Normalised Mutual Information (NMI) to report cluster overlap between approaches. The NMI is a variation of Mutual Information (MI) and serves as a measure of the correspondence between two clusters, providing a score normalised between 0 and 1. A score of 1 implies perfect correlation, whereas 0 indicates the two variables are completely independent.

The NMI is defined as follows.

$$NMI(X; Y) = \frac{MI(X; Y)}{\sqrt{H(X)H(Y)}} \quad (6.1)$$

with the MI between two discrete variables, X and Y , defined as:

$$MI(X; Y) = \sum_{x \in X} \sum_{y \in Y} p(x, y) \log \left(\frac{p(x, y)}{p(x)p(y)} \right) \quad (6.2)$$

and the entropy of a random variable X is defined as:

$$H(X) = - \sum_{x \in X} p(x) \log p(x) \quad (6.3)$$

6.2.5 Number of clusters

In the analysis of Accra in Chapter 4 and Chapter 5, I found that a cluster number (K) of 8 was effective in separating satellite images of Accra into distinct, intuitive clusters. However, the optimal K value might vary depending on the specific objectives of the analysis. Typically, a smaller K value groups the urban area into primary clusters representing built and natural environments, whereas an increased K value further divides these main clusters into more nuanced sub-clusters.

To ensure consistent comparisons across different cities and methodologies, I set a standard K value of 8 for all analyses. The comparison of results across various cluster methodologies and cities with differing K values would present challenges, especially as the MAD score generally decreases with an increased number of clusters. Maintaining a constant K value allows for a more manageable comparison.

However, for Approach B, which clusters all cities together, I also included results using a larger K value ($K=16$). This larger K value allows more clusters to emerge and has a chance of revealing a cluster that may be more common in one city even if not emerging across all three together. While I maintained a consistent number of clusters throughout this chapter, it is important to note that depending on specific applications, the number of clusters may need to be adjusted.

Hyperparameter experiments

Besides comparing cluster results between approaches, I used the MAD score to determine the optimal learning rate for each approach. As mentioned in Section 4.2.3, some learning rates returned the inherent structure of how the satellite images were stitched together as part of the commercial pre-processing by Maxar, which was also reflected in a high MAD score. Thus, to find interpretable and intuitive results that are useful in an urban analysis setting, I ran the DeepCluster model with varying learning rates. I compared the results of the models of all approaches run with learning rates set to 0.01, 0.001, 0.0001 and

0.00001 and I chose the learning rate with the lowest MAD score for each approach. The Table A.3 in the Appendix reports the various MAD scores computed to select learning rates for Approach A. Here, I compare the results of $K=8$, $K=12$ and $K=16$ across cities. I added $K=12$ to the hyperparameter experiments to observe how the MAD score changes between $K=8$ and $K=16$. Additionally, Tables A.4 through A.6 report the MAD scores for the separate models (Approach B) in each city. Table A.7 reports the MAD score corresponding to varying learning rates for the hybrid model (Approach C) across all cities. The MAD scores for Approach C in each city are detailed in Tables A.8 through A.11.

Computation

The training time for 20 epochs was approximately 24 hours for one city, and about 72-76 hours for all cities together. Additional training did not substantially change the clusters and their interpretation. I trained the model on 3 RTX6000 GPUs with 72GB memory for the separate city models (Approach B) and 250GB memory for joint and hybrid approaches (Approaches A and C), where all satellite images of all cities were fed into the CNN at the same time.

6.3 Results

6.3.1 Cluster results for joint, separate and hybrid models

I report the results of the different clustering approaches as outlined in Section 6.2.3 and visualised by Figure 6.2. In the following sections, I present the results of the joint one-stage model (Approach A), the separate one-stage model (Approach B) and the hybrid two-stage model (Approach C). Later in the chapter, I compare the cluster assignments across all approaches with the MAD score.

Approach A: Joint one-stage approach

Table 6.2 lists the different clusters for each city at $K=8$. The clusters were named based on what they represent given the demographic and urban environment data described in the Data Chapter 3.

I found that Approach A, as described in Section 6.2.3, predominantly separated one city from the other cities. Four of the eight clusters were mainly assigned to Accra. Only one cluster existed across all four cities, which is *cluster 3*. The remaining clusters belonged to either Accra or the other three cities as shown in Figure 6.3. The clusters assigned with Approach A represented less detailed features of the environment and mainly separated specific types of vegetation from all other features, thereby limiting the detail of features encapsulated within each cluster.

Figure 6.4 displays the *MedAD* score per cluster and urban characteristic for Approach A at $K=8$. The figure illustrates that among all cities, *cluster 3* — a cluster common to every city — exhibited the most variation (i.e. highest MedAD scores) in urban characteristics. This could indicate that the approach has pushed the clusters together and thus created one default cluster for anything that does not go into another cluster and did not actually capture similar urban neighbourhoods across cities. I expected that certain clusters with

Approach	Cluster number	Accra	Dakar	Dar es Salaam	Kigali
A	0		Dark vegetation (1%)	Dark vegetation (7%)	Dark vegetation (2%)
	1	Dark vegetation Accra (9%)			
	2	Densely populated, high building orientation (21%)			
	3	Mixed environments (6%)	Mixed environments (91%)	Mixed environments (86%)	Mixed environments (71%)
	4	Densely populated, low building orientation (42%)			
	5	Empty land Accra (19%)			
	6		Shrubs, agriculture, med. vegetation (8%)	Shrubs, agriculture, med. vegetation (7%)	Shrubs, agriculture, med. vegetation (27%)
	7	Water (3%)			

Table 6.2: Table of clusters across cities with $K=8$, Approach A. The percentages represent the proportion of each cluster group in the respective city.

distinct visual features would be picked up across cities. For example, the *Empty land* or *Vegetation* clusters in Accra (*cluster 5* and *cluster 1*) captured a landscape that is found in the other study cities, however, was clustered separately from these. A further explanation for the separation of Accra could also point to a different commercial pre-processing that was captured by this approach.

I further tested whether a higher number of clusters would lead to more city-specific clusters and reran the analysis with $K=16$. As shown in Figure 6.5, similar patterns emerged as with $K=8$. The algorithm discerned more detailed city characteristics, such as agriculture in Dakar and wetlands in Kigali. However, one group of seven clusters remained specific to Accra, while the other clusters were distributed among the other cities. There was one shared cluster between all cities (*cluster 12*), with tile examples



Figure 6.3: Approach A: Joint one-stage clustering with $K=8$. Clusters are plotted on the map of Accra, Dakar, Dar es Salaam, and Kigali, from left to right. The algorithm separated most of Accra from the other cities.

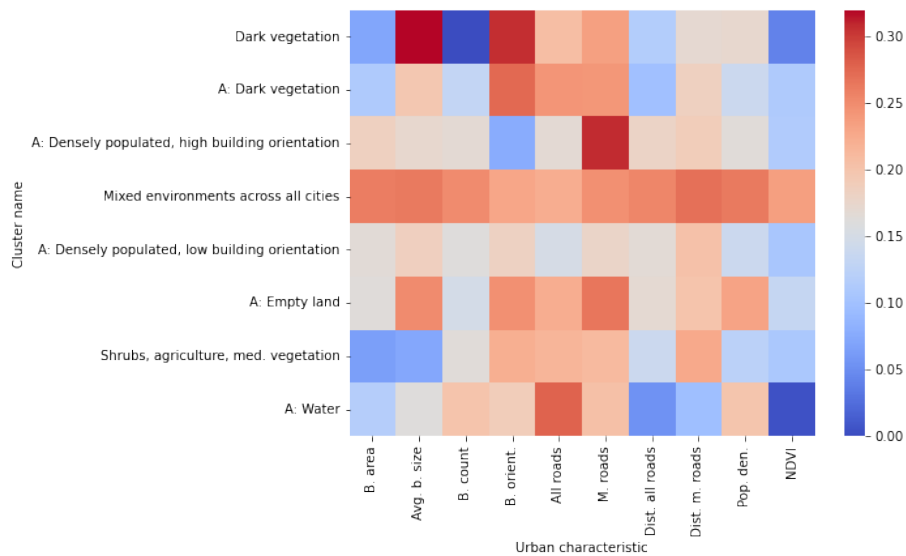


Figure 6.4: Heatmap of MedAD scores per cluster and urban characteristic for Approach A. Cluster names with A indicate that these only captured tiles of Accra. Cluster names without the A captured all other three cities, except for the *Mixed environments across all cities* cluster, which captured tile in all cities.

shown in Figure 6.6. Both figures showed no clear indication of why the tiles were clustered together and the results were challenging to interpret.



Figure 6.5: Approach A: Joint one-stage clustering with $K=16$. Clusters plotted on the map of Accra, Dakar, Dar es Salaam and Kigali, from left to right. The algorithm separated most of Accra from the other cities.



Figure 6.6: Shared cluster across cities for Approach A with $K=16$. Each row represents one city, from top to bottom: Accra, Dakar, Dar es Salaam and Kigali. The shared cluster is very difficult to interpret visually.

Approach B: Separate one-stage approach

For Approach B, I applied the feature extraction and clustering method described in Section 6.2.3 to all four cities separately. The cluster results are listed in Table 6.3. Compared to Approach A, the cluster numbers were not shared across cities, meaning the model assigned eight individual clusters to each city. The clusters were named based on what they represent given the demographic and urban environment data described in the Data Chapter 3.

In Approach B, as described in Section 6.2.3, each city is clustered individually. Although the analysis was conducted separately, I found that some clusters picked up similar urban environments when interpreted with external data. Across all cities, I found that three main classes of clusters emerged. These included clusters with vegetation, clusters capturing mixed environments and clusters capturing densely populated areas. Figure 6.7 displays some examples of these tiles for all cities. Even though the vegetation visually varies across the study cities due to different climates, the vegetation was clustered together in every city. The second set of clusters that I found across cities was densely populated areas. Each city had one or two clusters that captured the highest building count and population density per tile and cluster together with very low NDVI and low values for distance to roads. In Accra and Dakar, both cities with lower average vegetation compared to the other cities, the algorithm found two densely populated area clusters, with their distinguishing feature being building orientation with respect to cardinal directions. The third set of clusters captured was more complex and contained multidimensional environmental characteristics rather than any single dominant characteristic. These mixed environment clusters represent the median values of the urban characteristics across cities, such as population density (8-18 people per tile) and median building count (7-9 buildings per tile). In each city, there was one mixed environment cluster, and it was spread throughout the city, surrounding, and moving radially from densely populated areas in the city centre.

Cluster group	Accra	Dakar	Dar es Salaam	Kigali
Vegetation	7: Dark vegetation (5%)	6: Vegetation (8%)	4: Dark vegetation (3%)	2: Dark dense vegetation (4%)
	3: Light vegetation (6%)		2: Medium vegetation (4%)	7: Wetlands (4%)
			1: Vegetation and unpaved roads (4%)	6: Shrubs (8%)
Water	6: Water (2%)	5: Lake Retba (Water) (1%)		
Agriculture		1: Agriculture (3%)	7: Agriculture (4%)	3: Agriculture (23%)
		7: Shrubs and agriculture (10%)	5: Sparsely populated areas and agriculture (10%)	
Empty land	5: Empty land (10%)	2: Sand and empty land (4%)		
Densely populated	0: Densely populated, high building orientation (7%)	0: Densely populated, high building orientation (5%)	0: Densely populated areas (10%)	0: Densely populated areas (8%)
	1: Densely populated, low building orientation (21%)	4: Densely populated, low building orientation (14%)		
Other mixed environments	2: Sparse-moderately populated areas (36%)	3: Sparse-moderately populated areas (56%)	6: Moderately populated areas (34%)	1: Sparse-moderately populated areas, sloped area (5%)
	4: Buildings surrounded by vegetation (14%)		3: Moderately populated areas (33%)	5: Moderately populated areas (37%)
				4: Mixed environment, sparse-densely populated (22%)

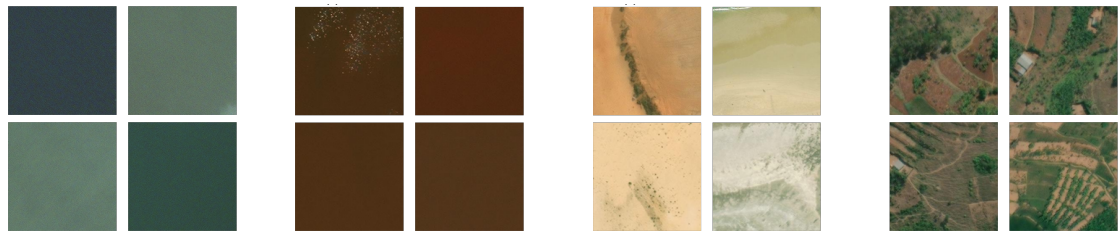
Table 6.3: Table of clusters across cities, Approach B. The percentages represent the proportion of each cluster group in the respective city. The number is the cluster number and is used in plots across Chapter 6 and Chapter 7 for identification.

Additionally, I found that Approach B also captured city-specific clusters such as water bodies. The water cluster in Accra picked up a large water reservoir (*Weija reservoir*),

as well as parts of the wetlands and river basin. Further, it clustered tiles with the pink water of *Lake Retba* together as well as sand and bare land in Dakar as seen in Figure 6.8.



Figure 6.7: Common clusters found across cities for Approach B. Each row represents one city; from top to bottom: Accra, Dakar, Dar es Salaam and Kigali. Dynamic mosaic ©2022 Maxar Technologies.



(a) **6**: Water, Accra. (b) **5**: Lake Retba, Dakar. (c) **2**: Sand and empty land, Dakar. (d) **3**: Agriculture, Kigali.

Figure 6.8: Examples of city-specific clusters for Approach B. The numbers refer to the cluster number in Table 6.3. Dynamic mosaic ©2022 Maxar Technologies.

Cluster group	Accra	Dakar	Dar es Salaam	Kigali
Vegetation	2: Light vegetation (7%)	5: Shrubs and unpaved roads (8%)	2: Medium vegetation (7%)	6: Shrubs (9%)
	4: Dark dense vegetation (6%)	6: Trees and bushy landscape (6%)	1: Vegetation and unpaved roads (6%)	2: Dark dense vegetation (6%)
			7: Dark dense vegetation (3%)	3: Wetlands (4%)
Water	7: Water (2%)			
Agriculture		7: Agriculture (13%)	6: Agriculture and sparsely populated areas (9%)	0: Agriculture and sparsely populated areas (15%)
			4: Agriculture (8%)	1: Trees and agriculture (9%)
Empty land	5: Empty land (10%)	1: Sand and empty land (8%)		
Densely populated	3: Densely populated, med. building orientation (22%)	3: Densely populated, med. building orientation (8%)	5: Densely populated, low building orientation (15%)	7: Densely populated areas (23%)
	0: Densely populated, low building orientation (19%)	4: Densely populated, low building orientation (7%)	3: Densely populated, high building orientation (12%)	
	6: Densely populated, high building orientation (13%)	2: Densely populated, high building orientation (5%)		
Other mixed environments	1: Other mixed environments (21%)	0: Other mixed environments (44%)	0: Other mixed environments (41%)	5: Other mixed environments, sparse-densely populated (29%)
				4: Mixed environment, sparsely populated; sloped area (5%)

Table 6.4: Table of clusters across cities, Approach C. The clusters are grouped into cluster categories such as vegetation and densely populated areas. The number is the cluster number, which is used in plots across Chapter 6 and Chapter 7 for identification. The percentages represent the proportion of each cluster group in the respective city.

Approach C: Hybrid two-stage approach

For Approach C, as described in Section 6.2.3, the image representations of the satellite images were learned together across all cities, however, clustered separately. The resultant clusters are detailed in Table 6.4. In contrast to Approach A, the clusters were not shared across cities, which means that the model assigned eight distinct clusters for each city. The naming of these clusters was derived from the interpretation of the demographic and urban environmental data discussed in the Data Chapter 3.

Approach C found analogous clusters to Approach B. These included three main cluster groups that were picked up in all cities: vegetation, densely populated areas and mixed environments as seen in Figure 6.9. Approach C also picked up city-specific clusters, such as Water in Accra and specific types of vegetation and agriculture in Dakar and Dar es Salaam, as depicted in Figure 6.10. Some city-specific features were lost, such as *Lake Retba* in Dakar.



Figure 6.9: Common clusters found across cities for Approach C. Each row represents one city; from top to bottom: Accra, Dakar, Dar es Salaam and Kigali. Dynamic mosaic ©2022 Maxar Technologies.

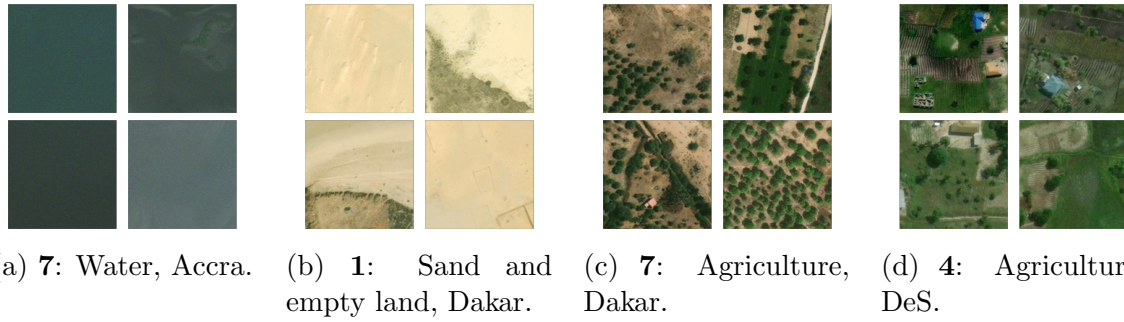


Figure 6.10: Examples of city-specific clusters for Approach C. The numbers refer to the cluster number in Table 6.4. Dynamic mosaic ©2022 Maxar Technologies.

Overall comparison of approaches

To compare all cluster results between cities, I computed the *MAD* score for each approach. The score is defined in Equation 6.2.4 in Section 6.2.4 and a lower *MAD* score indicates less variation within urban characteristics. Table 6.5 reports the *MAD* values compared between all three approaches and calculated on a city basis, as well as across cities. Approach C achieved the lowest *MAD* scores for Dakar, Dar es Salaam and Kigali, and these scores were closely followed by those of Approach B for all three cities. However, approach B achieved the best scores for Accra, closely followed by Approach C. Approach A (with $K=8$) attained the highest scores across cities. I further calculated the *MAD* across all cities, where the hybrid approach (Approach C) obtained the lowest *MAD* scores followed very closely by the separate approach (Approach A), with only a difference of 0.2474 in *MAD*.

Approach	Accra	Dakar	DeS	Kigali	All cities
A	2.0203	2.374	2.4905	2.3393	2.2021
B	1.9235	1.8118	2.2217	1.9675	1.6944
C	1.9709	1.7357	2.0411	1.9637	1.6125

Table 6.5: Comparing the overall *MAD* scores per city. Bold text indicates the best scores for each city. DeS: Dar es Salaam. Approach A refers to the $K=8$ model.

The visual analysis and external evaluation of Approach A demonstrated that the joint

one-stage clustering separated Accra from the other cities and did not generalise well across cities. Joint one-stage clustering (Approach A) with a higher number of K ($K=16$) resulted in a MAD score of 2.081 across all cities, which is still higher than the scores for the other approaches. Approach A did not satisfy the overall aim of this chapter, which was to create meaningful, interpretable and actionable clusters across cities. This was because the results of Approach A were very difficult to interpret across cities due to the separation of Accra from the other cities. It further clustered tiles together that were not visually similar. In the following section, I, therefore, focus on Approaches B and C. To understand how Approaches B and C differ, I investigated cluster changes between the approaches and report the z-scores for each urban characteristic per cluster as detailed in the next section.

Cluster changes between Approach B and C

The experimental outcomes indicated that Approach A did not offer substantial insight as it did not adequately capture variations across all cities and it split off Accra from the other cities. In this section, I thus first interpret the clusters of Approach B and C with demographic and environmental data and compare their main shared clusters in Figures 6.11 and 6.12. The figure displays a set of clusters that are found in both approaches, including clusters capturing vegetation, densely populated areas and mixed environments. Each city map is complimented with a radar chart that represents the urban characteristic of certain clusters. Similar shapes on the radar plots indicate that the clusters are capturing similar city environments based on demographic data and data on the built and natural environment. The radar charts showed that certain city features were picked up in both approaches, while others did not match as well. For example, the vegetation classified by Approach B is very similar to Approach C, however, the mixed environments captured differ significantly. The mixed environments captured by Approach B in Dar es Salaam are distributed evenly across the whole city, while this is not as much the case for Approach C.

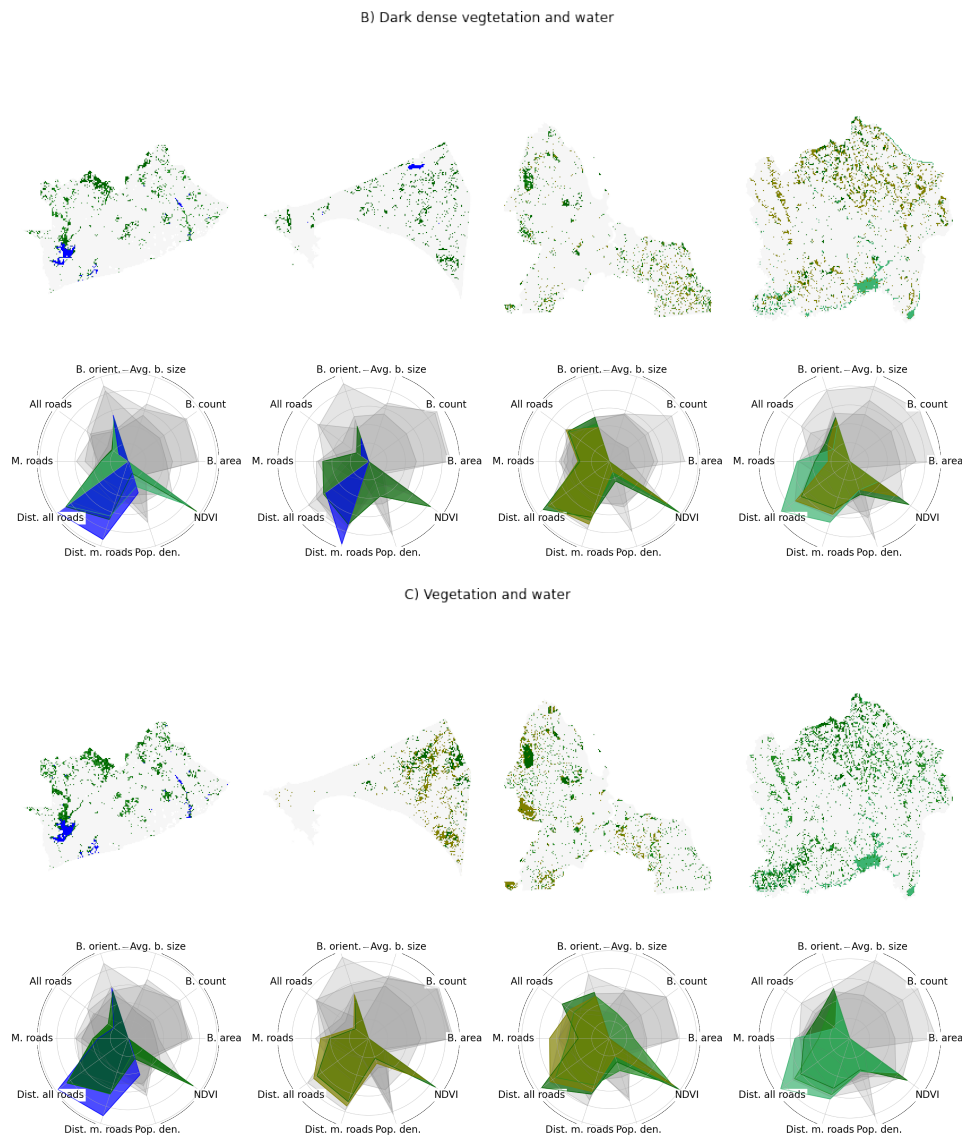


Figure 6.11: Water and vegetation clusters for Approach B (top) and Approach C (bottom). Each set of four city maps from left to right: Accra, Dakar, Dar es Salaam and Kigali. The radar charts depict the built and natural environment and demographic characteristics of each cluster. Similar shapes on the radar plots indicate that the clusters are capturing similar city environments based on demographic data and data on the built and natural environment. The different shades in colours represent different clusters. Each environmental and demographic variable is scaled with a quantile transformer, a non-parametric transformation to map the data to a uniform distribution with values between 0 and 1 (0.5 indicates the median value of a variable across all tiles in the entire image). Tiles with no buildings were included (as zeros) in summary statistics for building count and building area but excluded from the calculation of summary statistics for average building size and orientation so that zero is not used in the denominator. NDVI: Normalised Difference Vegetation Index.

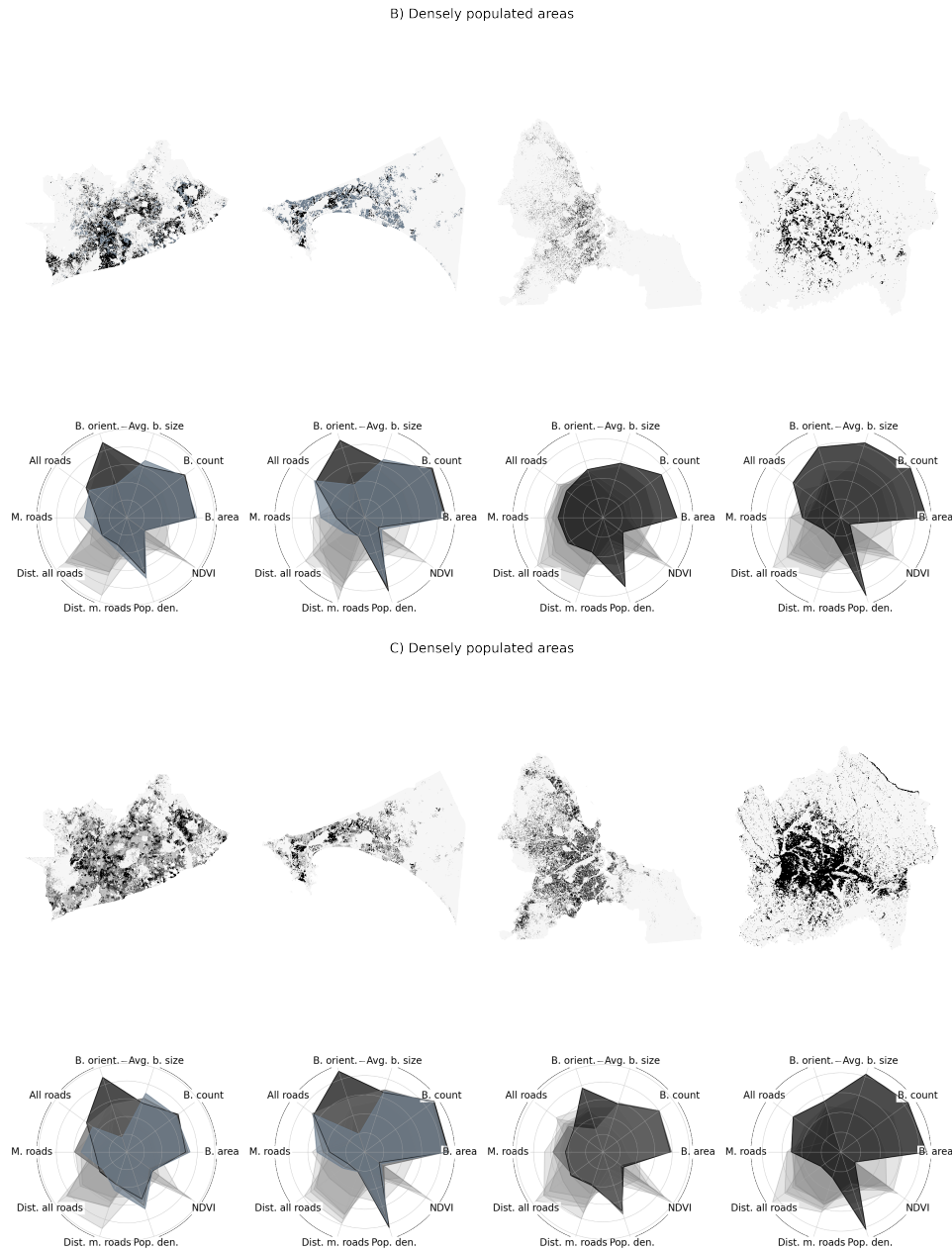


Figure 6.12: Densely populated areas clusters for Approach B (top) and Approach C (bottom). Each set of four city maps from left to right: Accra, Dakar, Dar es Salaam and Kigali. The radar charts depict the built and natural environment and demographic characteristics of each cluster. Similar shapes on the radar plots indicate that the clusters are capturing similar city environments based on demographic data and data on the built and natural environment. The different shades in colours represent different clusters. Each environmental and demographic variable is scaled with a quantile transformer, a non-parametric transformation to map the data to a uniform distribution with values between 0 and 1 (0.5 indicates the median value of a variable across all tiles in the entire image). Tiles with no buildings were included (as zeros) in summary statistics for building count and building area but excluded from the calculation of summary statistics for average building size and orientation so that zero is not used in the denominator. NDVI: Normalised Difference Vegetation Index.

To further understand how approaches B and C differ, I investigated cluster changes between the approaches and report the z-scores for each urban characteristic per cluster for each approach separately. In this context, the z-score is used to statistically quantify each urban characteristic per cluster, indicating how many standard deviations a characteristic is from the overall mean across all clusters. I further created Sankey plots for each city to visualise the cluster changes between approaches B and C as depicted in Figures 6.14 to 6.16. When comparing approaches B and C of the Accra analysis, the most consistent clusters corresponded to the natural environment. Specifically, these were the clusters capturing light vegetation and water bodies (*cluster 3* and *cluster 6*). Approach C yielded three distinct densely populated clusters, compared to the two derived from Approach B, primarily distinguished by differing building orientations. Assessing these through the MAD score revealed that Approach B formed more homogeneous clusters, indicating less deviation of these clusters from the overall cluster median. Consequently, naming and interpreting clusters resulting from Approach B for Accra was more intuitive relative to those generated from Approach C. In addition, I computed the NMI as defined in Equation 6.1, which yielded a value of 0.41.

For the Dakar analysis, the NMI between Approach A and C amounted to 0.40. Even though this was a rather moderate overlap, both approaches picked up similar urban characteristics. For example, Approach B separated the densely populated areas into two clusters, whereas Approach C split up one of these clusters (*cluster 4*) into two separate clusters (*cluster 3* and *4*) that differed mainly in building orientation. This also occurred in the analysis of Accra. Some city-specific clusters such as two agriculture clusters (*cluster 1* and *cluster 7*), one capturing trees, the other one low-level agriculture were merged together into a more general cluster that captures agriculture (*cluster 7*). Furthermore, clusters that captured very homogeneous landscapes of Approach B, such as the pink *Lake Retba* and sandy areas (*cluster 2*) were also merged together into one cluster in Approach C (*cluster 1*).

The difference in cluster membership between Approach B and C was greatest in Dar es Salaam, where the NMI between approaches amounted to 0.14. The cluster assignments plotted on top of the Dar es Salaam map (Figure 6.15) clearly showed this. The clusters of Approach B seemed very heterogeneous or noisy compared to the map showing the cluster results of Approach C. This difference was also represented in the MAD score reported in Table 6.5. The densely populated areas cluster in Approach B did not capture any difference in building orientation, whereas Approach C captured two distinct clusters capturing densely populated areas with either high or low building orientation. Approach C further clustered the densely vegetated *Pande Game Reserve* in the north of the city boundary and the shrubby area of *Kwembe* in the west of the city boundary into two separate clusters while Approach B assigns these parts to two mixed clusters. Both factors could explain why Approach C had a lower MAD score compared to Approach B.

The results of approaches B and C of Kigali were more intuitive to follow. The NMI between Approach B and Approach C for the city amounted to 0.39. Similar to the Accra analysis, the algorithm identified and separated densely populated areas and vegetation into separate clusters in both approaches. While Approach B captured two clusters with relatively high building area (*cluster 0* and *cluster 5*), Approach C grouped these together into one densely populated cluster (*cluster 7*). Other mixed clusters stayed the same, such as *cluster 1* and *cluster 7* for Approach B and *clusters 4* and *cluster 3* for Approach C.

Hyperparameter experiments

Upon investigating the outcomes of Approach A with varying cluster numbers ($K=[8,12,16]$), I observed that a learning rate of 0.0001 consistently yielded the lowest MAD scores. This learning rate mirrors the optimal rate identified in the single city clustering research, as discussed in Chapter 4. For Approach B, however, the learning rate necessitated adjustments on a city-by-city basis. For Approach B, the learning rate was adjusted per city and ranged widely. For Approach C, a learning rate of 0.00001 was best. All learning

rates are listed in Table 6.6 and MAD scores for all different approaches and learning rates are further listed in Appendix Section A.5.

Approach	K	City	Learning rate
A	8	All cities	0.00001
	12	All cities	0.0001
	16	All cities	0.0001
B	8	Accra	0.0001
	8	Dakar	0.00001
	8	DeS	0.01
	8	Kigali	0.00001
C	8	All cities	0.00001

Table 6.6: Selected learning rates across all approaches. For each model, I tested a set of different learning rates (0.01, 0.001, 0.0001, and 0.00001) and chose the learning rate with the lowest MAD score to present the results.

6.4 Discussion

The results of this chapter revealed that two out of the three approaches successfully identified distinct clusters that captured the diverse and complex patterns of urban environments, encompassing both single- and multi-featured environments. Approaches B and C, in particular, identified clusters that were shared between cities as well as city-specific and relevant for urban planning and policy-making. However, one of the approaches failed to achieve this comprehensive representation. The joint one-stage approach (Approach A) picked up some variances within cities, however, did not generalise well across cities and clustered images of Accra separately from the other cities. One possible explanation for this discrepancy could be variations in the pre-processing conducted by Maxar, to which I did not have direct insight. The Accra images were obtained through their Open Data initiative, while the images of the other city were acquired through their commercial branch, as detailed in Section 3.5.1. Notably, the pre-processing of images from the Open Data initiative exhibited issues with cloud coverage, suggesting a less meticulous pre-processing compared to the commercially purchased images. Despite visual analysis and scrutiny of metadata, no evidence was found to attribute the clustering differences to specific times of the day, seasons or other visible differences in the data.

The separate models (Approach B), although more time intensive, provided city-specific insights that could be valuable for urban planners. For example, the clustering approach identified empty sandy land or agricultural land, which could be useful for predicting city growth and productivity. While I found representative clusters, matching clusters between cities and retraining the network for each city is labour and time intensive. The joint approach (Approach A), on the other hand, was more efficient in terms of training time but returned results that were hard to interpret. The approach split off one city, which was the images that I acquired through Maxar's Open Data initiative described in Chapter 3. The separation could be due to commercial image processing, which I did not have any insight into, and would require further investigation.

The two-stage approach (Approach C) is the most time and computing-efficient. This is because the CNN does not require retraining, regardless of the number of clusters tested in the subsequent analysis. It showed that cities shared some features between them, while also exhibiting city-specific intricacies. I found that Approach C identified similar patterns to Approach B, making it a promising alternative. Some city-specific intricacies were lost, likely due to underrepresentation in the combined datasets. However, this is a trade-off that results in enhanced comparability between cities. At the same time, more specific clusters emerged through learning the features across cities. For example, in Accra, Dakar and Dar es Salaam, Approach C picked up densely populated areas with a specific building orientation that in Approach B was mostly clustered together into one cluster.

The clustering technique presents a promising tool for understanding variations in satellite images across various cities, particularly in regions where data availability is scarce. It thus stands as a useful aid for policy and decision-making. Nonetheless, there exist limitations in the current analysis, largely pertaining to geographic and temporal data availability. To address these, future studies could focus on incorporating more time points or expanding the geographical scope to additional cities. Moreover, the analysis relied on a single satellite image per city, susceptible to weather and seasonal variances, and the demographic and environmental data may not exactly align with the date of the satellite image capture. Furthermore, the interpretation of clusters could have been constrained by the specific datasets used in this study, particularly those related to the built environment. Incorporating more comprehensive data on such factors could enhance cluster interpretation in future studies. Further implications for tracking sustainable urban development are outlined and discussed in the next Chapter, where I describe the clusters identified with Approach C across all study cities in more detail.

6.5 Summary

An unsupervised deep clustering approach was applied to high-resolution satellite images of four cities in SSA. Three different approaches were tested, including joint, separate, and hybrid approaches. I found that two of my proposed unsupervised deep learning approach created effective visual features for clustering satellite images. Both, the separate approach (Approach B) and the hybrid approach (Approach C) detected similar patterns in each city, as well as captured city-specific intricacies. Approach C offered a compromise between computational efficiency as well as interpretability of features between cities and capturing city-specific intricacies. The joint approach (Approach A) split off one city from the other cities. The results were evaluated with demographic data and data on the built and natural environment. The cluster analysis revealed similar patterns across all cities but also highlighted the presence of city-specific clusters.

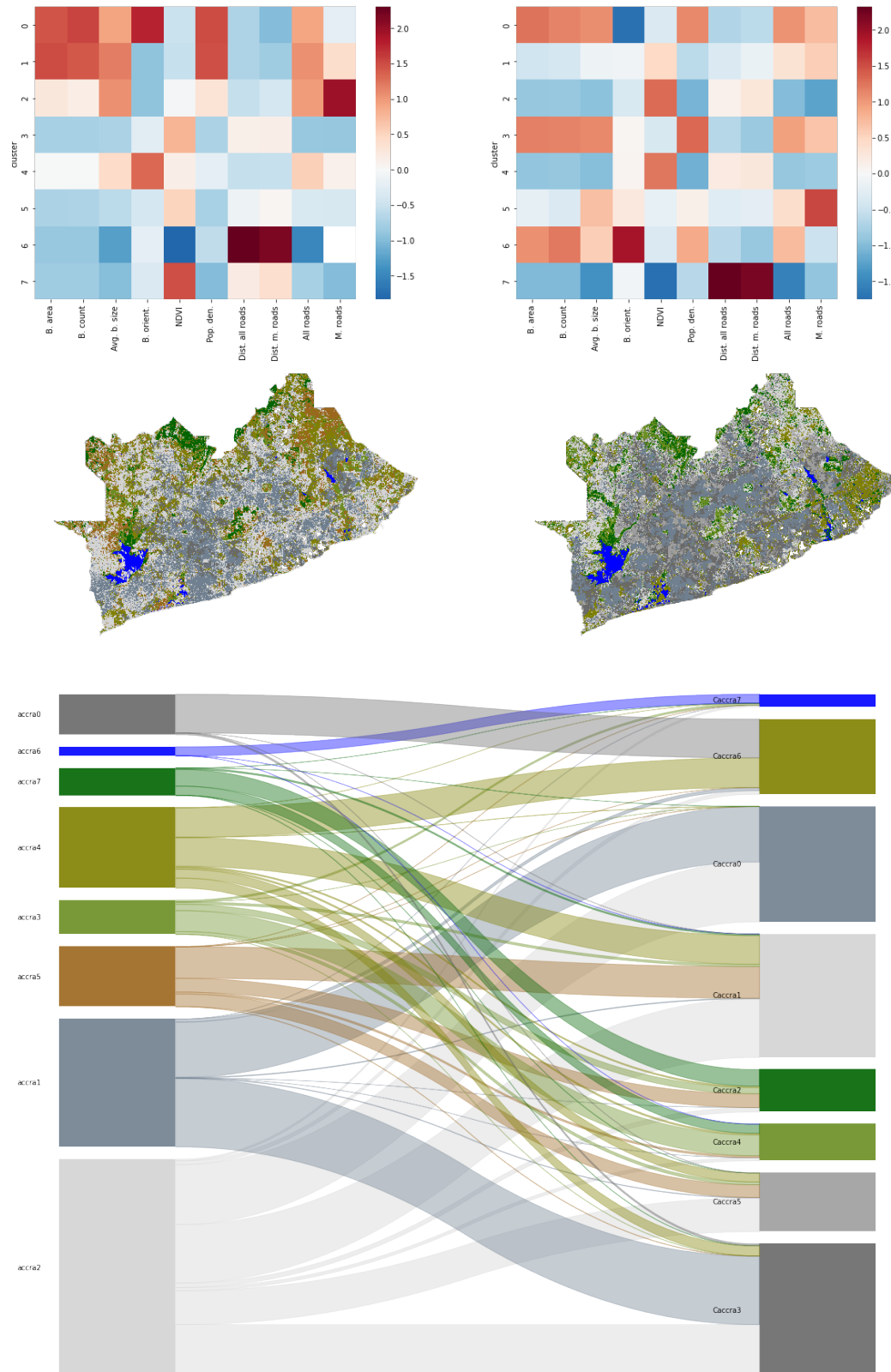


Figure 6.13: Top: z-scores of urban characteristics per cluster. Middle: city maps with each tile coloured by cluster membership. Bottom: cluster change of Approach B (left) and Approach C (right), Accra. The cluster numbers refer to the clusters listed in Table 6.3 and Table 6.4.

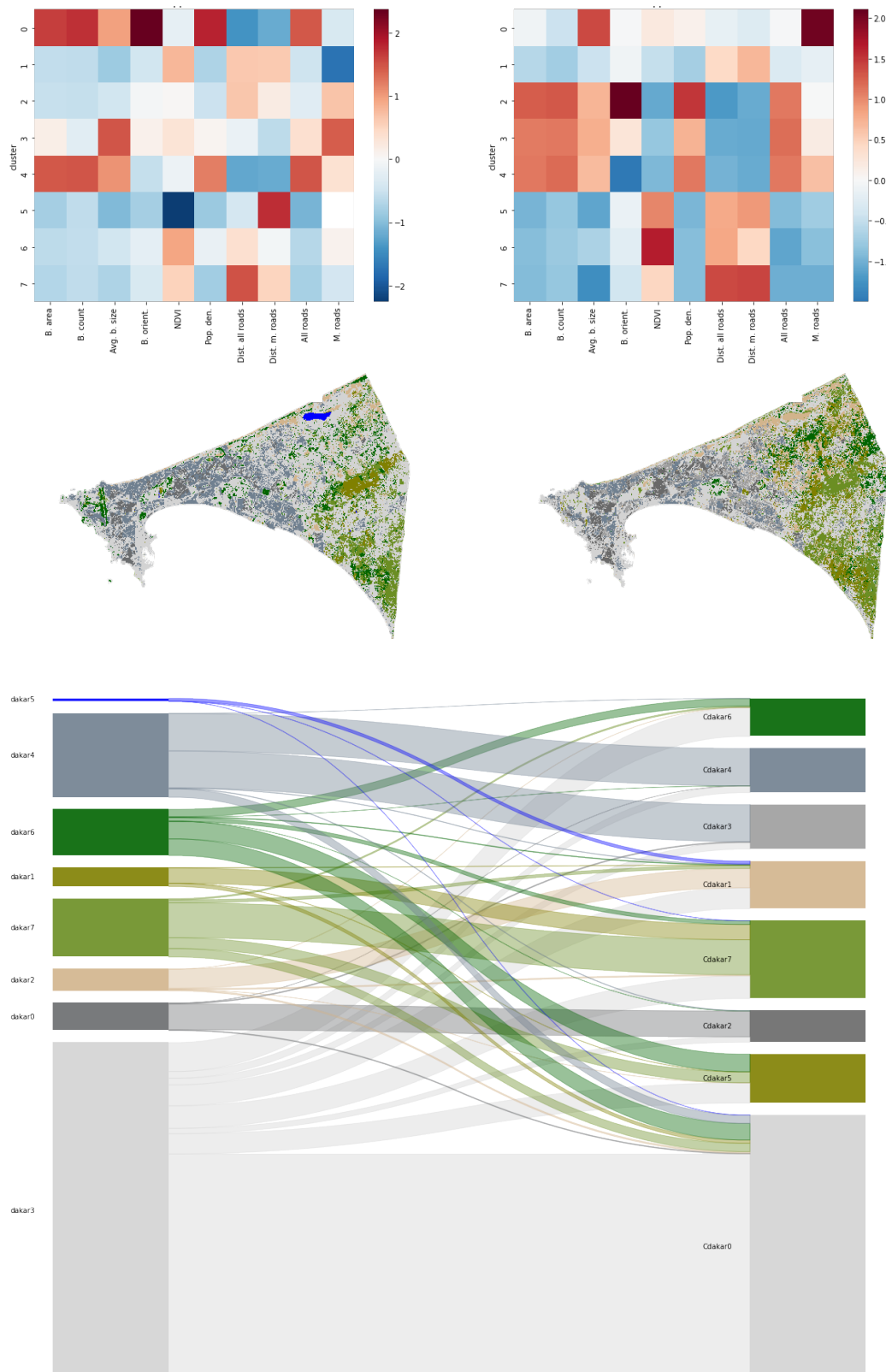


Figure 6.14: Top: z-scores of urban characteristics per cluster. Middle: city maps with each tile coloured by cluster membership. Bottom: cluster change of Approach B (left) and Approach C (right), Dakar. The cluster numbers refer to the clusters listed in Table 6.3 and Table 6.4.

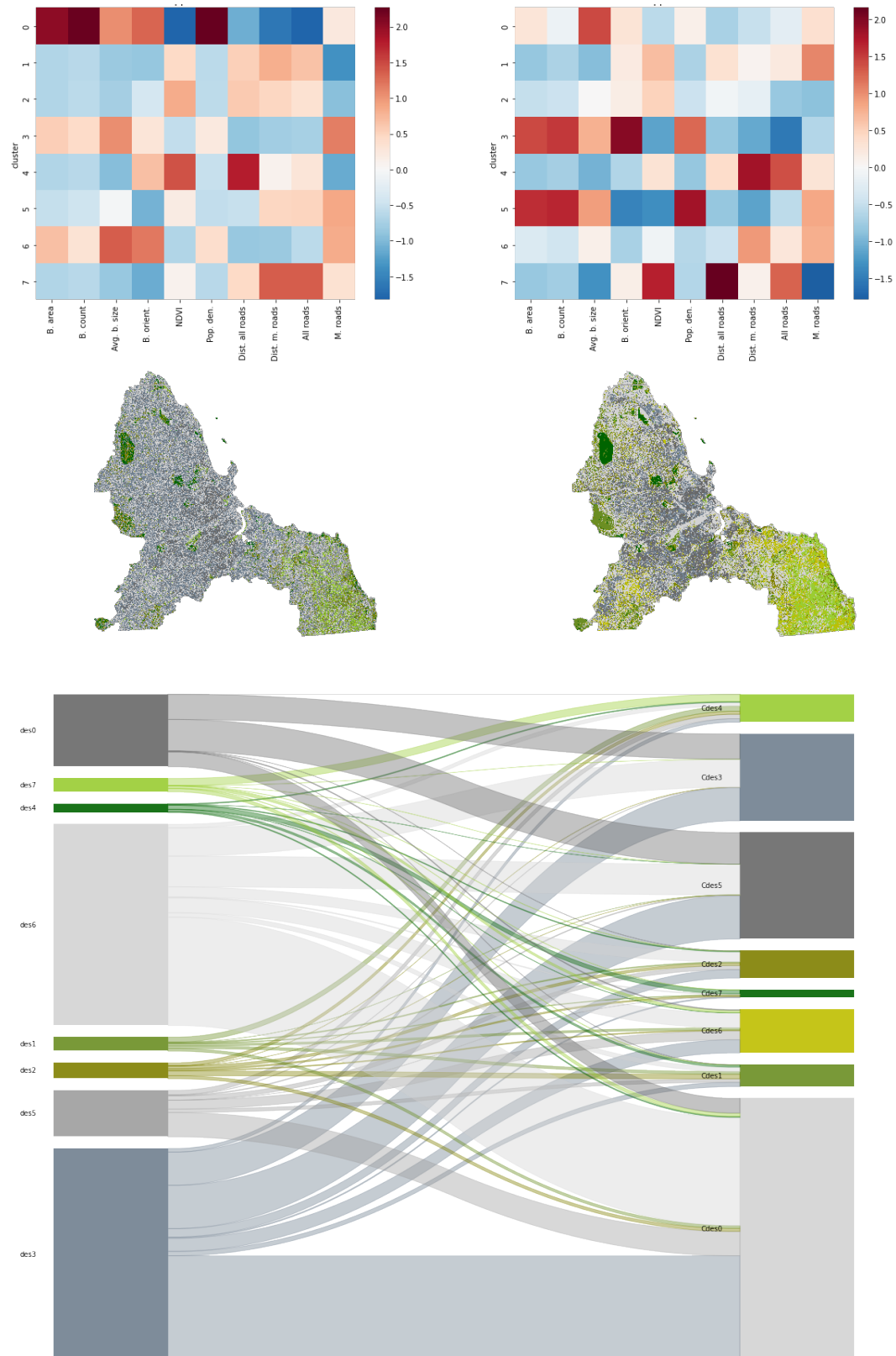


Figure 6.15: Top: z-scores of urban characteristics per cluster. Middle: city maps with each tile coloured by cluster membership. Bottom: cluster change of Approach B (left) and Approach C (right), Dar es Salaam. The cluster numbers refer to the clusters listed in Table 6.3 and Table 6.4.

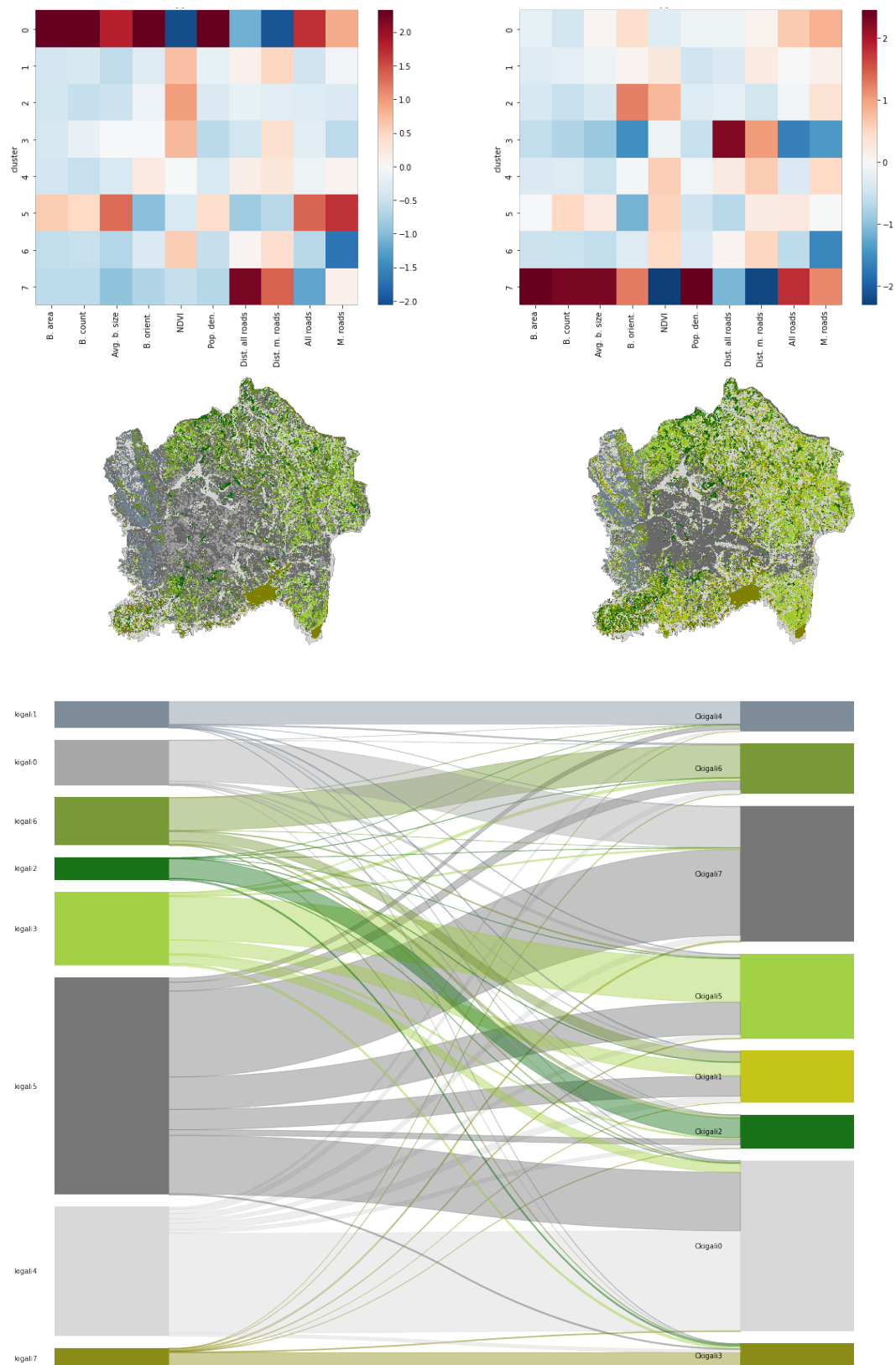


Figure 6.16: Top: z-scores of urban characteristics per cluster. Middle: city maps with each tile coloured by cluster membership. Bottom: cluster change of Approach B (left) and Approach C (right), Kigali. The cluster numbers refer to the clusters listed in Table 6.3 and Table 6.4.

Chapter 7

Phenotyping natural and built environments in cities in SSA

7.1 Overview and objectives

In this chapter, I present and interpret the cluster results of the hybrid two-stage clustering analysis described in Chapter 6. I state the objectives and describe the data, as well as external and internal cluster evaluation of all clusters across the four study cities. I fit a classifier to predict cluster membership with external urban characteristics to add further interpretability to the model.

The aim of this chapter is to apply the clustering framework to four study cities in SSA, identifying both cross-city and city-specific clusters. Using demographic, built and natural environment data, I evaluate these clusters, further delving into an internal analysis within the feature space. This exploration provides insights into how cities in SSA grow, shedding light on the shared and unique characteristics of city growth and structure in Accra, Dakar, Dar es Salaam and Kigali.

7.2 Methods

7.2.1 Data and pre-processing

The images and data on urban characteristics used for the analysis are described in Chapter 3. I follow the same pre-processing procedure as outlined in Chapter 6. I used 256 x 256 pixel satellite image tiles of four cities SSA, obtained urban characteristics of all four cities and computed per tile averages.

7.2.2 Hybrid two-stage approach

The analysis presented in Chapter 6, where I explored different methodologies for clustering satellite images across multiple cities, demonstrated that both the separate and hybrid approaches successfully identified meaningful clusters that were similar across various cities, while concurrently encapsulating city-specific nuances. Approach C struck a balance between computational efficiency and the interpretability of features across cities while preserving city-specific intricacies. Moreover, Approach C identified clusters with minimal deviation in urban characteristics, as indicated by the MAD score. Given these advantages, I continued the analysis with the hybrid approach.

As shown by Figure 6.2 in Chapter 6 as Approach C, I implemented a two-stage feature learning and clustering of the image representations. In the initial step, the number of clusters utilised for feature learning in the DeepCluster algorithm, denoted by k , was fixed at 1000. This decision was made to accommodate the extraction of a wide array of features and clusters across multiple cities. Following this, the extracted features were reduced to 256 dimensions with a principal component analysis and ℓ^2 normalised prior to their clustering. This final clustering phase was performed individually for each city, using the designated number of clusters, K . The training time for 20 epochs was approximately

72-76 hours and I trained the model on 3 RTX6000 GPUs with 250GB memory.

7.2.3 External analysis of clusters

I used data on the built environment, water, vegetation, and population (as described in Chapter 3) to quantify the characteristics of the image-based clusters and name the clusters based on what they represent. Figure A.7 to Figure A.10 depict the clusters plotted on the respective city map. In addition to the demographic data and data on the built and natural environment, I further compare the cluster results to the ESA land cover maps described in Section 3.6.2 and report the results in Section 7.3.1.

Prediction with external variables

I used the same data on the built environment, water, vegetation, and population to analyse the features of image-based clusters. For this, I employed the XGBoost machine learning algorithm [249] to predict cluster membership per city with all urban characteristics, similar to the analysis in Chapter 4. I trained a separate model for each study city. The data of each city was split into a 70:30 training-testing ratio, ensuring all clusters were equally represented. I calibrated the XGBoost gradient boosting classifier with accuracy as the primary evaluation metric. I used a Bayesian optimisation approach to tune hyperparameters using the *Hyperopt* library [252]. The accuracy of each classifier is listed in Table 7.1. The moderate to high accuracy levels indicate that the environmental and demographic variables incorporated into the SHAP analysis might only encompass a subset of the visible features. Other potential visual indicators, for which I did not have urban data, could include aspects like vehicles, distinct architectural features, specific vegetation types, and varied terrains [253].

I attained the SHAP values [251] to quantify the importance of each variable in cluster prediction. It is important to mention that while tree models handle correlated predictors

City	Accuracy
Accra	0.6484
Dakar	0.6401
Dar es Salaam	0.7541
Kigali	0.5106

Table 7.1: Accuracy scores of the classifiers trained on urban characteristics to predict specific cluster membership for each city.

well, SHAP importance scores might be dispersed among such features, potentially leading to an underestimation of their individual importance.

Comparison to ESA land cover maps

To understand how the clusters compare to global land cover maps, I used the 2021 world cover dataset by ESA [235] as described in Section 3.6.2. I downloaded the land cover maps of the specific country, cut out the city shapefile and re-projected the maps into the same coordinate reference system. Furthermore, I spatially joined the datasets for every city. To visualise which land cover labels are most prevalent within each cluster, I created co-occurrence plots normalised by cluster. I report the results in Section 7.3.1.

7.2.4 Internal analysis of clusters

In order to ascertain the degree to which DeepCluster successfully discerns clusters within and across cities based on intermediate image representations, I examined cluster associations within the PC-reduced feature space. This involved deriving the 256 principal components from image representations extracted from the penultimate fully connected layer of the CNN. These components collectively explained 99.5% of the variance of the 4096 features that the CNN utilises to represent each image. Since the image representations are learned across cities in the two-stage approach discussed in Section 7.2.2, the cluster comparison can be done across cities in the feature space. The uniformity of

tiles within each cluster in the feature space was assessed using the intra-cluster distance (Equation 4.3), which is the average distance from tiles to the corresponding cluster centroid. Cluster separability was computed using the inter-cluster distance, as defined in Equation 4.4. I created clustermaps (Figure 7.12 and Figure 7.13) to visualise how close and far away the clusters are in the feature space. The dendrogram is calculated using the Euclidean distances between cluster centroids.

7.2.5 Sensitivity to number of clusters (K)

I described the main phenotypes across all four cities by setting the cluster number to 8, as previously defined and discussed in Chapter 5 and Chapter 6. However, to examine how the clusters change with varying cluster numbers for each city, I also constructed clustergrams showing cluster development from ($K=1$ to $K=20$). A clustergram is a visualisation tool for high-dimensional data introduced by Schonlau [287]. This type of plot provides a graphical representation of the PC-reduced data over varying cluster numbers that can help highlight patterns or groups within the data set. The clustergrams and analysis are shown and reported in Section 7.3.3. I further tracked the CH score (Equation 5.1 in Chapter 5) and the silhouette score. The silhouette score is an additional measure of cluster quality. It indicates how similar a tile is to its own cluster compared to other clusters. The score is computed for each individual tile and ranges from -1 to 1, where a high value indicates that the tile is well-matched to its own cluster and poorly matched to neighbouring clusters.

The silhouette score, $s(i)$, for each tile i is defined as follows:

$$s(i) = \frac{b(i) - a(i)}{\max\{a(i), b(i)\}} \quad (7.1)$$

where $a(i)$ is the average distance from i to the other points in the same cluster and $b(i)$

is the minimum average distance from i to points in a different cluster, minimised over clusters.

The overall silhouette score is then the average of all individual scores for each tile:

$$S = \frac{1}{N} \sum_{i=1}^N s(i) \quad (7.2)$$

where N is the total number of observations (image tiles). This score provides an overall assessment of the clustering solution. The figures showing the changing CH and silhouette scores across varying K are reported in Section 7.3.3.

7.3 Results

7.3.1 Clusters of the urban environment

Figure 7.1 through Figure 7.4 show the clusters of the two-stage approach with $K=8$ for each city plotted on the respective map of each city. The clusters were named and coloured based on the type of environment that they represent. The clusters can be largely grouped into urban characteristics such as vegetation, water or densely populated areas, or more complex ones that capture a mix of urban characteristics. Figure 7.5 shows the proportions of cluster membership per city, the colours match the maps in Figure 7.1 to Figure 7.4.

The framework identified densely populated areas across all cities, typically aligning with metropolitan centres hosting markets and transport hubs. These densely populated clusters gradually thin out as they extend from the city centre. However, pockets of dense populations in regions removed from the urban core were present, typically clustered around major roads or trade areas. Visual analysis of satellite image tiles confirmed these as informal settlements characterised by high population density, small buildings, and

corrugated iron roofing. In Accra and Dakar, three distinct clusters represented densely populated areas, with the main discriminator being building orientation. Dar es Salaam, on the other hand, had two clusters capturing densely populated areas, distinguished by high and low building orientation. The main difference in these clusters was building orientation and alignment of buildings. Kigali presented a greener landscape with only one densely populated cluster, meaning there was no discrimination between building orientation. In cities where building orientation served as a factor for cluster assignment, there is typically a history of colonial city planning as discussed in Section 2.3.1. Unlike these cities, Kigali exhibits a more scattered housing layout. This lack of neighbourhoods with systematic building orientation could also be attributed to the city's varying elevation and geographic nuances.

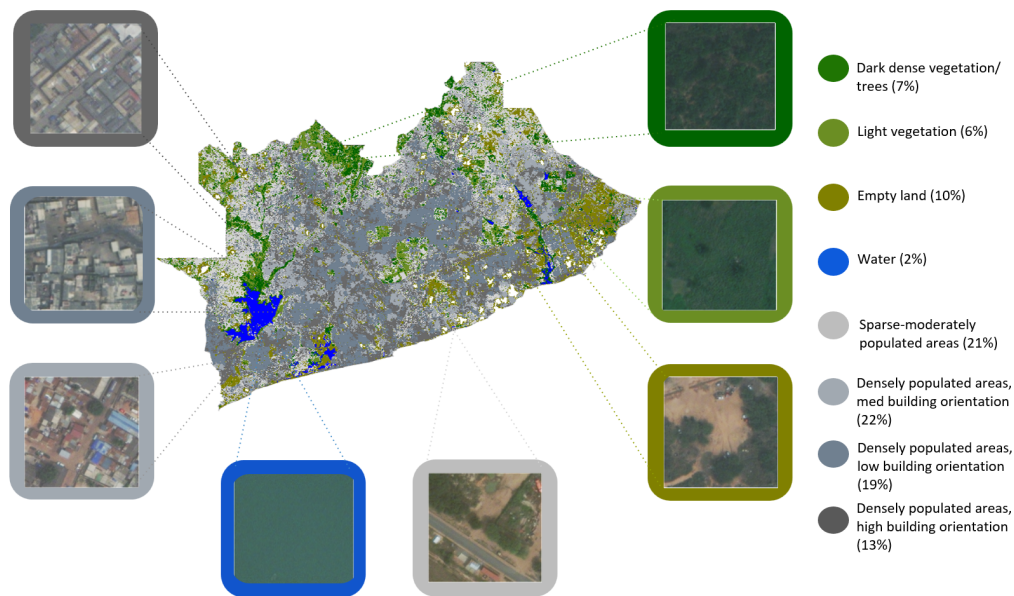


Figure 7.1: Two-stage multi-city clustering with $K=8$. Clusters plotted on the map of Accra.

Aspects of vegetation, water and agriculture were captured across all cities. These clusters varied depending on the cities' climate and city boundary size. These clusters were mostly located outside the dense city centre and next to peri-urban areas and sparsely populated areas. The vegetation and agriculture clusters also captured urban green patches and urban agriculture in isolated cases within the densely populated metropolitan areas across

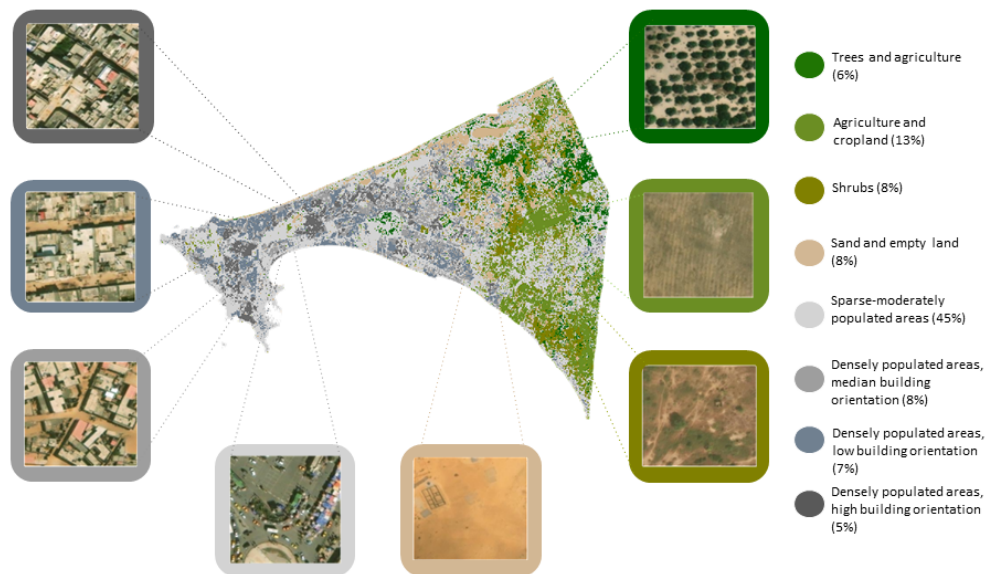


Figure 7.2: Two-stage multi-city clustering with $K=8$. Clusters plotted on the map of Dakar.

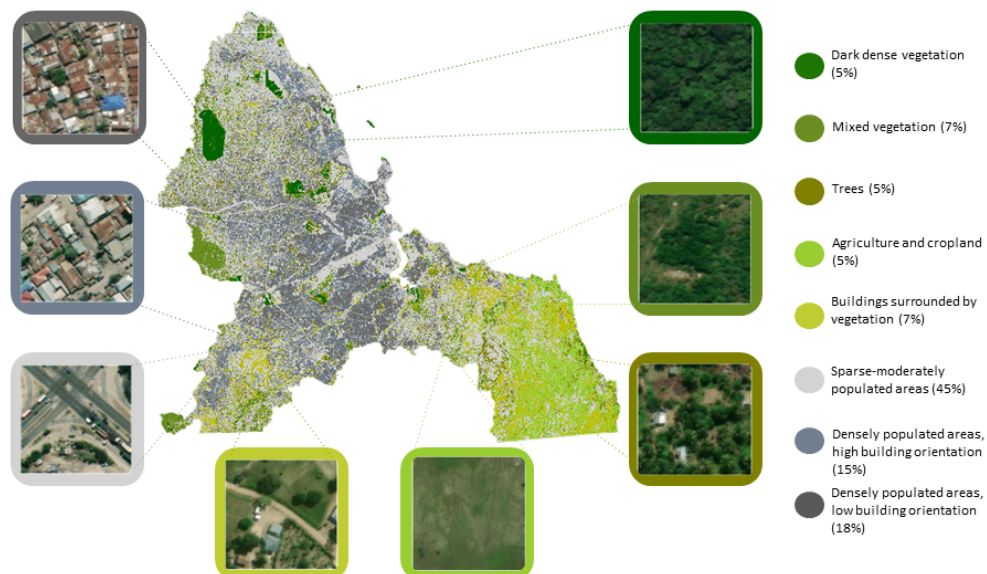


Figure 7.3: Two-stage multi-city clustering with $K=8$. Clusters plotted on the map of Dar es Salaam.

all study cities. This was more common in tropical climate zones, especially in Kigali and Dar es Salaam. As detailed in the Data Chapter (Chapter 3), Dakar's climate, which is drier compared to the other cities, was evident in its cluster formations. In Dakar, the cluster capturing the highest vegetation levels encompassed tree-filled and agricultural lands that were not as lush as the dense vegetation clusters in the other cities. Instead

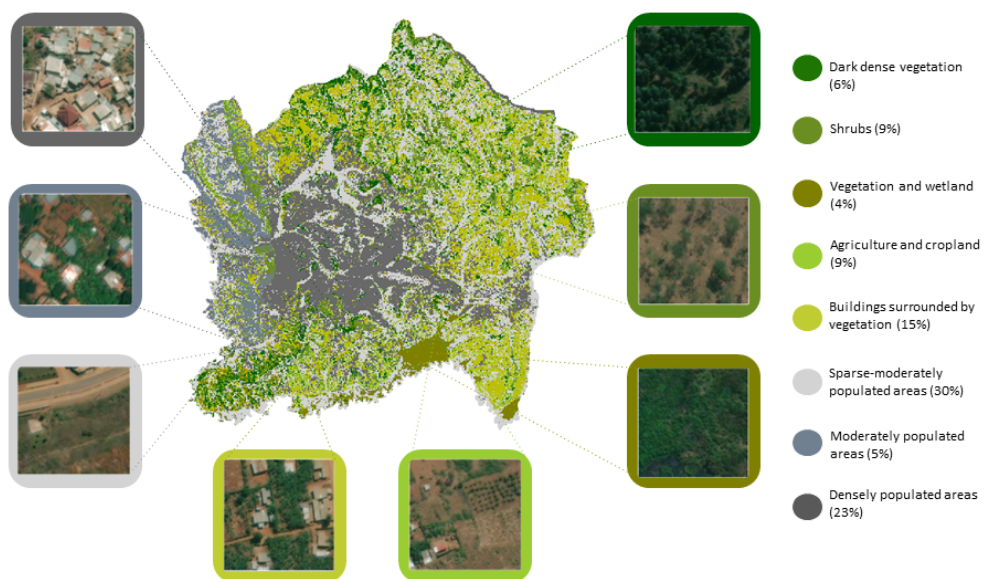


Figure 7.4: Two-stage multi-city clustering with $K=8$. Clusters plotted on the map of Kigali.

of an additional vegetation cluster in Dakar, the algorithm assigned tiles with sand and empty land into one cluster. Both Dar es Salaam and Kigali exhibited a significant proportion of agricultural land or cropland, a point which I delve deeper into when I compare the clusters with the ESA land cover maps in Section 7.3.1.

Water clusters were only separated in Accra, despite Dakar having several lakes within its city boundary. The water bodies instead were assigned to the cluster that captures sand and empty land, which all captured homogeneous landscapes in unvaried colour hues. A reason for not separating the water bodies might be the varying colours of the water bodies in Dakar, such as *Lake Retba*, which had a distinctive pink hue at the time the satellite images were taken. Other water bodies had a stark green colour in the images, especially the ones close to the *Mbeubeuss Landfill*. The landfill - one of the biggest landfills in West Africa - is a 115-hectare waste disposal site that, despite its environmental impact, provides a livelihood for a community of waste pickers who salvage and sell recyclable materials [288].

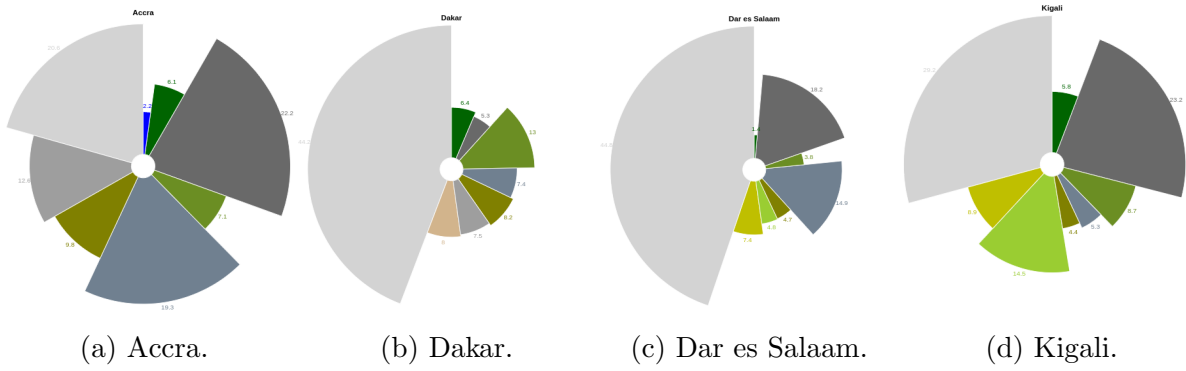


Figure 7.5: Pie charts representing proportions of cluster membership per city. Each pie section and radius depicts the proportional cluster membership. The colours of the pie sections refer to the cluster maps in Figures A.7 to A.10

The framework also identified other mixed environments across all cities. One cluster captured sparse-moderately populated areas and tiles that included built and natural environments. As depicted in the pie charts (Figure 7.5), this cluster was the largest cluster in all cities, except in Accra. However, Accra’s sparse-moderately populated areas cluster is just one percentage point smaller relative to the largest cluster (*Densely populated area, med. building orientation*). Accra had the largest proportion of densely populated areas relative to the other cities, as depicted in the pie charts (Figure 7.5). This could likely be attributed to the studied area being the GAMA, as outlined in Section 3.4, excluding the more rural parts of the Accra region.

Comparison to ESA land cover maps

Figure 7.6 shows the cluster membership compared to the ESA land cover maps. The created co-occurrence plots were normalised by cluster (row) to visualise what ESA land cover classes co-existed with the clusters. The clusters are named by what they represent and correspond to the cluster maps in Figure 7.1 to Figure 7.4. There was a moderate overlap between certain clusters and ESA land cover maps. In general, the densely populated areas clusters overlapped mostly with the built-up land cover class for all cities. The cluster analysis allowed for a more nuanced interpretation, discerning different levels of building density and building orientation. I further found that for some cities two

different ESA land cover classes visually look very similar, such as the *Trees covered area* and the *Shrub covered area* classes. Tile examples are visualised in the Appendix A.7. The plotted tiles were sampled from a shuffled list of tiles that were assigned the specific label. The co-occurrence map of Kigali (Figure 7.6) showed that a large part of the *Dark dense vegetation* cluster was labelled as *Shrubs* with the ESA land cover data. Most of the tiles did not look like shrubs, but rather more lush and green (Appendix A.7, Figure A.3). This was also very similar to the analysis in Accra, where a large proportion of the *Dark dense vegetation* cluster was labelled as *Shrubs* by the ESA land cover maps. Figure A.4 in the Appendix depicts randomly selected tiles that were assigned to the *Dark dense vegetation* cluster that visually do not indicate any shrub-like areas.

A further example of Dakar is shown in Figure A.5 in the Appendix. A large part of the *Sand/bare areas and water* cluster was labelled as *Cropland* by the ESA land cover maps. The tiles that captured *Bare areas* as defined by the ESA land cover dataset also included built environment and water, even though there is also an ESA land cover class named *Water*. The *Water* cluster in Accra was picked up by both the clustering algorithm and labelled as *Water* by the ESA land cover map. This was also highlighted by the co-occurrence plot for Accra in Figure 7.6. In general, the clusters picked up patterns that were not represented in the land cover data and added further detail and sub-groups to specific land cover classes.

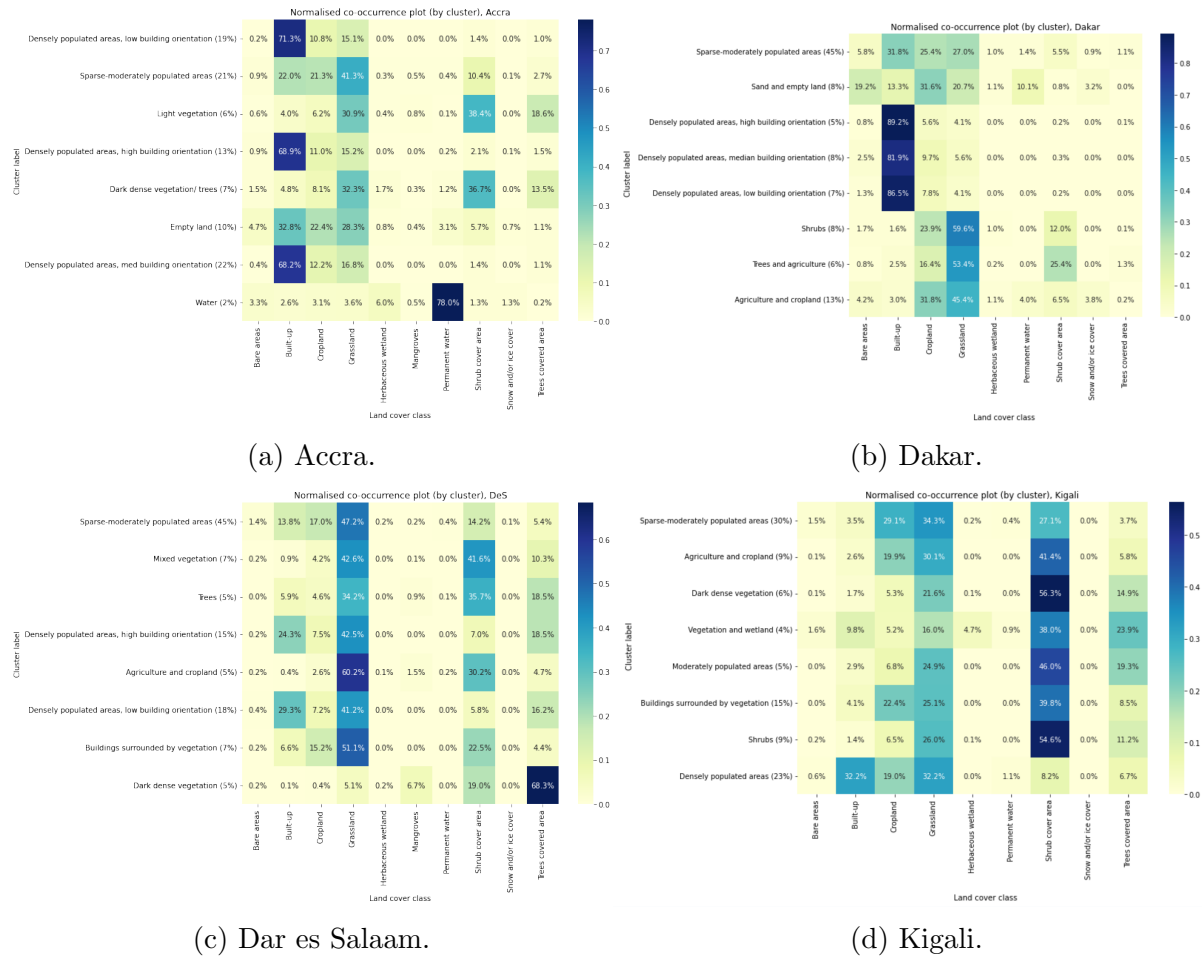


Figure 7.6: Normalised co-occurrence plots (by rows) of Accra, Dakar, Dar es Salaam and Kigali.

7.3.2 Cluster interpretation with feature prediction

I further added interpretability to the cluster results by predicting cluster membership with the external data on the built and natural environment as described in Section 7.2.4. The fitted model was used to extract SHAP values to recognise which environmental characteristics had the most impact on the cluster prediction. Figures 7.7 through 7.10 present the SHAP analyses for individual cities. For these city-specific analyses, I used SHAP values derived from the XGBoost classifier, trained distinctively using data on the built and natural environment solely for each city.

The SHAP analysis for the city of Accra showed that building orientation was an important feature for cluster membership prediction. This is corroborated by the pie chart displaying cluster membership proportions, which demonstrates that the majority of clusters were assigned to the built environment or mixed environments (Figure 7.5a). NDVI was further an important predictor for cluster assignment, especially for the natural environment clusters. The third most important urban characteristic was all road length, which is followed by building area. This is similar to the analysis in Chapter 4, however, the ranking of the top characteristics changed.

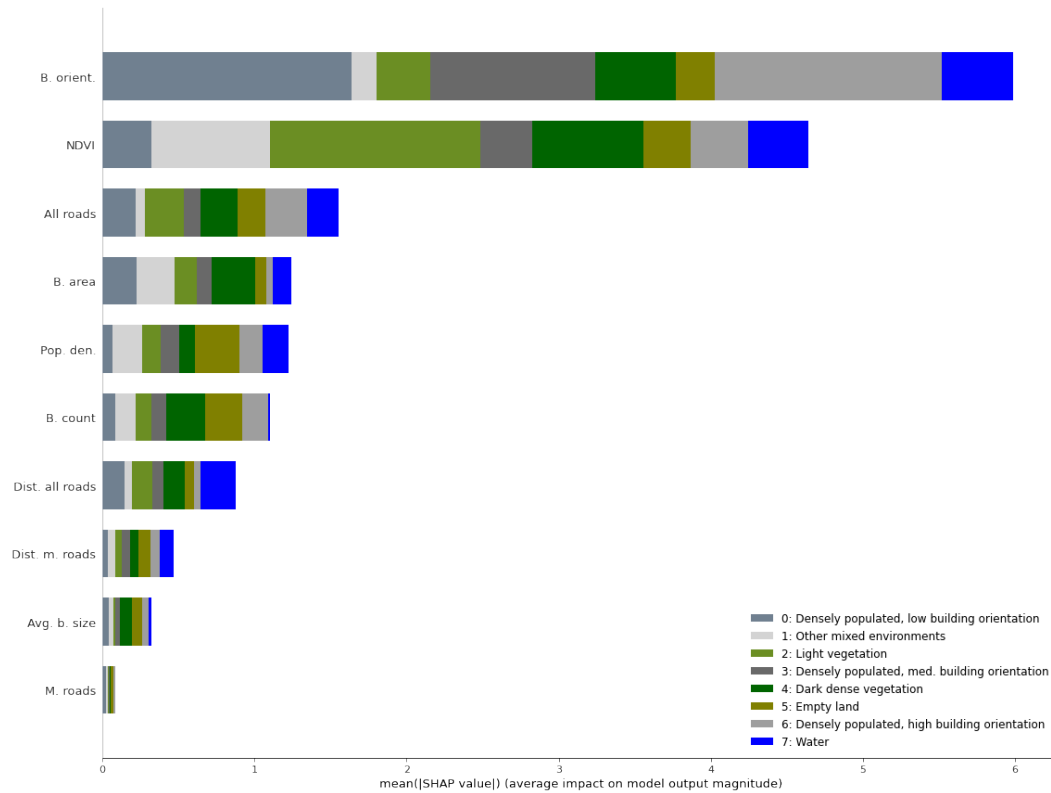


Figure 7.7: SHAP summary for Accra. The figure shows SHapley Additive exPlanations (SHAP) [251] values, obtained by fitting XGBoost [249] classifier to predict cluster membership by environmental and demographic variables. The SHAP value for each variable indicates its predictive power for assignment to various clusters and hence identifies the measures of urban form (buildings and roads), environment (water and vegetation), and population that externally differentiate clusters that were generated based on images alone. The mean SHAP values from the XGBoost classifier were calculated for each environmental and demographic variable as described in Methods. The total length of each bar, which is the mean absolute SHAP value, represents the overall importance of each variable for predicting cluster membership, and the different colours represent the importance for assignment to each cluster.

Building orientation was also very important predictor for prediction cluster membership in Dakar, especially the densely populated areas clusters. This was followed by building count and NDVI. NDVI was especially important for the natural environment clusters and building count was important for the built environment but also the agriculture cluster. The classifier, as explained in Section 7.2.3, can handle missing values (NaN values) meaning that the absence of a value (in this case building count) might be used by the classifier to make a prediction.

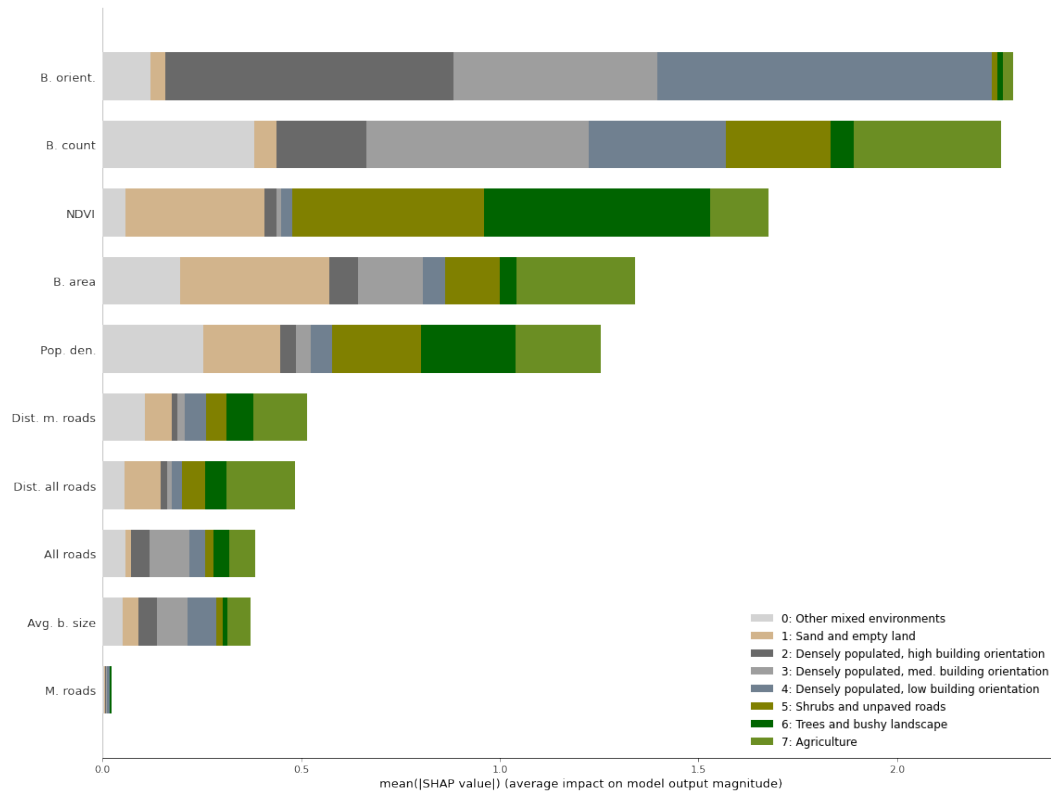


Figure 7.8: SHAP summary for Dakar.

Based on the SHAP analysis, NDVI and population density were the most important features for predicting the cluster membership of satellite images of Dar es Salaam. NDVI was especially important for clusters capturing the natural environment or agriculture and cropland. This was closely followed by building information, namely building count and building area.

The most important features for the prediction of cluster membership for clusters of Kigali were building area and NDVI. NDVI was not only important for predicting cluster assignment of clusters capturing natural environments but also clusters capturing the built environment. This was followed by the urban characteristics of population density and average building size. Overall, in all the separate models building information, NDVI and population density seemed to be important features in predicting cluster membership.

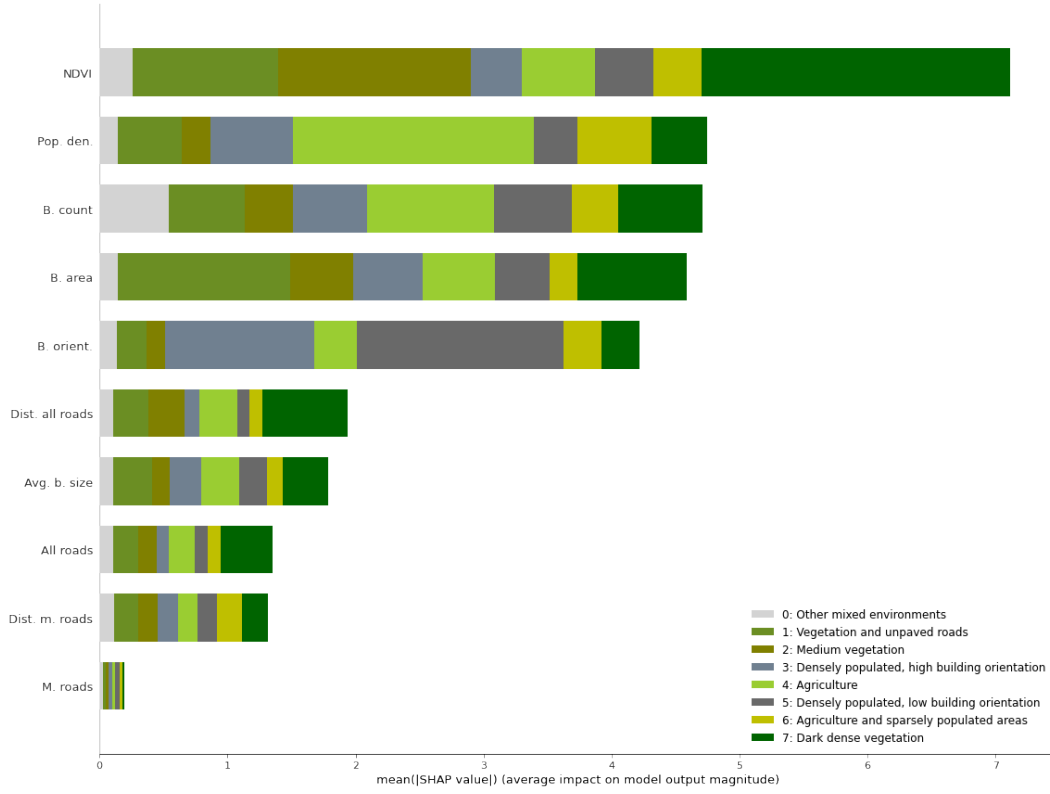


Figure 7.9: SHAP summary for Dar es Salaam.

7.3.3 Feature space analysis

My internal cluster analysis involved examining the cluster assignment of image representations within the feature space, as detailed in Section 7.2.4. In Figure 7.11, the cluster centroids were visualised based on the first, second, and third principal components. The figure reveals that the image tiles from Accra and Dakar are further apart in the feature space compared to other cities. This distinction is largely attributed to the distance between Accra’s natural landscapes (especially the water cluster) and Dakar’s high-density areas filled with boxy concrete structures [289]. Notably, Dakar does not have as many metal roof sheets as the other cities, which are commonly seen in the region [290]. Figure 7.12 showcases the clustermap of inter-cluster distances across all cities. The first row indicates which city a certain column belongs to. For a comprehensive view of intra-cluster and inter-cluster distances, refer to Appendix Figure A.11.

Figure 7.13 represents the inter-cluster distances with z-score normalisation. Here, some

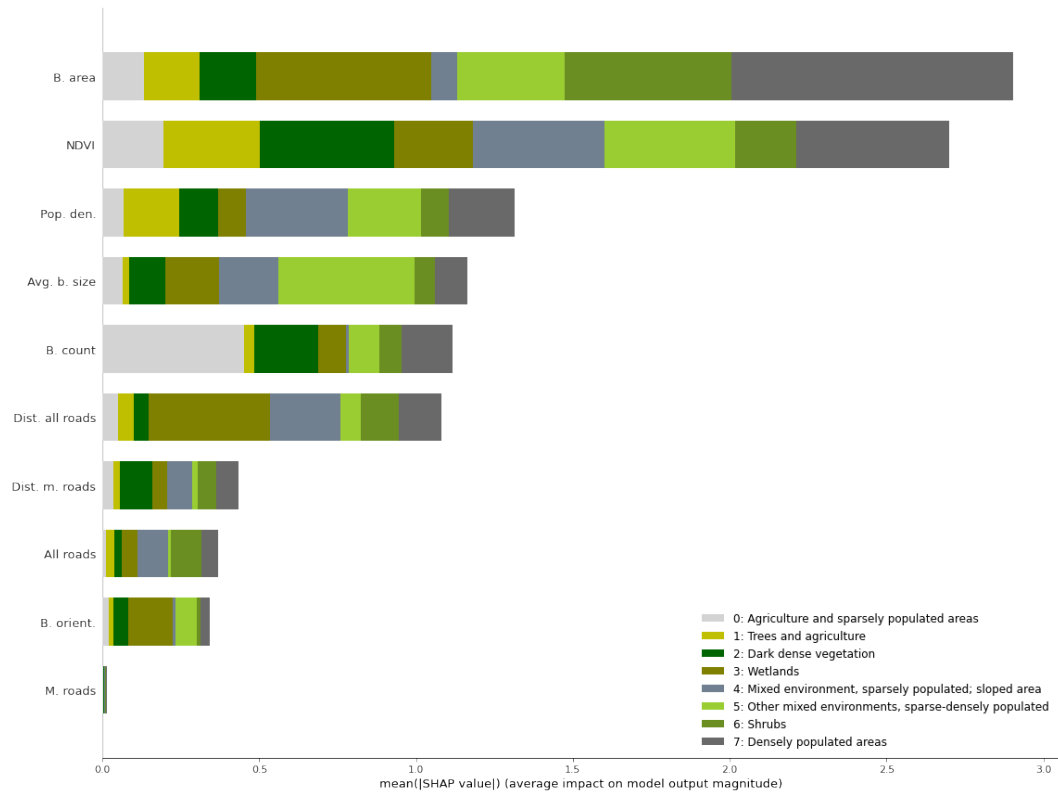


Figure 7.10: SHAP summary for Kigali.

patterns emerged that were not visible in Figure 7.12. Clusters capturing agriculture (*D7: Agri.*, *DeS4: Agri.*, *K5: Mixed and Agri.*, *DeS6: Agri.*) were all close together in the feature space. The clusters capturing mixed environments (*K0: Mixed*, *D0: Mixed*, *DeS0: Mixed*) were also closer together in the feature space. These clusters were also close to the densely populated cluster of Kigali (*K7: Dense pop.*). Inspecting the satellite image tiles, I observed that the densely populated cluster in Kigali was not as densely populated compared to clusters in other cities, both visually and in terms of absolute building characteristics. In fact, it exhibited similarities to the mixed environment clusters identified in other cities. Other clusters, like the Water cluster of Accra (*A7: Water*) were very far away from all other clusters. The built environment clusters of Dakar were further apart from the other clusters that captured densely populated areas of the other cities. Upon visual inspection of the tiles representing densely populated areas in Dakar, it was apparent that the buildings showed great differences in aspects such as height and building materials compared to those in other cities. The natural environment clusters of

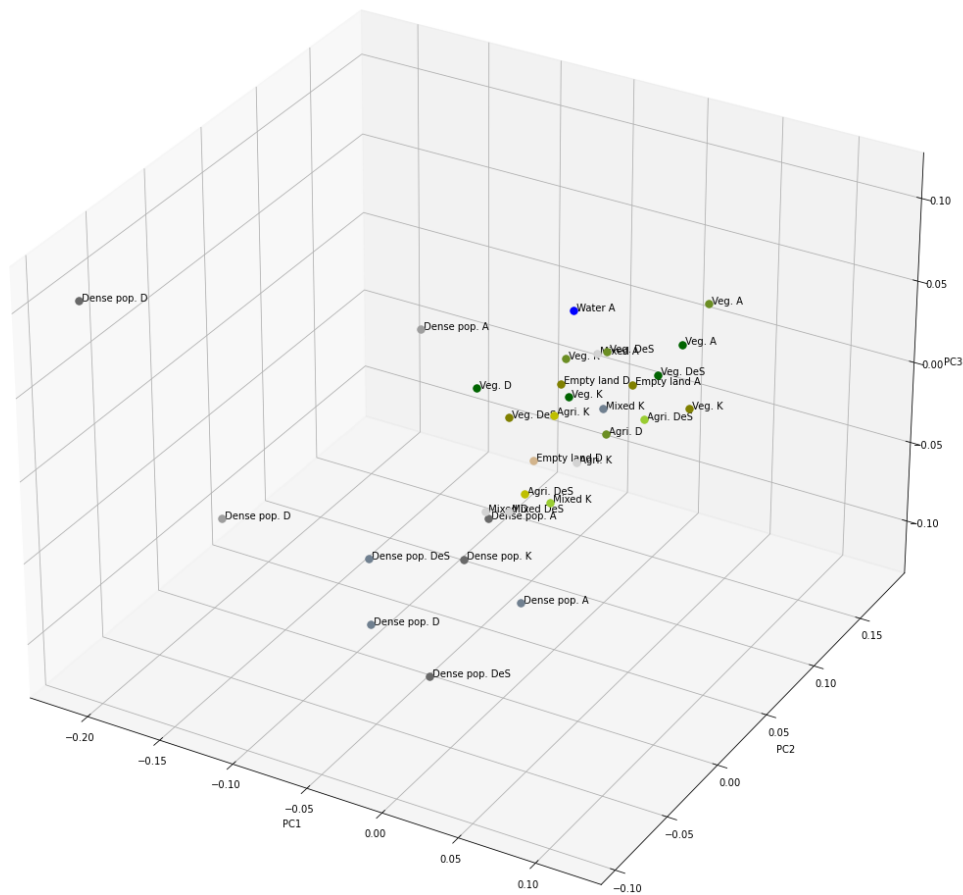


Figure 7.11: Cluster centroids seen in the CNN-extracted feature space. The cluster centroids were plotted on the first, second and third principal component of the feature space. The colours of the points refer to the maps in Figure 7.1 to Figure 7.4 and the clusters are named based on their super-cluster category. The letters refer to the study cities, A: Accra, D: Dakar, DeS: Dar es Salaam and K: Kigali.

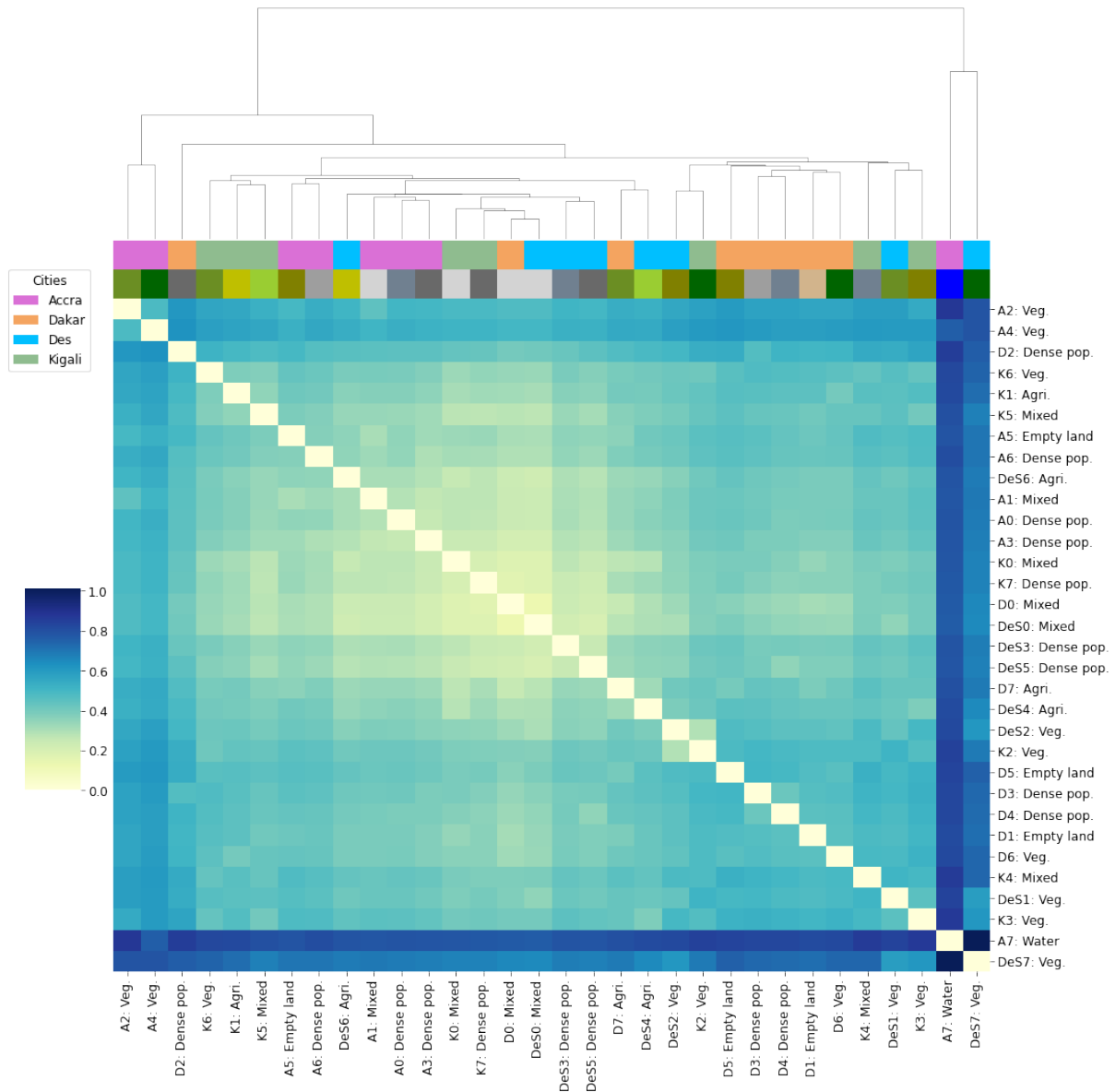


Figure 7.12: Clustermap of inter-cluster distances. The clusters are ordered by distance between cluster centroids, with the dendrogram indicating cluster distances. Light hues denote close proximities in the feature space, while blue shades indicate greater distances between clusters. The letters refer to the study cities, A: Accra, D: Dakar, DeS: Dar es Salaam and K: Kigali. The cluster numbers refer to Table 6.4.

Dakar were closer to the other cities in the feature space.

This hierarchical approach allows for grouping the clusters into larger groups of clusters based on the distances between the cluster centroids. When cutting the dendrogram of distances between cluster centres at a higher point (i.e. lower cluster number), super-groups emerge that resemble the grouping of the clusters, which I used for reporting cluster results in Chapter 6. These super-groups include vegetation, densely populated areas, agriculture, and other mixed environments. Surprisingly, the densely populated areas clusters of Dakar cluster separately from the other densely populated areas. This could be due to the dry and sandy nature of the city, which manifests in different colour ranges in vegetation and built environment compared to the other cities.

Sensitivity to cluster number

I further analysed cluster development from $K=1$ to $K=20$ per city to evaluate how the clusters split and if a different cluster number would be favourable. Figure A.7 to Figure A.10 depicts the clustergrams [287] of the image representations with varying cluster numbers. On the horizontal axis, we can observe the number of clusters. Each point symbolises the centre of a respective cluster, weighted by the first principal component to enhance the readability of the diagram. The lines that join these points, along with their respective thickness, denote the movement of observations between clusters. As a result, it aids in analysing when new clusters evolve as a division from an existing single class, or when they originate from observations drawn from two separate clusters.

The clustermaps revealed that there were a few clusters that split off very early and stayed stable over changing values for K . This observation is corroborated by Figures 7.15 and Figure 7.14, which show the CH scores across changing K and reveal analogous patterns. As previously highlighted in Chapter 5, the image representations cluster

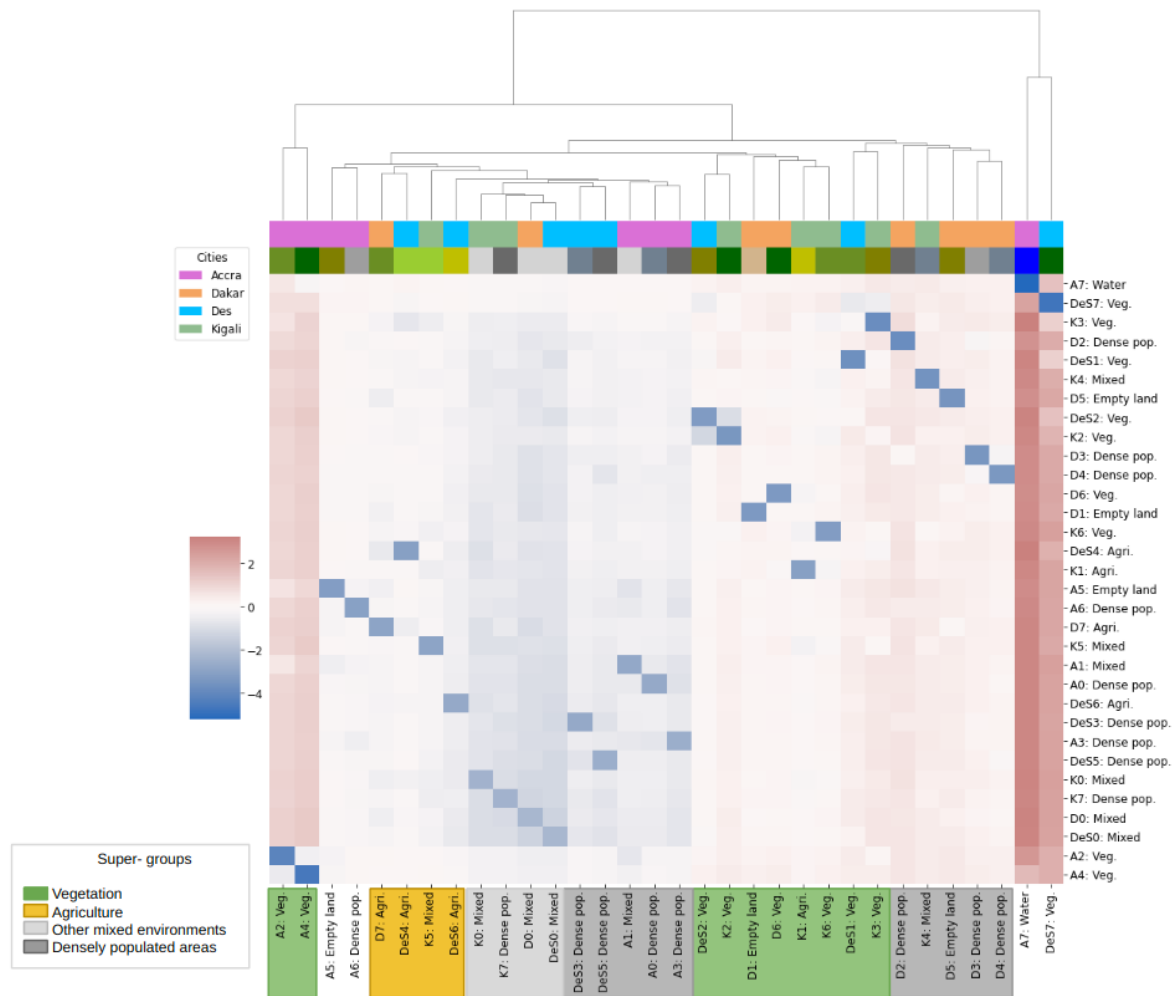


Figure 7.13: Clustermap of inter-cluster distances with z-score normalisation. Light hues denote close proximities in the feature space, while darker shades indicate greater distances between clusters. The colour scale ranges from cool (blue) to warm (red), indicating negative to positive standard deviations from the mean, respectively. The letters refer to the study cities, A: Accra, D: Dakar, DeS: Dar es Salaam and K: Kigali. The cluster numbers refer to Table 6.4.

densely within the feature space, leading to a decrease in the CH score with an increase in K . Nevertheless, across all cities, the curve started to flatten between $K=5$ and $K=8$. Simultaneously, there was an upward trajectory in the silhouette score with increasing K , following an initial descent around $K=3$. The clustergrams tell a similar story, indicating the emergence of approximately 4 to 6 super-clusters. The clustergrams for Dar es Salaam and Dakar depict well-entrenched cluster strands that primarily diverged further from previous clusters. The clustergram of Kigali is the sole exception, displaying some instability before creating more established clusters at around $K=8$. To conclude,

although larger cluster numbers yield more detailed results, they also pose a significant challenge when it comes to human interpretation. That said, this analysis suggests that for all cities, a value of $K=8$ (or similar) struck an optimal balance. It provided sufficiently granular clustering to accurately capture the structure of the data while remaining comprehensible and analysable by human researchers. This balance is crucial as it ensures that the technical aspects of the analysis do not overshadow the interpretability of the results, thus making the data actionable for decision-makers and urban planners.

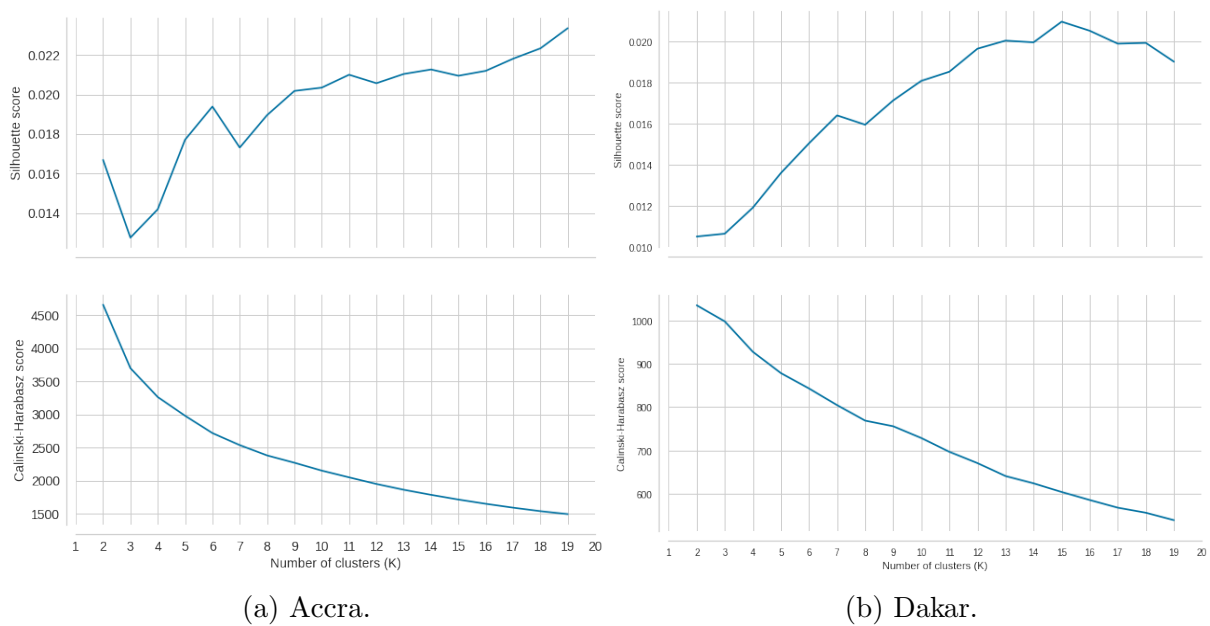


Figure 7.14: Silhouette and CH score of Accra and Dakar.

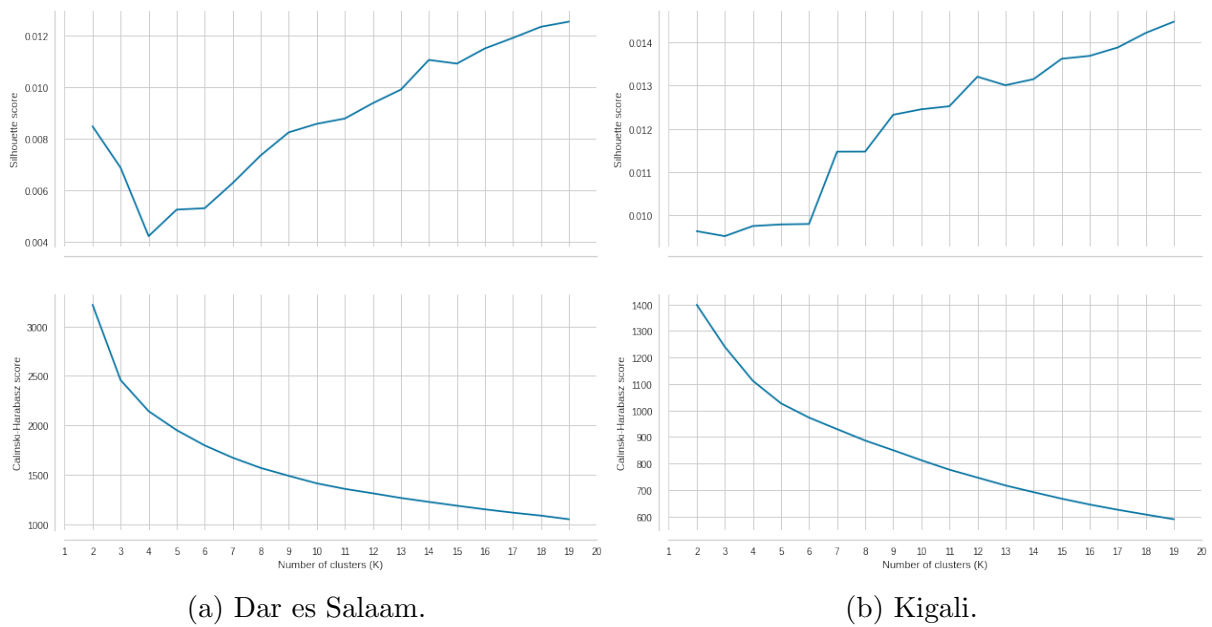


Figure 7.15: Silhouette and CH score of Dar es Salaam and Kigali.

7.4 Discussion

When we close our eyes and imagine large cities across the world, the image that frequently emerges is a skyline filled with tall, high-rise buildings. This common visualisation often reinforces the perception that cities are increasingly becoming more uniform and homogeneous across the world. Yet, cities in SSA diverge from this expectation. As explored in Chapter 2, the growth pattern of SSA cities deviates from the stereotypical upward trend; instead, they expand outwards, with the city growth often manifesting in sprawl or leapfrog development, leaving parts of the city disconnected in terms of services and transport [4].

While urban sprawl and demographic growth are characteristic of many cities within SSA, this chapter has demonstrated that there are additional shared attributes between the cities studied in this thesis. Using unsupervised feature extraction and clustering, this work highlighted that there are some common ways in which cities in Africa have grown. Notably, Accra, Dakar, and Dar es Salaam, despite their varied climates and historical backgrounds, show remarkable urban similarities, predominantly influenced by colonial legacies and visible in the distribution of their built environments. On the other hand, Kigali stands apart, being a nascent city without a pronounced colonial imprint. However, Kigali also shares certain urban characteristics with its coastal counterparts, such as intertwined urban-rural landscapes and the presence of agricultural lands at the peripheries.

A salient result of this research is the footprint of colonial urban planning, prominently observed in Accra, Dakar, and Dar es Salaam. Regardless of their distinct colonial rulers, these cities exhibit strikingly analogous urban designs. Central to each of these cities is the positioning of the city centre, which was captured by the clustering of densely populated areas. These dense pockets in the city core are adjacent to the harbour, characterised by expansive public squares and boulevards. As captured by the *Sparse-moderately pop-*

ulated areas clusters, these broad arteries, vestiges of colonial urban design, have over time metamorphosed into commercial hubs and marketplaces that have been set up along the wide roads [291]. They are vital in terms of transport and commerce, connecting the city with the parts outside the crowded metropolitan hub. However, they grapple with the challenges posed by burgeoning traffic, often failing to cater to the growing demands [292, 293]. Furthermore, the mixed environments (also captured by the *Sparse-moderately populated areas* clusters) in all coastal cities of the study further encompass industrial and commercial areas. These areas are predominantly situated in proximity to the harbour and typically within or adjacent to city centres, exert added pressure on the already taxed road network. For example, in Dar es Salaam, industrial lands are mainly distributed along the industrial corridor that encircles the city's railway infrastructure and reaches into the city centre [294]. Yet, Accra has decentralised some of the congestion by relocating its primary shipping operations to Tema Port, both Dakar and Dar es Salaam maintain their primary cargo facilities within their central metropolitan area.

In all three coastal cities, the colonial influence manifests in the architectural orientations in densely populated parts of the cities. Neighbourhoods, especially within Accra and Dakar, but also present in Dar es Salaam, exhibit grid-like road structures inherent to colonial urban design [70]. These areas are all captured by the *Densely populated areas* clusters that pick up rows of buildings that are aligned in a specific way. However, such uniform patterns diminish in the peri-urban zones, which display more organic, unstructured road and building alignments. This is represented by a larger intermix of the clusters capturing specific building orientations.

Another salient feature across these cities is the location of informal settlements. These are also captured by the *Densely populated areas* clusters in this chapter. The prevalence of informal settlements is often intricately linked to underlying land tenure patterns or land rights, which play a defining role in shaping the spatial characteristics and development of these communities [4]. Distributed as small dense pockets across metropolitan areas, these settlements often cluster near coastal zones or adjacent to wetlands [295]. In

the studied cities, these wetlands also serve as places for accessing water and sanitation. However, their positioning, especially those along coastlines or adjoining river basins, inherently subjects them to elevated flood risks [201, 296, 297].

Expanding upon the shared urban characteristics of cities in SSA, one cannot overlook the presence of affluent residential zones situated in proximity to the city centre. These upscale neighbourhoods stand out due to their expansive free-standing residences complemented by verdant gardens, and in some instances, swimming pools. Contrasting sharply with the built-up density of the city centre, these enclaves present high levels of urban green and are well connected to primary arterial roads [298]. This is represented by a mix of densely populated clusters intermixed by clusters that capture vegetation. For example, in Dar es Salaam, affluent neighbourhoods have tree canopies covering on average over 20% of the surface area, compared to less than 0.5% in poorer areas like *Namanga* [298].

The delineation of city boundaries significantly influenced the analysis. Cities with boundaries that included rural areas, such as Kigali and Dar es Salaam, exhibit large areas with agricultural lands at their peripheries, yet in Accra and Dakar these parts are a lot smaller. As determined by the clusters, most agricultural zones in all study cities are located either in rural areas or at the rural-urban interface, farthest from city centres. In some cases, densely populated pockets neighbour these agricultural or undeveloped lands. This juxtaposition illustrates the intertwined nature of urban and rural landscapes in these cities [299, 300, 301], as evidenced by the prevalent wildland-urban and rural-urban intermix in SSA cities, particularly noticeable in Dar es Salaam and Kigali [294, 302]. In addition, the peri-urban zones often feature scattered, and at times unfinished buildings surrounded by vegetation, indicative of potential urban expansion in the future. Such free standing building structures, more commonly found in peri-urban areas rather than dense city cores, were mostly captured by clusters that represent multiple features of the urban environment such as the *Buildings surrounded by vegetation* clusters in Dar es Salaam and Kigali. These incomplete structures may not only signal future urban growth, but

could also reflect financial limitations like lack of access to credit or land tenure issues, which often prevent the completion of construction projects [303, 57].

Despite the listed commonalities, the cities have a few distinctions worth highlighting. Firstly, differences related to climate and therefore vegetation. The climatic distinctions across cities manifest in their green spaces. Dar es Salaam and Kigali, in particular, abound with green patches within their urban area. Regardless of their dense urban cores, these cities retain interspersed green spaces and agricultural pockets. Accra, and especially Dakar, did not have as much urban green, and large urban green areas such as parks within the metropolitan core, were almost non-existent. Secondly, there were some architectural differences related to building materials and climate. Dakar exhibits a distinct architectural footprint. Its buildings, predominantly cubic in design and constructed from concrete, feature flat roofs often used for drying laundry [289]. This is in stark contrast to the ubiquitous metal sheet roofing in Accra and notably in East African cities, which serve as a shield against rain and sun [290]. Thirdly, there were some density differences between the three coastal cities and Kigali. Kigali's urban layout is heavily influenced by its diverse topographical features. The city's elevation variations translate to a lower building density relative to other studied cities [304, 294]. On the other hand, Dakar's urban core stands out for its very high density, a likely vestige of French colonial urban planning [70].

In summary, cities within SSA often exhibit parallel growth trajectories. While there are inherent complexities and diversities in these cities, their urban forms frequently display resemblances, predominantly resulting from shared histories and accelerated urbanisation experiences. These similarities underscore the consistent patterns of urban growth throughout the region. However, it is also essential to recognise the unique influences of geography, colonial past, and urban policies that differentiate each city's growth. As we seek solutions for urban challenges in SSA, we must approach them with a detailed

comprehension of these commonalities and disparities. The use of satellite imagery combined with machine learning presents a valuable avenue for tracing these shared urban growth patterns. Although the approach requires certain decision-making processes, such as determining the number of clusters, this chapter has shown that the clusters it generates are intuitive and interpretable, and can be used across a wide range of settings and applications, offering a robust and flexible tool for studying urban areas.

7.5 Summary

A hybrid two-stage feature extraction and clustering algorithm, as outlined in Chapter 6, was applied to high-resolution satellite images of Accra, Dakar, Dar es Salaam and Kigali. City-specific and shared clusters between cities are described with external data not used for clustering. An ensemble classifier was fitted to predict cluster membership across cities with urban characteristics to add interpretability to the analysis. The image representations were further analysed in the feature space together with cluster assignments, to identify similar and dissimilar clusters. The clusters were further analysed in an urban development context, identifying similarities and differences in how cities grow in SSA.

Chapter 8

Discussion and conclusion

8.1 Summary

I obtained high-resolution satellite images of four major cities in SSA. I further acquired demographic data and data on the built and natural environment as described in Chapter 3. First, I applied a deep learning clustering framework to Accra, Ghana as described in Chapter 4. I further investigated various model choices of the approach, including the scale of analysis and the number of clusters defined as part of sensitivity analyses, as per Chapter 5. In Chapter 6, I explored how to effectively scale and apply the deep learning clustering framework to multiple cities in SSA, in order to obtain interpretable and meaningful clusters. In Chapter 7, I applied the hybrid two-stage feature extraction and clustering approach to four distinct cities in SSA, namely Accra, Dakar, Dar es Salaam and Kigali. I further discuss similarities and differences in the resulting urban phenotypes in the four study cities, as well as in the context of cities in SSA.

8.2 Strengths and limitations

8.2.1 Strengths and novelty

My thesis makes a number of novel contributions to measuring and tracking complex urban environments. First, I developed a novel approach for using high-resolution satellite images without external data to capture variations in multiple natural and built environment of cities and can provide timely data to support sustainable and healthy urban development. The approach coherently integrates different features of the urban environment, which have traditionally been analysed in isolation [10, 94, 95, 96, 16, 97, 98, 99, 100, 101, 102, 103]. These elements are interrelated and display complex patterns across different scales which are relevant for policy implementation at different administrative levels.

The algorithm combines feature extraction and clustering, making its implementation easy and efficient. I used very high-resolution satellite images, which show objects as small as cars and trees, and therefore allow the model to use information on high-level features of the urban environment. The results were interpreted against built and natural environment and demographic characteristics that demonstrate the interpretability of the model to help its wider use. I also investigated the image representations in the feature space to better understand cluster formation and assessed the sensitivity of the approach to key choices such as tile size and cluster number. The approach used in this thesis picks up more detailed clusters than pixel-level analysis (e.g., land classification by NDVI) could do.

Moreover, I analyse different approaches for applying the developed framework to high-resolution satellite images of multiple cities to obtain meaningful and interpretable clusters. I compared the clusters across cities and identify clusters that are similar as well as clusters that are city-specific. Finally, I address how this method can inform policies related to sustainable urban development and be used in other cities in the developing

world, to help close the global data gap.

8.2.2 Limitations

Data limitations

The primary limitations of this analysis stem from constraints in spatial and temporal data availability. For future research, it is recommended to either examine additional points in time or extend the analysis to encompass more cities. As highlighted in Chapter 6, there was potentially a mismatch in pre-processing quality between the satellite image of Accra compared to the other cities. This turned out to be a limitation, since it was picked up by the clustering algorithm. Future work could therefore include repeating the analysis with images of Accra that are purchased and pre-processed in the same way. In addition I only used one satellite image that captured the city at a specific point in time, and so weather and season could have an impact on pixel intensities. While the data on built and natural environment and demographic characteristics were obtained in the same year, there could be a mismatch between the exact date the satellite image was taken and these data were gathered or estimated. Additionally, the analysis was limited by the datasets used to interpret the clusters. Further data on the built and natural environment, such as on building height or agricultural land, might have improved the cluster interpretation.

Modelling limitations

There are some limitations related to pre-processing the data, such as deciding on a tile size for the analysis. In Chapter 5, I investigated key choices based on zoning and scale to better understand cluster changes in the Accra analysis. An additional sensitivity analysis for all study cities was outside the scope of this thesis. A further limitation is the computational cost compared to a simpler pixel-level analysis. However, the framework in

this thesis is able to capture more detailed information that goes beyond what is captured by one pixel. At the same time, computing power is increasing and becoming available at low cost, which will make the analysis likely faster and more accessible in the future.

As highlighted in Chapter 5 and Chapter 6, I considered multiple model choices, such as learning rate, cluster number, and analysis approaches for the Accra analysis, but also the multi-city analyses. However, I could have investigated additional CNN network architectures for feature learning or other clustering algorithms for clustering the image representations. This, however, would have gone beyond the scope of this thesis since every model change comes with additional hyperparameter calibrations.

8.3 Impact and implications for tracking urban development and policy

The implications of this work are significant in urban development contexts, and they extend in two primary directions. Firstly, the research underlines the potential of using unsupervised machine learning and satellite imagery as effective tools in urban analysis. Secondly, it fosters a comparative understanding of urban change across cities in SSA. By analysing urban development, and identifying both similarities and differences among these cities, this thesis deepens our understanding of urban growth dynamics in the region. Consequently, it can guide the formulation of targeted urban policies that address prevalent regional trends.

8.3.1 Leveraging unsupervised machine learning and satellite images

Numerous international frameworks and institutions have underlined the urgent need for comprehensive and specialised data on urban environments [82]. However, the urban sciences that exist today are predominantly skewed towards the Global North [6]. This thesis contributes towards addressing this gap by creating a framework for the integrated analysis of high-resolution satellite imagery. This new methodology provides actionable and interpretable clusters of four cities in SSA. Not only does it bring forth meaningful information for the selected cities, but it also sets a precedent for urban decision-making strategies in other locations.

Furthermore, data collection for urban characteristics, such as buildings, often proves to be both time-intensive and costly. Procuring a comprehensive set of such urban characteristics over multiple time points presents an even greater challenge, especially for cities in SSA. This thesis demonstrates that satellite imagery can serve as a viable alternative to traditional data sources, enabling the tracking of urban characteristics over time and in combination. Unsupervised algorithms, in particular, can discern patterns in complex data without predefined criteria and labels. This capacity is key to unravelling the complexities and heterogeneities of cities' environments [20] and propelling data-driven policy decisions.

The flexibility of the framework allows it to be applied to cities globally, thereby paving the way for large-scale urban comparisons as well as specific city-focused analysis. The framework demonstrates its effectiveness by capturing unique and discernible urban phenotypes related to the natural (vegetation and water) and constructed (building count, size, density, and orientation; length and arrangement of roads) environments, as well as population dynamics. These phenotypes can represent either distinctive defining at-

tributes (e.g., water bodies or dense vegetation) or combinations thereof (e.g., buildings surrounded by vegetation or sparsely populated areas intermixed with roads). Thus, this work offers evidence and maps that can act as base for effective urban planning and public health interventions as discussed in the following section.

8.3.2 Implications for urban policy and planning

The proposed framework offers a versatile tool for systematically monitoring urban development in regions with limited data availability. Its practicality, cost-effectiveness, and efficiency in implementation make it a valuable resource for urban planning and policy-making. Specifically, the framework proves useful in addressing urban inequalities, as identified clusters align with issues such as disparities in housing access, sanitation, and the presence of urban green spaces [4].

Though, the applicability of the framework extends to various contexts, enabling the measurement of urban density and the tracking of informal settlements, especially in areas susceptible to encroachment and flooding. Additionally, it aids in monitoring urban planning decisions, including the implementation of green belts and zoning strategies. The method can also facilitate the tracking of urban green spaces and agriculture, thereby supporting the mitigation of urban air and noise pollution. Moreover, it assists in identifying areas with high building density and robust transport connectivity, which proves valuable in understanding spatial distribution concerning employment opportunities and essential services.

The scalability of this approach allows for its implementation in diverse cities, not only within specific regions but also across borders. Overcoming challenges associated with disparate urban data, the framework streamlines the comparative analysis of cities by capitalising on standardised satellite imagery formats captured uniformly across borders. This simplifies cross-city comparisons, enhancing the feasibility of assessing policy efficacy

and understanding urban changes on a regional or global scale.

Moreover, the framework supports longitudinal analysis, enabling the tracking of urban sprawl, land-use transformations, and changes in green spaces over time. These insights are crucial for urban planning, policy decisions, and health initiatives. The integration of unsupervised machine learning and high-resolution satellite imagery addresses the critical need for comprehensive urban data in SSA and beyond. The framework not only provides insights for specific cities but also establishes a tool for global urban decision-making.

8.3.3 Urban phenotypes in SSA and their implications

Examining the city structures in SSA, this thesis emphasises the irregular growth patterns typical of the region. Cities do not simply expand from a central core; they burgeon in sporadic pockets, some densely populated, even at distances from the dense city centre [4, 51]. While these high-density areas often correspond with lower economic status compared to other city regions, their strategic positioning near transportation and commerce hubs makes them centres for both formal and informal trade. Yet, this centrality often comes with the drawback of increased noise and pollution [267, 53, 268, 305]. Sprawling tendencies, however, add complexity to the urban landscape. Extensive distances between structures are a defining challenge in SSA cities [304]. Here, the unsupervised framework introduced in this thesis can aid in analysing city density shifts, spotlighting areas ready for vertical construction or new transport hubs' introduction. Furthermore, recognising the rapid expansion of cities in SSA and simultaneously informal settlements, it is vital to view informality as a potential asset, particularly concerning housing. The goal should not be to eliminate informality but to elevate the quality of the outputs it produces [306]. In this context, this work can aid in identifying pockets of informal settlements and guiding strategic interventions tailored to specific locations.

As cities continue to evolve, a trend towards taller structures in these study areas be-

comes apparent, contrasting with previous urban sprawl patterns observed in African cities [4, 270]. Such vertical developments may amplify economic activities but also carry the risk of displacing existing residents if housing needs are not addressed. The visible shifts in the urban landscape, especially the move towards verticality, can be observed using the approach detailed in this thesis.

Additionally, the research highlights a prevailing trend: commercial centres of cities in SSA are predominantly located near the city centre in the studied cities. This trend is especially pronounced in cities with their primary commercial harbour adjacent to the city centre, such as Dakar and Dar es Salaam. The closeness of these centres to the city's core significantly influences inland trade routes and the efficient distribution of goods. As cities have expanded outwards, the population has been pushed farther from these hubs, which were once strategic. However, in today's context, these locations may no longer offer optimal efficiency for trade and transportation, leading to potential traffic and transport inefficiencies [307]. Recognising the potential challenges and inefficiencies of such placements, some cities are taking strategic actions. For example, Kigali's master plan includes the establishment of a special economic zone in the peri-urban area near the airport. Such zones, foreseen to undergo substantial growth in Kigali [308, 309], can redirect major trade routes and reduce traffic congestion in city centres.

Moreover, the urban cluster maps from this thesis provide a detailed overview of city structures and potential growth areas. The combination of densely populated areas together with mixed environment clusters, capture most of the paved and unpaved roads and provide a good representation of the connectivity and accessibility of the cities. These clusters can be used to identify areas that are poorly connected. Yet, merely expanding the road network is not the sole solution. Recognising its inherent limitations [310], other viable alternatives include the introduction of rapid rail or bus systems [293], or transforming the cityscape to promote walkability and bikeability.

Finally, a key observation from this work is that without clear designation as urban green

spaces, areas frequently succumb to development pressures. Across all city boundaries analysed in the thesis, the only lush and dense green spaces within metropolitan areas were those officially marked as parks or forests. The conservation of these urban green spaces is paramount not only for the welfare of urban dwellers, providing vital spaces for recreation and mental well-being, but also for combating environmental challenges such as urban heat islands [311] or reducing the risk of flooding [296, 312, 297]. The clustering method developed through this thesis can assist in monitoring changes in urban green spaces or urban agriculture over time, offering a means to identify and safeguard areas that require protection and conservation. Preservation of urban green spaces could significantly contribute to sustainable and resilient urban development strategies in the SSA region [313]. The framework further identified areas of urban agriculture in Dar es Salaam and Kigali. These areas can help to increase urban green spaces and promote urban biodiversity within the city. Additionally, they can boost food security and improve the living standards of the residents [314, 315]. However, for this potential to be realised, there is a need to delve deeper into how urban agriculture influences biodiversity enhancement in African cities. Yet, further research is required to understand its potential [316].

Endlessly sprawling urban landscapes are unsustainable, yet a dense skyline of skyscrapers is not the ideal alternative due to potential inequalities and strain on green or open spaces. Cities in SSA face the challenge of harmonising equity, economic considerations, and sustainability. This encompasses addressing demands for building materials, addressing the carbon footprint inherent in construction, and forecasting future energy needs [304, 317]. The optimal approach might differ based on the unique climates, histories and regional variances across SSA. Yet, striving for a varied mix of land use, including horizontal and vertical development strategies will be key.

8.4 Future work

There are several directions for future work that would continue to build on these findings. First, future studies could benefit from using satellite images captured by a range of different sensors. This research relied on high-resolution satellite images of the same resolution and from the same provider. Repeating the analysis with images from multiple sensors with different resolutions or spectral bands could provide additional insight into the robustness of the method.

Second, the scope of cities included in this thesis could be expanded. The two-stage framework developed and particularised in Chapter 6 could be applied to more cities across SSA and potentially even globally. By comparing the patterns and results across diverse cities, researchers can gain a broader perspective on urbanisation and understand how African cities differ from cities in the Global North. Moreover, it would permit adjustments for variations in the images that could be attributable to seasonality or differences in the timing of image capture.

Thirdly, integrating more external data sources into the analysis could provide further insights and refine cluster interpretation. This could include further data on the built environment such as building height, population data, economic indices, including inequality measures, weather and climate data, and socio-demographic data, among others. This additional data can enrich the understanding of the urban patterns revealed in the satellite images and contribute to a more holistic understanding of urbanisation trends in SSA.

Lastly, the arena of self-supervised deep learning and remote sensing is rapidly evolving, making it a promising avenue for future research. During the course of my postgraduate research, I have witnessed exponential growth in machine learning datasets for remote sensing, stimulated by the emergence of new sensors, and enhanced capabilities for handling large geospatial data [318, 319]. There has been an impressive improvement in the performance of models and the availability of datasets that could be leveraged for transfer

learning, such as land cover classification [96]. While the models chosen in this thesis were based on existing evidence and performance metrics, the potential advantages of other unsupervised or self-supervised network architectures and emerging paradigms like vision transformers cannot be discounted. In this rapidly evolving field, future research stands to gain substantially by frequently updating and refining models to align with state-of-the-art developments. Such optimisation could enhance the performance of the clustering, thereby leading to a more precise and detailed understanding of urbanisation trends.

8.5 Conclusion

The potential for urban growth in African cities to reduce poverty, enhance health, and promote both local and global sustainability is significant [36, 320, 321, 322]. However, in order to effectively create and adjust policies that maximise this potential, it is critical to monitor the scope and features of urban development at various scales. Currently, this is a challenging task due to the insufficient and scattered nature of existing data. My research suggests that using unlabelled satellite images, in conjunction with unsupervised deep learning techniques, can help fill these data gaps across time and space, offering a scalable solution to track urbanisation in low-income countries. As satellite images become more affordable and accessible, and as computational power increases, this method will become progressively more efficient and cost-effective. It also offers the possibility of automation, allowing for nearly real-time tracking of urban growth in various cities. In doing this, my method can help narrow the data gap between the Global North and the Global South, offering a more equal deep learning approach that is not solely dependent on labelled image data, which is predominantly collected in the high and upper-middle income countries.

8.6 Open science and code availability

Most of the code written for this thesis is available on my personal GitHub page https://github.com/barbarametzler/clustering_satellite_images. The DeepCluster algorithm, which I use for my analysis and described in further detail in Chapter 4 was published by Facebook research and is also openly available at <https://github.com/facebookresearch/deepcluster>. Additionally, some of the satellite images and all of the external data sources are openly available as listed in this chapter (Chapter 3), where I describe all data sources.

Bibliography

- [1] A. B. Metzler, R. Nathvani, V. Sharmanska, W. Bai, E. Muller, S. Moulds, C. Agyei-Asabere, D. Adjei-Boadi, E. Kyere-Gyeabour, J. D. Tetteh, G. Owusu, S. Agyei-Mensah, J. Baumgartner, B. E. Robinson, R. E. Arku, and M. Ezzati, “Phenotyping urban built and natural environments with high-resolution satellite images and unsupervised deep learning,” *Science of The Total Environment*, vol. 893, p. 164794, Oct. 2023.
- [2] “Open Data Program | Disaster Response Geospatial Analytics.”
- [3] United Nations, Department of Economic and Social Affairs, Population Division, *World Urbanization Prospects: The 2018 Revision*. New York: United Nations, 2019.
- [4] S. V. Lall, J. V. Henderson, and A. J. Venables, *Africa’s cities: Opening doors to the world*. The World Bank, 2017.
- [5] G. Mills, “Cities as agents of global change,” *International Journal of Climatology*, vol. 27, no. 14, pp. 1849–1857, 2007.
- [6] M. Acuto and S. Parnell, “Leave no city behind,” *Science*, vol. 352, pp. 873–873, May 2016. Publisher: American Association for the Advancement of Science.
- [7] W. Sirko, S. Kashubin, M. Ritter, A. Annkah, Y. S. E. Bouchareb, Y. Dauphin, D. Keysers, M. Neumann, M. Cisse, and J. Quinn, “Continental-scale building detection from high resolution satellite imagery,” *arXiv preprint arXiv:2107.12283*, 2021.
- [8] M. Hallas, “Mapping Africa: How Ecopia.ai and Maxar Mapped Every Building and Road in sub-Saharan Africa Using High-Resolution Satellite Imagery,” *AGU Fall Meeting Abstracts*, vol. 11, Dec. 2019.
- [9] “Building Footprints - Bing Maps,” Feb. 2020.
- [10] S. Angel, A. M. Blei, J. Parent, P. Lamson-Hall, N. Galarza-Sanchez, D. L. Civco, and K. Thom, “Atlas of urban expansion—2016 edition,” *Areas and densities*. Cambridge, MA: NYU Urban Expansion Program at New York University, UN-Habitat, and the Lincoln Institute of Land Policy, 2016.

- [11] U. Serajuddin, H. Uematsu, C. Wieser, N. Yoshida, and A. Dabalen, “Data Deprivation,” Apr. 2015. Publisher: World Bank, Washington, DC.
- [12] J. Espey, E. Swanson, S. Badiie, Z. Christensen, A. Fischer, M. Levy, G. Yetman, A. Sherbinin, R. Chen, and Q. you, “Data for Development: A Needs Assessment for SDG Monitoring and Statistical Capacity Development. Sustainable Development Solutions Network,” 2015.
- [13] S. Giest and A. Samuels, “‘For good measure’: data gaps in a big data world,” *Policy Sciences*, vol. 53, pp. 559–569, Sept. 2020.
- [14] M. R. Ibrahim, J. Haworth, and T. Cheng, “Understanding cities with machine eyes: A review of deep computer vision in urban analytics,” *Cities*, vol. 96, p. 102481, Jan. 2020.
- [15] Z. Zhu, Y. Zhou, K. C. Seto, E. C. Stokes, C. Deng, S. T. A. Pickett, and H. Taubenböck, “Understanding an urbanizing planet: Strategic directions for remote sensing,” *Remote Sensing of Environment*, vol. 228, pp. 164–182, July 2019.
- [16] B. Huang, B. Zhao, and Y. Song, “Urban land-use mapping using a deep convolutional neural network with high spatial resolution multispectral remote sensing imagery,” *Remote Sensing of Environment*, vol. 214, pp. 73–86, Sept. 2018.
- [17] L. Rokach and O. Maimon, “Clustering Methods,” in *Data Mining and Knowledge Discovery Handbook* (O. Maimon and L. Rokach, eds.), pp. 321–352, Boston, MA: Springer US, 2005.
- [18] A. K. Jain, M. N. Murty, and P. J. Flynn, “Data clustering: a review,” *ACM Computing Surveys*, vol. 31, pp. 264–323, Sept. 1999.
- [19] D. Xu and Y. Tian, “A Comprehensive Survey of Clustering Algorithms,” *Annals of Data Science*, vol. 2, pp. 165–193, June 2015.
- [20] J. Wang and F. Biljecki, “Unsupervised machine learning in urban studies: A systematic review of applications,” *Cities*, vol. 129, p. 103925, Oct. 2022.
- [21] F. Hu, G.-S. Xia, J. Hu, and L. Zhang, “Transferring Deep Convolutional Neural Networks for the Scene Classification of High-Resolution Remote Sensing Imagery,” *Remote Sensing*, vol. 7, pp. 14680–14707, Nov. 2015. Number: 11 Publisher: Multidisciplinary Digital Publishing Institute.
- [22] Y. Gu, Y. Wang, and Y. Li, “A Survey on Deep Learning-Driven Remote Sensing Image Scene Understanding: Scene Classification, Scene Retrieval and Scene-Guided Object Detection,” *Applied Sciences*, vol. 9, May 2019.
- [23] M. Bearak, D. Moriarty, and J. Ledur, “Africa’s rising cities,” *Washington Post*, Nov. 2021.
- [24] S. Barrios, L. Bertinelli, and E. Strobl, “Climatic change and rural–urban migration: The case of sub-Saharan Africa,” *Journal of Urban Economics*, vol. 60, pp. 357–371, Nov. 2006.

- [25] M. Maurel and M. Tuccio, "Climate Instability, Urbanisation and International Migration," *The Journal of Development Studies*, vol. 52, pp. 735–752, May 2016. Publisher: Routledge _eprint: <https://doi.org/10.1080/00220388.2015.1121240>.
- [26] D. Castells-Quintana and H. Wenban-Smith, "Population Dynamics, Urbanisation without Growth, and the Rise of Megacities," *The Journal of Development Studies*, vol. 56, pp. 1663–1682, Sept. 2020. Publisher: Routledge _eprint: <https://doi.org/10.1080/00220388.2019.1702160>.
- [27] K. Hommann and S. V. Lall, *Which Way to Livable and Productive Cities?: A Road Map for Sub-Saharan Africa*. Washington, DC: World Bank Group, Mar. 2019.
- [28] P. B. Cobbinah, M. O. Erdiaw-Kwasie, and P. Amoateng, "Africa's urbanisation: Implications for sustainable development," *Cities*, vol. 47, pp. 62–72, Sept. 2015.
- [29] V. Oluwole, "Top 10 fastest growing cities in Africa 2022," June 2022. Section: lifestyle.
- [30] OECD, U. N. E. C. f. Africa, and A. D. Bank, *Africa's Urbanisation Dynamics 2022*. 2022.
- [31] J. V. Henderson and M. A. Turner, "Urbanization in the Developing World: Too Early or Too Slow?," *Journal of Economic Perspectives*, vol. 34, pp. 150–173, Aug. 2020.
- [32] B. G. Nworgu and L. N. Nworgu, "Urban-rural disparities in achievement at the basic education level: The plight of the rural child in a developing country," *Developing Country Studies*, 3 (14), 128–140. Downloaded on 30th August, 2017 from, 2013.
- [33] M. P. Blimpo and M. Cosgrove-Davies, *Electricity access in Sub-Saharan Africa: Uptake, reliability, and complementary factors for economic impact*. World Bank Publications, 2019.
- [34] D. A. Leon, "Cities, urbanization and health," *International Journal of Epidemiology*, vol. 37, pp. 4–8, Feb. 2008.
- [35] C. Dye, "Health and urban living," *Science (New York, N. Y.)*, vol. 319, pp. 766–769, Feb. 2008.
- [36] M. Ezzati, C. J. Webster, Y. G. Doyle, S. Rashid, G. Owusu, and G. M. Leung, "Cities for global health," *BMJ*, vol. 363, p. k3794, Oct. 2018. Publisher: British Medical Journal Publishing Group Section: Analysis.
- [37] D. Kadobera, B. Sartorius, H. Masanja, A. Mathew, and P. Waiswa, "The effect of distance to formal health facility on childhood mortality in rural Tanzania, 2005–2007," *Global Health Action*, vol. 5, p. 10.3402/gha.v5i0.19099, Nov. 2012.
- [38] Y. B. Okwaraji, S. Cousens, Y. Berhane, K. Mulholland, and K. Edmond, "Effect of Geographical Access to Health Facilities on Child Mortality in Rural Ethiopia: A Community Based Cross Sectional Study," *PLoS ONE*, vol. 7, p. e33564, Mar. 2012.

- [39] F. A. Armah, B. Ekumah, D. O. Yawson, J. O. Odoi, A.-R. Afitiri, and F. E. Nyieku, "Access to improved water and sanitation in sub-Saharan Africa in a quarter century," *Heliyon*, vol. 4, p. e00931, Nov. 2018.
- [40] D. Headey, D. Stifel, L. You, and Z. Guo, "Remoteness, urbanization, and child nutrition in sub-Saharan Africa," *Agricultural Economics*, vol. 49, no. 6, pp. 765–775, 2018. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/agec.12458>.
- [41] D. E. Bloom, D. Canning, and G. Fink, "Urbanization and the Wealth of Nations," *Science*, vol. 319, pp. 772–775, Feb. 2008. Publisher: American Association for the Advancement of Science.
- [42] V. Henderson, "The Urbanization Process and Economic Growth: The So-What Question," *Journal of Economic Growth*, vol. 8, no. 1, pp. 47–71, 2003. Publisher: Springer.
- [43] A. Ciccone and R. E. Hall, "Productivity and the Density of Economic Activity," Apr. 1993.
- [44] M. R. Montgomery, R. Stren, B. Cohen, and H. E. Reed, *Cities transformed: demographic change and its implications in the developing world*. Routledge, 2013.
- [45] J. V. Henderson, A. J. Venables, T. Regan, and I. Samsonov, "Building functional cities," *Science*, vol. 352, pp. 946–947, May 2016. Publisher: American Association for the Advancement of Science.
- [46] U. N. H. S. Programme, "World cities report 2020: The value of sustainable urbanization," 2020.
- [47] "Population living in slums (% of urban population) - Sub-Saharan Africa | Data."
- [48] J. Du and A. Mahendra, "Too Many Cities Are Growing Out Rather than Up. 3 Reasons That's a Problem," Jan. 2019.
- [49] S. Moulds, W. Buytaert, M. R. Templeton, and I. Kanu, "Modeling the Impacts of Urban Flood Risk Management on Social Inequality," *Water Resources Research*, vol. 57, no. 6, p. e2020WR029024, 2021. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1029/2020WR029024>.
- [50] B. I. Musah, L. Peng, and Y. Xu, "Urban Congestion and Pollution: A Quest for Cogent Solutions for Accra City," *IOP Conference Series: Earth and Environmental Science*, vol. 435, p. 012026, Jan. 2020. Publisher: IOP Publishing.
- [51] S. Lall, M. Lebrand, H. Park, D. Sturm, and A. Venables, "Pancakes to Pyramids," 2021. Publisher: World Bank, Washington, DC.
- [52] "The growth of Africa's towns and small cities is transforming the continent," *The Economist*, Sept. 2023.

- [53] S. N. Clark, A. S. Alli, R. Nathvani, A. Hughes, M. Ezzati, M. Brauer, M. B. Toledano, J. Baumgartner, J. E. Bennett, J. Nimo, J. Bedford Moses, S. Baah, S. Agyei-Mensah, G. Owusu, B. Croft, and R. E. Arku, "Space-time characterization of community noise and sound sources in Accra, Ghana," *Scientific Reports*, vol. 11, p. 11113, May 2021. Number: 1 Publisher: Nature Publishing Group.
- [54] D. Schwela, "Review of Urban Air Quality in Sub-Saharan Africa Region," 2012. Publisher: World Bank, Washington, DC.
- [55] G. Nakweya, "Improving air quality in Africa," *Nature Africa*, Jan. 2022. Bandiera_abtest: a Cg_type: News Feature Publisher: Nature Publishing Group.
- [56] S. Iddi, K. Muindi, H. Gitau, and B. Mberu, "Characterization of Healthy Housing in Africa: Method, Profiles, and Determinants," *Journal of Urban Health : Bulletin of the New York Academy of Medicine*, vol. 99, pp. 146–163, Feb. 2022.
- [57] L. S. Tusting, D. Bisanzio, G. Alabaster, E. Cameron, R. Cibulskis, M. Davies, S. Flaxman, H. S. Gibson, J. Knudsen, C. Mbogo, F. O. Okumu, L. v. Seidlein, D. J. Weiss, S. W. Lindsay, P. W. Gething, and S. Bhatt, "Mapping changes in housing in sub-Saharan Africa from 2000 to 2015," *Nature*, vol. 568, pp. 391–394, Apr. 2019.
- [58] K. Muindi, S. Iddi, H. Gitau, and B. Mberu, "Housing and health outcomes: evidence on child morbidities from six Sub-Saharan African countries," *BMC Pediatrics*, vol. 23, p. 219, May 2023.
- [59] H. Bixby, J. E. Bennett, A. A. Bawah, R. E. Arku, S. K. Annim, J. D. Anum, S. E. Mintah, A. M. Schmidt, C. Agyei-Asabere, B. E. Robinson, A. Cavanaugh, S. Agyei-Mensah, G. Owusu, M. Ezzati, and J. Baumgartner, "Quantifying within-city inequalities in child mortality across neighbourhoods in Accra, Ghana: a Bayesian spatial analysis," *BMJ Open*, vol. 12, p. e054030, Jan. 2022. Publisher: British Medical Journal Publishing Group Section: Epidemiology.
- [60] C. A. Mensah, "Urban green spaces in Africa: Nature and challenges," 2014. Accepted: 2022-01-26T09:56:17Z Publisher: University of Cape Coast.
- [61] E. Duku, C. A. Mensah, I. Amadu, and W. K. Adzigbli, "Changes in urban green spaces in coastal cities and human well-being: The case of Cape Coast Metropolis, Ghana," *Geo: Geography and Environment*, vol. 10, no. 1, p. e00119, 2023. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/geo2.119>.
- [62] B. Freund, *The African City: A History*. New Approaches to African History, Cambridge: Cambridge University Press, 2007.
- [63] T. Förster and C. Ammann, "African Cities and the Development Conundrum," *International Development Policy | Revue internationale de politique de développement*, pp. 3–25, Oct. 2018. ISBN: 9782940600076 Number: 10 Publisher: Institut de hautes études internationales et du développement.

- [64] M. Nielsen, "Inverse governmentality: The paradoxical production of peri-urban planning in Maputo, Mozambique," *Critique of Anthropology*, vol. 31, no. 4, pp. 329–358, 2011. Publisher: Sage Publications Sage UK: London, England.
- [65] A. L. Mabogunje, "Urban Planning and the Post-Colonial State in Africa: A Research Overview," *African Studies Review*, vol. 33, no. 2, pp. 121–203, 1990. Publisher: Cambridge University Press.
- [66] C. A. d. Savoires, "National urban policy : Sub saharan Africa report," Tech. Rep. 978-92-1-132739-7, Sept. 2022.
- [67] B. Solomon-Ayeh, R. King, and I. Decardi-Nelson, "Street Vending and the Use of Urban Public Space in Kumasi, Ghana," 2011.
- [68] J. Beek and A. Thiel, "Orders of trade: regulating Accra's Makola market," *The Journal of Legal Pluralism and Unofficial Law*, vol. 49, pp. 34–53, Jan. 2017. Publisher: Routledge _eprint: <https://doi.org/10.1080/07329113.2017.1289358>.
- [69] A. Brown, M. Lyons, and I. Dankoco, "Street Traders and the Emerging Spaces for Urban Voice and Citizenship in African Cities," *Urban Studies*, vol. 47, no. 3, pp. 666–683, 2010. Publisher: Temporary Publisher.
- [70] N. G. Baruah, J. V. Henderson, and C. Peng, "Colonial legacies: Shaping African cities," *Journal of Economic Geography*, vol. 21, pp. 29–65, Nov. 2017.
- [71] "Colonialism and Urbanisation," in *The African City: A History* (B. Freund, ed.), New Approaches to African History, pp. 65–106, Cambridge: Cambridge University Press, 2007.
- [72] J. R. Weeks, A. Hill, D. Stow, A. Getis, and D. Fugate, "Can we spot a neighborhood from the air? Defining neighborhood structure in Accra, Ghana," *GeoJournal*, vol. 69, no. 1-2, pp. 9–22, 2007.
- [73] "The Urban Remnants of Colonial Planning in Africa: Dar es Salaam and Nairobi," Sept. 2021.
- [74] C. Coquéry-Vidrovitch, *Racial and social zoning in African cities from colonization to postindependence*. Brill, Jan. 2012. Pages: 267-286 Section: Beyond Empire and Nation.
- [75] A. Durand-Lasserve and H. Selod, "The Formalization of Urban Land Tenure in Developing Countries," in *Urban Land Markets: Improving Land Management for Successful Urbanization* (S. V. Lall, M. Freire, B. Yuen, R. Rajack, and J.-J. Helluin, eds.), pp. 101–132, Dordrecht: Springer Netherlands, 2009.
- [76] D. Koenig, "Conflicting Interests of Residents and Planners in West African Cities," *Proceedings of the African Futures Conference*, vol. 1, no. 1, pp. 158–159, 2016. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.2573-508X.2016.tb00043.x>.

- [77] M. Norris, G. Klabbers, A. B. Pembe, C. Hanson, U. Baker, K. Aung, M. Mmweteni, R. S. Mfaume, and L. Beňová, “A growing disadvantage of being born in an urban area? Analysing urban-rural disparities in neonatal mortality in 21 African countries with a focus on Tanzania,” *BMJ global health*, vol. 7, p. e007544, Jan. 2022.
- [78] D. Inkoom, O. Amponsah, and E. Adabor, “Half a Decade of Implementation of Ghana’s Urban Policy: Stocktaking the Successes and Failures of Flagship Projects,” *GUTT Background Paper*, 2019.
- [79] H. Burns, “A Smart City Masterplan, Kigali,” *Urban Planning for Transitions*, pp. 153–169, 2021. Publisher: Wiley Online Library.
- [80] United Republic of Tanzania Ministry of Lands, Housing and Human Settlements Development, *Dar es Salaam Masterplan 2012-2032*. Dar es Salaam: United Republic of Tanzania, 2013.
- [81] L. L. Peter and Y. Yang, “Urban planning historical review of master plans and the way towards a sustainable city: Dar es Salaam, Tanzania,” *Frontiers of Architectural Research*, vol. 8, pp. 359–377, Sept. 2019.
- [82] “Goal 11 | Department of Economic and Social Affairs.”
- [83] M. Acuto, S. Parnell, and K. C. Seto, “Building a global urban science,” *Nature Sustainability*, vol. 1, pp. 2–4, Jan. 2018. Number: 1 Publisher: Nature Publishing Group.
- [84] B. Hillier, *Space is the machine: a configurational theory of architecture*. London, UK: Space Syntax, 2007. Publication Title: Space Syntax: London, UK. (2007).
- [85] S. Brock, “Goethe’s contribution to philosophy: the morphology of individuality,” *SATS*, vol. 20, pp. 35–52, July 2019. Publisher: De Gruyter.
- [86] J. Dibble, A. Prelorndjos, O. Romice, M. Zanella, E. Strano, M. Pagel, and S. Porta, “On the origin of spaces : morphometric foundations of urban form evolution,” *Environment and Planning B: Planning and Design*, vol. 46, pp. 707–730, May 2019. Num Pages: 24 Number: 4.
- [87] H. Taubenböck, H. Debray, C. Qiu, M. Schmitt, Y. Wang, and X. X. Zhu, “Seven city types representing morphologic configurations of cities across the globe,” *Cities*, vol. 105, p. 102814, Oct. 2020.
- [88] H. Taubenböck, C. Gerten, K. Rusche, S. Siedentop, and M. Wurm, “Patterns of Eastern European urbanisation in the mirror of Western trends – Convergent, unique or hybrid?,” *Environment and Planning B: Urban Analytics and City Science*, vol. 46, pp. 1206–1225, Sept. 2019. Publisher: SAGE Publications Ltd STM.
- [89] J. Jacobs, *Dark Age Ahead: Author of The Death and Life of Great American Cities*. Vintage Canada, 2010.

- [90] N. AlSayyad and A. Roy, *Urban Informality: Transnational Perspectives from the Middle East, Latin America, and South Asia*. Lexington Books, Nov. 2003. Google-Books-ID: mf0eAAAAQBAJ.
- [91] G. Boeing, “Measuring the complexity of urban form and design,” *URBAN DESIGN International*, vol. 23, pp. 281–292, Nov. 2018.
- [92] B. Hillier, “Space is the machine: A configurational theory of architecture,” 1996.
- [93] F. E. Horton and D. R. Reynolds, “Effects of Urban Spatial Structure on Individual Behavior,” *Economic Geography*, vol. 47, no. 1, pp. 36–48, 1971.
- [94] X. Liu, A. de Sherbinin, and Y. Zhan, “Mapping urban extent at large spatial scales using machine learning methods with VIIRS nighttime light and MODIS daytime NDVI data,” *Remote Sensing*, vol. 11, p. 1247, Jan. 2019. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.
- [95] K. Karra, C. Kontgis, Z. Statman-Weil, J. C. Mazzariello, M. Mathis, and S. P. Brumby, “Global land use/land cover with Sentinel 2 and deep learning,” in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, pp. 4704–4707, IEEE, 2021.
- [96] H. Alemohammad and K. Booth, “LandCoverNet: A global benchmark land cover classification training dataset,” Dec. 2020. arXiv:2012.03111 [cs].
- [97] A. Boguszewski, D. Batorski, N. Ziemba-Jankowska, T. Dziedzic, and A. Zambrzycka, “LandCover.ai: Dataset for Automatic Mapping of Buildings, Woodlands, Water and Roads from Aerial Imagery,” Apr. 2022. arXiv:2005.02264 [cs].
- [98] C. Huang, J. Yang, N. Clinton, L. Yu, H. Huang, I. Dronova, and J. Jin, “Mapping the maximum extents of urban green spaces in 1039 cities using dense satellite images,” *Environmental Research Letters*, vol. 16, p. 064072, June 2021. Publisher: IOP Publishing.
- [99] R. Mathieu, C. Freeman, and J. Aryal, “Mapping private gardens in urban areas using object-oriented techniques and very high-resolution satellite imagery,” *Landscape and Urban Planning*, vol. 81, pp. 179–192, June 2007.
- [100] P. Das and S. Chand, “Extracting road maps from high-resolution satellite imagery using refined DSE-LinkNet,” *Connection Science*, vol. 33, pp. 278–295, Apr. 2021. Publisher: Taylor & Francis eprint: <https://doi.org/10.1080/09540091.2020.1807466>.
- [101] G. Cadamuro, A. Muhebwa, and J. Taneja, “Assigning a Grade: Accurate Measurement of Road Quality Using Satellite Imagery,” Dec. 2018. arXiv:1812.01699 [cs, stat].
- [102] E. Brewer, J. Lin, P. Kemper, J. Hennin, and D. Runfola, “Predicting road quality using high resolution satellite imagery: A transfer learning approach,” *PLoS ONE*, vol. 16, p. e0253370, July 2021.

- [103] T. G. Tiecke, X. Liu, A. Zhang, A. Gros, N. Li, G. Yetman, T. Kilic, S. Murray, B. Blankespoor, E. B. Prydz, and H.-A. H. Dang, *Mapping the World Population One Building at a Time*. World Bank, 2017. eprint: <https://elibrary.worldbank.org/doi/pdf/10.1596/33700>.
- [104] N. Naik, S. D. Kominers, R. Raskar, E. L. Glaeser, and C. A. Hidalgo, “Computer vision uncovers predictors of physical urban change,” *Proceedings of the National Academy of Sciences*, vol. 114, pp. 7571–7576, July 2017.
- [105] S. M. Arietta, A. A. Efros, R. Ramamoorthi, and M. Agrawala, “City Forensics: Using Visual Elements to Predict Non-Visual City Attributes,” *IEEE transactions on visualization and computer graphics*, vol. 20, pp. 2624–2633, Dec. 2014.
- [106] T. Gebru, J. Krause, Y. Wang, D. Chen, J. Deng, E. L. Aiden, and L. Fei-Fei, “Using deep learning and Google Street View to estimate the demographic makeup of neighborhoods across the United States,” *Proceedings of the National Academy of Sciences*, vol. 114, pp. 13108–13113, Dec. 2017. Publisher: Proceedings of the National Academy of Sciences.
- [107] R. J. Sampson and S. W. Raudenbush, “Systematic Social Observation of Public Spaces: A New Look at Disorder in Urban Neighborhoods,” *American Journal of Sociology*, vol. 105, pp. 603–651, Nov. 1999.
- [108] M. Kuffer, K. Pfeffer, and R. Sliuzas, “Slums from Space—15 Years of Slum Mapping Using Remote Sensing,” *Remote Sensing*, vol. 8, p. 455, June 2016.
- [109] X. X. Zhu, D. Tuia, L. Mou, G.-S. Xia, L. Zhang, F. Xu, and F. Fraundorfer, “Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, pp. 8–36, Dec. 2017. Conference Name: IEEE Geoscience and Remote Sensing Magazine.
- [110] L. Zhang, L. Zhang, and B. Du, “Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art,” *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, pp. 22–40, June 2016.
- [111] E. Suel, J. W. Polak, J. E. Bennett, and M. Ezzati, “Measuring social, environmental and health inequalities using deep learning and street imagery,” *Scientific Reports*, vol. 9, pp. 1–10, Apr. 2019.
- [112] M. Helbich, Y. Yao, Y. Liu, J. Zhang, P. Liu, and R. Wang, “Using deep learning to examine street view green and blue spaces and their associations with geriatric depression in Beijing, China,” *Environment International*, vol. 126, pp. 107–117, May 2019.
- [113] X. Li, C. Zhang, W. Li, R. Ricard, Q. Meng, and W. Zhang, “Assessing street-level urban greenery using Google Street View and a modified green view index,” *Urban Forestry & Urban Greening*, vol. 14, pp. 675–685, Jan. 2015.
- [114] D. R. Richards and P. J. Edwards, “Quantifying street tree regulating ecosystem services using Google Street View,” *Ecological Indicators*, vol. 77, pp. 31–40, June 2017.

- [115] I. Seiferling, N. Naik, C. Ratti, and R. Proulx, “Green streets Quantifying and mapping urban trees with street-level imagery and computer vision,” *Landscape and Urban Planning*, vol. 165, pp. 93–101, Sept. 2017.
- [116] Y. Yang, D. He, Z. Gou, R. Wang, Y. Liu, and Y. Lu, “Association between street greenery and walking behavior in older adults in Hong Kong,” *Sustainable Cities and Society*, vol. 51, p. 101747, Nov. 2019.
- [117] P. J. Villeneuve, R. L. Ysseldyk, A. Root, S. Ambrose, J. DiMuzio, N. Kumar, M. Shehata, M. Xi, E. Seed, X. Li, M. Shooshtari, and D. Rainham, “Comparing the Normalized Difference Vegetation Index with the Google Street View Measure of Vegetation to Assess Associations between Greenness, Walkability, Recreational Physical Activity, and Health in Ottawa, Canada,” *International Journal of Environmental Research and Public Health*, vol. 15, p. 1719, Aug. 2018. Number: 8 Publisher: Multidisciplinary Digital Publishing Institute.
- [118] J. M. Keralis, M. Javanmardi, S. Khanna, P. Dwivedi, D. Huang, T. Tasdizen, and Q. C. Nguyen, “Health and the built environment in United States cities: measuring associations using Google Street View-derived indicators of the built environment,” *BMC Public Health*, vol. 20, p. 215, Feb. 2020.
- [119] E. Suel, M. Sorek-Hamer, I. Moise, M. von Pohle, A. Sahasrabhojane, A. A. Asanjan, R. E. Arku, A. S. Alli, B. Barratt, S. N. Clark, A. Middel, E. Deardorff, V. Lingenfelter, N. C. Oza, N. Yadav, M. Ezzati, and M. Brauer, “What You See Is What You Breathe? Estimating Air Pollution Spatial Variation Using Street-Level Imagery,” *Remote Sensing*, vol. 14, p. 3429, Jan. 2022. Number: 14 Publisher: Multidisciplinary Digital Publishing Institute.
- [120] F.-Y. Gong, Z.-C. Zeng, E. Ng, and L. K. Norford, “Spatiotemporal patterns of street-level solar radiation estimated using Google Street View in a high-density urban environment,” *Building and Environment*, vol. 148, pp. 547–566, Jan. 2019.
- [121] A. Ganji, L. Minet, S. Weichenthal, and M. Hatzopoulou, “Predicting Traffic-Related Air Pollution Using Feature Extraction from Built Environment Images,” *Environmental Science & Technology*, vol. 54, pp. 10688–10699, Sept. 2020. Publisher: American Chemical Society.
- [122] N. Naik, R. Raskar, and C. Hidalgo, “Cities Are Physical Too: Using Computer Vision to Measure the Quality and Impact of Urban Appearance †,” *American Economic Review*, vol. 106, pp. 128–132, May 2016.
- [123] E. Muller, E. Gemmell, I. Choudhury, R. Nathvani, A. B. Metzler, J. Bennett, E. Denton, S. Flaxman, and M. Ezzati, “Mapping City-Wide Perceptions of Neighbourhood Quality using Street View Images,” 2022.
- [124] Q. C. Nguyen, M. Sajjadi, M. McCullough, M. Pham, T. T. Nguyen, W. Yu, H.-W. Meng, M. Wen, F. Li, K. R. Smith, K. Brunisholz, and T. Tasdizen, “Neighbourhood looking glass: 360^o automated characterisation of the built environment for neighbourhood effects research,” *J Epidemiol Community Health*, vol. 72, pp. 260–266, Mar. 2018. Publisher: BMJ Publishing Group Ltd Section: Research report.

- [125] R. Nathvani, V. D., S. N. Clark, A. S. Alli, E. Muller, H. Coste, J. E. Bennett, J. Nimo, J. B. Moses, S. Baah, A. Hughes, E. Suel, A. B. Metzler, T. Rashid, M. Brauer, J. Baumgartner, G. Owusu, S. Agyei-Mensah, R. E. Arku, and M. Ezzati, “Beyond here and now: Evaluating pollution estimation across space and time from street view images with deep learning,” *Science of The Total Environment*, vol. 903, p. 166168, Dec. 2023.
- [126] R. Nathvani, S. N. Clark, E. Muller, A. S. Alli, J. E. Bennett, J. Nimo, J. B. Moses, S. Baah, A. B. Metzler, M. Brauer, E. Suel, A. F. Hughes, T. Rashid, E. Gemmell, S. Moulds, J. Baumgartner, M. Toledano, E. Agyemang, G. Owusu, S. Agyei-Mensah, R. E. Arku, and M. Ezzati, “Characterisation of urban environment and activity across space and time using street images and deep learning in Accra,” *Scientific Reports*, vol. 12, p. 20470, Nov. 2022. Number: 1 Publisher: Nature Publishing Group.
- [127] S. Weichenthal, M. Hatzopoulou, and M. Brauer, “A picture tells a thousand...exposures: Opportunities and challenges of deep learning image analyses in exposure science and environmental epidemiology,” *Environment International*, vol. 122, pp. 3–10, Jan. 2019.
- [128] R. Wang, Y. Yuan, Y. Liu, J. Zhang, P. Liu, Y. Lu, and Y. Yao, “Using street view data and machine learning to assess how perception of neighborhood safety influences urban residents’ mental health,” *Health & Place*, vol. 59, p. 102186, Sept. 2019.
- [129] D. Hristova, L. M. Aiello, and D. Quercia, “The New Urban Success: How Culture Pays,” *Frontiers in Physics*, vol. 6, 2018.
- [130] C. Ye, F. Zhang, L. Mu, Y. Gao, and Y. Liu, “Urban function recognition by integrating social media and street-level imagery,” *Environment and Planning B: Urban Analytics and City Science*, vol. 48, pp. 1430–1444, July 2021. Publisher: SAGE Publications Ltd STM.
- [131] F. Biljecki and K. Ito, “Street view imagery in urban analytics and GIS: A review,” *Landscape and Urban Planning*, vol. 215, p. 104217, Nov. 2021.
- [132] N. J. W. THROWER, “Annals Map Supplement Number Twelve: Land Use in the Southwestern United States—from Gemini and Apollo-Imagery,” *Annals of the Association of American Geographers*, vol. 60, pp. 208–209, Mar. 1970. Publisher: Routledge eprint: <https://doi.org/10.1111/j.1467-8306.1970.tb00714.x>.
- [133] F. Eugenio and J. Marcello, *Very High Resolution (VHR) Satellite Imagery: Processing and Applications*. MDPI, Nov. 2019.
- [134] T. Esch, H. Taubenböck, W. Heldens, M. Thiel, M. Wurm, and S. Dech, “How can earth observation support the sustainable development of urban environments?,” *Urban Remote Sensing*, 2010.

- [135] E. A. Wentz, S. Anderson, M. Fragkias, M. Netzbald, V. Mesev, S. W. Myint, D. Quattrochi, A. Rahman, and K. C. Seto, "Supporting Global Environmental Change Research: A Review of Trends and Knowledge Gaps in Urban Remote Sensing," *Remote Sensing*, vol. 6, pp. 3879–3905, May 2014. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.
- [136] Y. Forget, M. Shimoni, M. Gilbert, and C. Linard, "Mapping 20 Years of Urban Expansion in 45 Urban Areas of Sub-Saharan Africa," *Remote Sensing*, vol. 13, p. 525, Jan. 2021. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
- [137] Y. Jia, Y. Ge, F. Ling, X. Guo, J. Wang, L. Wang, Y. Chen, and X. Li, "Urban Land Use Mapping by Combining Remote Sensing Imagery and Mobile Phone Positioning Data," *Remote Sensing*, vol. 10, p. 446, Mar. 2018. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
- [138] J. Aryal, C. Sitaula, and S. Aryal, "NDVI Threshold-Based Urban Green Space Mapping from Sentinel-2A at the Local Governmental Area (LGA) Level of Victoria, Australia," *Land*, vol. 11, p. 351, Mar. 2022. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
- [139] E. Suel, S. Bhatt, M. Brauer, S. Flaxman, and M. Ezzati, "Multimodal deep learning from satellite and street-level imagery for measuring income, overcrowding, and environmental deprivation in urban areas," *Remote Sensing of Environment*, vol. 257, p. 112339, May 2021.
- [140] N. Jean, M. Burke, M. Xie, W. M. Davis, D. B. Lobell, and S. Ermon, "Combining satellite imagery and machine learning to predict poverty," *Science*, vol. 353, pp. 790–794, Aug. 2016.
- [141] A. Head, M. Manguin, N. Tran, and J. E. Blumenstock, "Can Human Development be Measured with Satellite Imagery?," in *Proceedings of the Ninth International Conference on Information and Communication Technologies and Development*, ICTD '17, (New York, NY, USA), pp. 1–11, Association for Computing Machinery, Nov. 2017.
- [142] L. Alonso and F. Renard, "Integrating Satellite-Derived Data as Spatial Predictors in Multiple Regression Models to Enhance the Knowledge of Air Temperature Patterns," *Urban Science*, vol. 3, p. 101, Dec. 2019. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.
- [143] D. M. Varade, A. K. Maurya, and O. Dikshit, "Development of Spectral Indexes in Hyperspectral Imagery for Land Cover Assessment," *IETE Technical Review*, vol. 36, pp. 475–483, Sept. 2018. Publisher: Taylor & Francis.
- [144] M. Li, S. Zang, B. Zhang, S. Li, and C. Wu, "A Review of Remote Sensing Image Classification Techniques: the Role of Spatio-contextual Information," *European Journal of Remote Sensing*, vol. 47, pp. 389–411, Jan. 2014. Publisher: Taylor & Francis .eprint: <https://doi.org/10.5721/EuJRS20144723>.

- [145] L.-K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *IEEE Transactions on geoscience and remote sensing*, vol. 37, no. 2, pp. 780–795, 1999. Publisher: IEEE.
- [146] G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia, "Remote Sensing Image Scene Classification Meets Deep Learning: Challenges, Methods, Benchmarks, and Opportunities," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 3735–3756, 2020. arXiv:2005.01094 [cs].
- [147] Q. Zou, L. Ni, T. Zhang, and Q. Wang, "Deep Learning Based Feature Selection for Remote Sensing Scene Classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, pp. 1–5, Nov. 2015.
- [148] T. Blaschke and J. Strobl, "What's wrong with pixels? Some recent developments interfacing remote sensing and GIS," 2001.
- [149] E. R. Davies, *Computer Vision - 5th Edition*. Academic Press, 2017.
- [150] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016.
- [151] X. Zhao, B. Yu, Y. Liu, Z. Chen, Q. Li, C. Wang, and J. Wu, "Estimation of Poverty Using Random Forest Regression with Multi-Source Data: A Case Study in Bangladesh," *Remote Sensing*, vol. 11, p. 375, Jan. 2019.
- [152] J. E. Steele, P. R. Sundsøy, C. Pezzulo, V. A. Alegana, T. J. Bird, J. Blumenstock, J. Bjelland, K. Engø-Monsen, Y.-A. de Montjoye, A. M. Iqbal, K. N. Hadiuzzaman, X. Lu, E. Wetter, A. J. Tatem, and L. Bengtsson, "Mapping poverty using mobile phone and satellite data," *Journal of The Royal Society Interface*, vol. 14, p. 20160690, Feb. 2017.
- [153] J. C. Duque, J. E. Patino, L. A. Ruiz, and J. E. Pardo-Pascual, "Measuring intra-urban poverty using land cover and texture metrics derived from remote sensing data," *Landscape and Urban Planning*, vol. 135, pp. 11–21, Mar. 2015.
- [154] D. Newhouse, "Mapping Poverty by Satellite," p. 4.
- [155] P. Zhang, Y. Ke, Z. Zhang, M. Wang, P. Li, and S. Zhang, "Urban Land Use and Land Cover Classification Using Novel Deep Learning Models Based on High Spatial Resolution Satellite Imagery," *Sensors (Basel, Switzerland)*, vol. 18, Nov. 2018.
- [156] B. Babenko, J. Hersh, D. Newhouse, A. Ramakrishnan, and T. Swartz, "Poverty Mapping Using Convolutional Neural Networks Trained on High and Medium Resolution Satellite Images, With an Application in Mexico," *arXiv:1711.06323 [cs, stat]*, Nov. 2017. arXiv: 1711.06323.
- [157] C. Yeh, A. Perez, A. Driscoll, G. Azzari, Z. Tang, D. Lobell, S. Ermon, and M. Burke, "Using publicly available satellite imagery and deep learning to understand economic well-being in Africa," *Nature Communications*, vol. 11, pp. 1–11, May 2020. Number: 1 Publisher: Nature Publishing Group.

- [158] L. D. Knibbs, M. G. Hewson, M. J. Bechle, J. D. Marshall, and A. G. Barnett, "A national satellite-based land-use regression model for air pollution exposure assessment in Australia," *Environmental Research*, vol. 135, pp. 204–211, Nov. 2014.
- [159] S. B. Henderson, M. Brauer, Y. C. MacNab, and S. M. Kennedy, "Three Measures of Forest Fire Smoke Exposure and Their Associations with Respiratory and Cardiovascular Health Outcomes in a Population-Based Cohort," *Environmental Health Perspectives*, vol. 119, pp. 1266–1271, Sept. 2011. Publisher: Environmental Health Perspectives.
- [160] M. Sorek-Hamer, M. Von Pohle, A. Sahasrabhojane, A. Akbari Asanjan, E. Deardorff, E. Suel, V. Lingenfelter, K. Das, N. C. Oza, M. Ezzati, and M. Brauer, "A Deep Learning Approach for Meter-Scale Air Quality Estimation in Urban Environments Using Very High-Spatial-Resolution Satellite Imagery," *Atmosphere*, vol. 13, p. 696, May 2022. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.
- [161] V. Lingenfelter, M. Sorek-Hamer, E. Deardorff, M. von Pohle, A. Sahasrabhojane, A. Akbari Asanjan, M. Brauer, and H. Davies, "Satellite Imagery-Based Urban Noise Prediction Model - Case study over Vancouver, Canada," vol. 2020, pp. A043–0014, Dec. 2020. Conference Name: AGU Fall Meeting Abstracts ADS Bibcode: 2020AGUFMA043.0014L.
- [162] R. S. DEFRIES and J. R. G. TOWNSHEND, "NDVI-derived land cover classifications at a global scale," *International Journal of Remote Sensing*, vol. 15, pp. 3567–3586, Nov. 1994. Publisher: Taylor & Francis eprint: <https://doi.org/10.1080/01431169408954345>.
- [163] S. Lang, E. Schöpfer, D. Hölbling, T. Blaschke, M. Moeller, T. Jekel, and E. Kloyber, "Quantifying and Qualifying Urban Green by Integrating Remote Sensing, GIS, and Social Science Method," in *Use of Landscape Sciences for the Assessment of Environmental Security* (I. Petrosillo, F. Müller, K. B. Jones, G. Zurlini, K. Krauze, S. Victorov, B.-L. Li, and W. G. Kepner, eds.), NATO Science for Peace and Security Series C: Environmental Security, (Dordrecht), pp. 93–105, Springer Netherlands, 2008.
- [164] O.-H. Kwon, I. Hong, J. Yang, D. Y. Wohn, W.-S. Jung, and M. Cha, "Urban green space and happiness in developed countries," *EPJ Data Science*, vol. 10, p. 28, Dec. 2021. Number: 1 Publisher: Springer Berlin Heidelberg.
- [165] R. Xu and D. Wunsch, "Survey of clustering algorithms," *IEEE Transactions on Neural Networks*, vol. 16, pp. 645–678, May 2005.
- [166] A. S. Shirkhorshidi, S. Aghabozorgi, and T. Wah, "A Comparison Study on Similarity and Dissimilarity Measures in Clustering Continuous Data," *PLOS ONE*, vol. 10, p. e0144059, Dec. 2015.
- [167] D. Comaniciu and P. Meer, "Mean shift analysis and applications," in *Proceedings of the seventh IEEE international conference on computer vision*, vol. 2, pp. 1197–1203, IEEE, 1999.

- [168] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 152, pp. 166–177, June 2019.
- [169] G. Grekousis, "Artificial neural networks and deep learning in urban geography: A systematic review and meta-analysis," *Computers, Environment and Urban Systems*, vol. 74, pp. 244–256, Mar. 2019.
- [170] Z. Ullah, F. Al-Turjman, L. Mostarda, and R. Gagliardi, "Applications of Artificial Intelligence and Machine learning in smart cities," *Computer Communications*, vol. 154, pp. 313–323, Mar. 2020.
- [171] Y. Zheng, L. Capra, O. Wolfson, and H. Yang, "Urban Computing: Concepts, Methodologies, and Applications," *ACM Transactions on Intelligent Systems and Technology*, vol. 5, pp. 38:1–38:55, Sept. 2014.
- [172] K. El Boucheffry and R. S. de Souza, *Learning in Big Data: Introduction to Machine Learning*. Apr. 2020. Pages: 225-249 Publication Title: Knowledge Discovery in Big Data from Astronomy and Earth Observation ADS Bibcode: 2020kdbd.book..225E.
- [173] D. Arribas-Bel and M. Fleischmann, "Spatial Signatures - Understanding (urban) spaces through form and function," *Habitat International*, vol. 128, p. 102641, Oct. 2022.
- [174] X. X. Zhu, C. Qiu, J. Hu, Y. Shi, Y. Wang, M. Schmitt, and H. Taubenböck, "The urban morphology on our planet – Global perspectives from space," *Remote Sensing of Environment*, vol. 269, p. 112794, Feb. 2022.
- [175] S. Pauleit and F. Duhme, "Assessing the environmental performance of land cover types for urban planning," *Landscape and Urban Planning*, vol. 52, pp. 1–20, Nov. 2000.
- [176] S. Borra, R. Thanki, and N. Dey, "Satellite Image Clustering," in *Satellite Image Analysis: Clustering and Classification* (S. Borra, R. Thanki, and N. Dey, eds.), SpringerBriefs in Applied Sciences and Technology, pp. 31–52, Singapore: Springer, 2019.
- [177] B. Usman, "Satellite Imagery Land Cover Classification using K-Means Clustering Algorithm Computer Vision for Environmental Information Extraction," p. 5, 2013.
- [178] A. K. Goswami, S. Sharma, and P. V. Kumar, "Nearest Clustering Algorithm for Satellite Image Classification in Remote Sensing Applications," 2014.
- [179] A. H. T. Al-Ghraihi, Z. H. Abed, F. H. Fadhil, and F. K. Naser, "Classification of Satellite Images Based on Color Features Using Remote Sensing," *International Journal of Computer (IJC)*, vol. 31, pp. 42–52, Oct. 2018.
- [180] M. Gašparović, M. Zrinjski, and M. Gudelj, "Automatic cost-effective method for land cover classification (ALCC)," *Computers, Environment and Urban Systems*, vol. 76, pp. 1–10, July 2019.

- [181] M. K. Dhodhi, J. A. Saghri, I. Ahmad, and R. Ul-Mustafa, "D-ISODATA: A Distributed Algorithm for Unsupervised Classification of Remotely Sensed Data on Network of Workstations," *Journal of Parallel and Distributed Computing*, vol. 59, pp. 280–301, Nov. 1999.
- [182] B. Sowmya and B. Sheelarani, "Land cover classification using reformed fuzzy C-means," *Sadhana*, vol. 36, pp. 153–165, Apr. 2011.
- [183] S. Berisha, F. F. Shahraki, D. Mayerich, and S. Prasad, "Deep Learning for Hyperspectral Image Analysis, Part I: Theory and Algorithms," in *Hyperspectral Image Analysis: Advances in Machine Learning and Signal Processing* (S. Prasad and J. Chanussot, eds.), Advances in Computer Vision and Pattern Recognition, pp. 37–68, Cham: Springer International Publishing, 2020.
- [184] J. Block, M. Yazdani, M. Nguyen, D. Crawl, M. Jankowska, J. Graham, T. Defanti, and I. Altintas, "An Unsupervised Deep Learning Approach for Satellite Image Analysis with Applications in Demographic Analysis," in *2017 IEEE 13th International Conference on e-Science (e-Science)*, pp. 9–18, Oct. 2017.
- [185] S. Benz, H. Park, J. Li, D. Crawl, J. Block, M. Nguyen, and I. Altintas, "Understanding a Rapidly Expanding Refugee Camp Using Convolutional Neural Networks and Satellite Imagery," in *2019 15th International Conference on eScience (eScience)*, pp. 243–251, Sept. 2019.
- [186] M. Yazdani, M. H. Nguyen, J. Block, D. Crawl, N. Zurutuza, D. Kim, G. Hanson, and I. Altintas, "Scalable Detection of Rural Schools in Africa Using Convolutional Neural Networks and Satellite Imagery," in *2018 IEEE/ACM International Conference on Utility and Cloud Computing Companion (UCC Companion)*, pp. 142–147, Dec. 2018.
- [187] K. L. de Jong and A. S. Bosman, "Unsupervised Change Detection in Satellite Images Using Convolutional Neural Networks," *arXiv:1812.05815 [cs]*, Mar. 2019. arXiv: 1812.05815.
- [188] E. Kalinicheva, D. Ienco, J. Sublime, and M. Trocan, "Unsupervised Change Detection Analysis in Satellite Image Time Series Using Deep Learning Combined With Graph-Based Approaches," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 1450–1466, 2020.
- [189] S. Saha, F. Bovolo, and L. Bruzzone, "Change Detection in Image Time-Series Using Unsupervised LSTM," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022. Conference Name: IEEE Geoscience and Remote Sensing Letters.
- [190] J. Huang, Q. Dong, S. Gong, and X. Zhu, "Unsupervised Deep Learning by Neighbourhood Discovery," *arXiv:1904.11567 [cs]*, May 2019. arXiv: 1904.11567.
- [191] A. Chatterjee, J. Saha, J. Mukherjee, S. Aikat, and A. Misra, "Unsupervised Land Cover Classification of Hybrid and Dual-Polarized Images Using Deep Convolutional Neural Network," *IEEE Geoscience and Remote Sensing Letters*, vol. 18, pp. 969–973, June 2021. Conference Name: IEEE Geoscience and Remote Sensing Letters.

- [192] M. Stevenson, C. Mues, and C. Bravo, “Deep residential representations: Using unsupervised learning to unlock elevation data for geo-demographic prediction,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 187, pp. 378–392, May 2022.
- [193] Ghana Statistical Service, “Central Data Catalog,” 2019. <http://www2.statsghana.gov.gh/nada/index.php/catalog>.
- [194] S. Agyei-Mensah and A. de Graft Aikins, “Epidemiological transition and the double burden of disease in Accra, Ghana,” *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, vol. 87, pp. 879–897, Sept. 2010.
- [195] S. Agyei-Mensah and G. Owusu, “Segregated by neighbourhoods? A portrait of ethnic diversity in the neighbourhoods of the Accra Metropolitan Area, Ghana,” *Population, Space and Place*, vol. 16, no. 6, pp. 499–516, 2010. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/psp.551>.
- [196] R. Engstrom, C. Ofiesh, D. Rain, H. Jewell, and J. Weeks, “Defining neighborhood boundaries for urban health research in developing countries: a case study of Accra, Ghana,” *Journal of Maps*, vol. 9, pp. 36–42, Mar. 2013.
- [197] J. R. Weeks, A. G. Hill, and J. Stoler, *Spatial inequalities*. Springer, 2013.
- [198] F. Byamugisha, *Securing Africa’s land for shared prosperity* : World Bank,, 2013.
- [199] C. Toulmin, “Securing Land and Property Rights in Sub-Saharan Africa: The Role of Local Institutions,” *Land Use Policy*, vol. 26, pp. 10–19, Jan. 2009.
- [200] A. Mahendra and K. C. Seto, “Upward and outward growth: managing urban expansion for more equitable cities in the global south,” *World Resources Report: Towards a More Equal City*, 2019.
- [201] B. Addae and N. Oppelt, “Land-Use/Land-Cover Change Analysis and Urban Growth Modelling in the Greater Accra Metropolitan Area (GAMA), Ghana,” *Urban Science*, vol. 3, p. 26, Mar. 2019. Number: 1 Publisher: Multidisciplinary Digital Publishing Institute.
- [202] O. Cissé, “Challenges of Urbanisation in Dakar,” Tech. Rep. 978-2-493256-01-0, Friedrich-Ebert-Stiftung, Dakar-Fann, 2022.
- [203] “Senegal Urbanization Review How can cities contribute to the “Senegal Emerging Plan”?,” tech. rep., World Bank, 2015.
- [204] “Dakar city profile,” 2021.
- [205] support, “Dakar, A City In Constant Evolution,” Apr. 2023.
- [206] M. Le Pape, “Sinou, Alain. - Comptoirs et villes coloniales du Sénégal. Saint-Louis, Gorée, Dakar,” *Cahiers d’Études africaines*, vol. 33, no. 132, pp. 682–683, 1993. Publisher: Persée - Portail des revues scientifiques en SHS.

- [207] A. Njoh, *Planning Power: Town Planning and Social Control in Colonial Africa*. London: UCL Press, Dec. 2006.
- [208] M. Nyssölä, T. Kelsall, and T. Ndezi, “Dar es Salaam: City scoping study,” p. 8, 2021.
- [209] A. Harris and G. Beane, “Coastal Cities of the Western Indian Ocean Region and the Blue Economy: City Case Study - Dar es Salaam.,” Tech. Rep. ISSN 2799-2217, WIOMSA and UN-Habitat, 2021. ISSN 2799-2217.
- [210] “Sewerage | Dar-es-salaam Water Supply and Sanitation Authority.”
- [211] V. Manirakiza, L. Mugabe, A. Nsabimana, and M. Nzayirambaho, “City Profile: Kigali, Rwanda,” *Environment and Urbanization ASIA*, vol. 10, pp. 290–307, Sept. 2019. Publisher: SAGE Publications India.
- [212] G. Baffoe, S. Ahmad, and R. Bhandari, “The road to sustainable Kigali: A contextualized analysis of the challenges,” *Cities*, vol. 105, p. 102838, Oct. 2020.
- [213] S. Krishnasamy, L. Chhetri, D. Thosani, and A. Mchugh, “Kigali Master Plan 2050 - Master Plan Report,” tech. rep., Kigali, Rwanda, 2020.
- [214] I. Turok, “Cities as platforms for progress: Local drivers of Rwanda’s success,” *Local Economy*, vol. 34, no. 3, pp. 221–227, 2019. Publisher: SAGE Publications Sage UK: London, England.
- [215] “Ghana Statistical Services..”
- [216] “Agence Nationale de la Statistique et de la Démographie (ANSD) du Sénégal | Un centre d’excellence dans un Système statistique national fort.”
- [217] “National Bureau of Statistics - Home.”
- [218] “Home.”
- [219] C. Beam, “Soon, satellites will be able to watch you everywhere all the time,” June 2019. Library Catalog: www.technologyreview.com.
- [220] N. Baghdadi and Z. Mehrez, *Optical Remote Sensing of Land Surface*. Elsevier, 2016.
- [221] “OpenStreetMap.”
- [222] WorldPop, “Global 100m Population,” 2018.
- [223] E. R. O. A. S. E. Center, “Collection-1 U.S. Landsat Analysis Ready Data (ARD) Level-2 Data Product,” 2017. Medium: tiff.
- [224] F. H. Abanda and L. Byers, “An investigation of the impact of building orientation on energy consumption in a domestic building using emerging BIM (Building Information Modelling),” *Energy*, vol. 97, pp. 517–527, Feb. 2016.

- [225] B. Andersson, W. Place, R. Kammerud, and M. P. Scofield, “The impact of building orientation on residential heating and cooling,” *Energy and Buildings*, vol. 8, pp. 205–224, Aug. 1985.
- [226] A. Doodoo and J. Ayarkwa, “Effects of Climate Change for Thermal Comfort and Energy Performance of Residential Buildings in a Sub-Saharan African Climate,” *Buildings*, vol. 9, p. 215, Oct. 2019. Number: 10 Publisher: Multidisciplinary Digital Publishing Institute.
- [227] M. Haase and A. Amato, “An investigation of the potential for natural ventilation and building orientation to achieve thermal comfort in warm and humid climates,” *Solar Energy*, vol. 83, pp. 389–399, Mar. 2009.
- [228] S. A. Zaki, H. J. Toh, F. Yakub, A. S. Mohd Saudi, J. A. Ardila-Rey, and F. Muhammad-Sukki, “Effects of Roadside Trees and Road Orientation on Thermal Environment in a Tropical City,” *Sustainability*, vol. 12, p. 1053, Jan. 2020. Number: 3 Publisher: Multidisciplinary Digital Publishing Institute.
- [229] M. Mulero-Pázmány, L. Rollán, M. D’Amico, and M. González-Suárez, “Road orientation affects the impact of roads on wildlife,” *Wildlife Research*, vol. 50, pp. 39–46, July 2022. Publisher: CSIRO Publishing.
- [230] A. Cao, Q. Li, and Q. Meng, “Effects of Orientation of Urban Roads on the Local Thermal Environment in Guangzhou City,” *Procedia Engineering*, vol. 121, pp. 2075–2082, Jan. 2015.
- [231] G. Boeing, “Urban spatial order: street network orientation, configuration, and entropy,” *Applied Network Science*, vol. 4, pp. 1–19, Dec. 2019. Number: 1 Publisher: SpringerOpen.
- [232] “momepy.Orientation — momepy 0.5.2 documentation.”
- [233] C. Brelsford, T. Martin, J. Hand, and L. M. A. Bettencourt, “Toward cities without slums: Topology and the spatial evolution of neighborhoods,” *Science Advances*, vol. 4, p. eaar4644, Aug. 2018.
- [234] D. Zanaga, R. Van De Kerchove, D. Daems, W. De Keersmaecker, C. Brockmann, G. Kirches, J. Wevers, O. Cartus, M. Santoro, S. Fritz, and others, “ESA WorldCover 10 m 2021 v200,” 2022. Publisher: Zenodo.
- [235] “WorldCover Viewer.”
- [236] C. O. Dumitru, G. Schwarz, and M. Datcu, “Land Cover Semantic Annotation Derived from High-Resolution SAR Images,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, pp. 2215–2232, June 2016.
- [237] C. O. Dumitru, G. Schwarz, S. Cui, and M. Datcu, “Improved image classification by proper patch size selection: TerraSAR-X vs. Sentinel-1A,” in *2016 International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 1–4, IEEE, 2016.

- [238] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *Advances in Neural Information Processing Systems 25* (F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, eds.), pp. 1097–1105, Curran Associates, Inc., 2012.
- [239] J. Brownlee, “Best Practices for Preparing and Augmenting Image Data for CNNs,” May 2019. Library Catalog: machinelearningmastery.com.
- [240] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition,” in *International Conference on Learning Representations*, 2015.
- [241] “Imagery Basemaps.”
- [242] M. Caron, P. Bojanowski, A. Joulin, and M. Douze, “Deep Clustering for Unsupervised Learning of Visual Features,” in *ECCV*, 2018.
- [243] L. Bottou, “Stochastic Gradient Descent Tricks,” in *Neural Networks: Tricks of the Trade* (G. Montavon, G. B. Orr, and K.-R. Müller, eds.), vol. 7700, pp. 421–436, Berlin, Heidelberg: Springer Berlin Heidelberg, 2012. Series Title: Lecture Notes in Computer Science.
- [244] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, pp. 2278–2324, Nov. 1998. Conference Name: Proceedings of the IEEE.
- [245] M. Oquab, L. Bottou, I. Laptev, and J. Sivic, “Learning and transferring mid-level image representations using convolutional neural networks,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1717–1724, 2014.
- [246] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell, “A deep convolutional activation feature for generic visual recognition,” *UC Berkeley & ICSI, Berkeley, CA, USA*, vol. 1.
- [247] J. MacQueen, “Some methods for classification and analysis of multivariate observations,” The Regents of the University of California, 1967. ISSN: 0097-0433.
- [248] F. Lin and W. W. Cohen, “Power Iteration Clustering,” *ICML*, p. 8, 2010.
- [249] T. Chen and C. Guestrin, “XGBoost: A Scalable Tree Boosting System,” in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD ’16, (New York, NY, USA), pp. 785–794, Association for Computing Machinery, Aug. 2016.
- [250] J. H. Friedman, “Greedy function approximation: a gradient boosting machine,” *Annals of statistics*, pp. 1189–1232, 2001. Publisher: JSTOR.
- [251] S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS’17, (Red Hook, NY, USA), pp. 4768–4777, Curran Associates Inc., 2017.

- [252] J. Bergstra, D. Yamins, and D. Cox, “Making a science of model search: Hyperparameter optimization in hundreds of dimensions for vision architectures,” in *International conference on machine learning*, pp. 115–123, PMLR, 2013.
- [253] T. Lillesand, R. W. Kiefer, and J. Chipman, *Remote sensing and image interpretation*. John Wiley & Sons, 2015.
- [254] J. Yosinski, J. Clune, A. Nguyen, T. Fuchs, and H. Lipson, “Understanding Neural Networks Through Deep Visualization,” *arXiv:1506.06579 [cs]*, June 2015. arXiv: 1506.06579.
- [255] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, “Grad-CAM: Visual Explanations from Deep Networks via Gradient-based Localization,” *arXiv:1610.02391 [cs]*, Mar. 2017. arXiv: 1610.02391.
- [256] A. Chattopadhyay, A. Sarkar, P. Howlader, and V. N. Balasubramanian, “Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks,” in *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 839–847, Mar. 2018.
- [257] A. Amoah and K. Korle, “Forest depletion in Ghana: the empirical evidence and associated driver intensities,” *Forestry Economics Review*, vol. 2, pp. 61–80, Jan. 2020. Publisher: Emerald Publishing Limited.
- [258] B. Fisher, “African exception to drivers of deforestation,” *Nature Geoscience*, vol. 3, pp. 375–376, June 2010. Number: 6 Publisher: Nature Publishing Group.
- [259] D. Tuffour-Mills, P. Antwi-Agyei, and P. Addo-Fordjour, “Trends and drivers of land cover changes in a tropical urban forest in Ghana,” *Trees, Forests and People*, vol. 2, p. 100040, Dec. 2020.
- [260] C. Alliance, “Future Cities Africa,” tech. rep., 2016.
- [261] W. R. Commission, “Riparian Buffer Zone Policy,” tech. rep., 2013.
- [262] J. Frimpong, R. Adamtey, A. B. Pedersen, E. Wahaga, A. Jensen, E. Obuobie, and B. Ampomah, “A review of the design and implementation of Ghana’s National Water Policy (2007),” *Water Policy*, vol. 23, pp. 1170–1188, July 2021.
- [263] M. Poku-Boansi, C. Amoako, J. K. Owusu-Ansah, and P. B. Cobbinah, “What the state does but fails: Exploring smart options for urban flood risk management in informal Accra, Ghana,” *City and Environment Interactions*, vol. 5, p. 100038, Mar. 2020.
- [264] C. Amoako and E. Frimpong Boamah, “The three-dimensional causes of flooding in Accra, Ghana,” *International Journal of Urban Sustainable Development*, vol. 7, pp. 109–129, Jan. 2015. Publisher: Taylor & Francis eprint: <https://doi.org/10.1080/19463138.2014.984720>.

- [265] C. S. Wemegah, E. I. Yamba, J. N. A. Aryee, F. Sam, and L. K. Amekudzi, "Assessment of urban heat island warming in the greater accra region," *Scientific African*, vol. 8, p. e00426, July 2020.
- [266] "Global Human Settlement - Datasets - European Commission."
- [267] A. S. Alli, S. N. Clark, A. Hughes, J. Nimo, J. Bedford-Moses, S. Baah, J. Wang, J. Vallarino, E. Agyemang, B. Barratt, A. Beddows, F. Kelly, G. Owusu, J. Baumgartner, M. Brauer, M. Ezzati, S. Agyei-Mensah, and R. E. Arku, "Spatial-temporal patterns of ambient fine particulate matter (PM_{2.5}) and black carbon (BC) pollution in Accra," *Environmental Research Letters*, vol. 16, p. 074013, June 2021. Publisher: IOP Publishing.
- [268] J. Wang, A. S. Alli, S. Clark, A. Hughes, M. Ezzati, A. Beddows, J. Vallarino, J. Nimo, J. Bedford-Moses, S. Baah, G. Owusu, E. Agyemang, F. Kelly, B. Barratt, S. Beevers, S. Agyei-Mensah, J. Baumgartner, M. Brauer, and R. E. Arku, "Nitrogen oxides (NO and NO₂) pollution in the Accra metropolis: Spatiotemporal patterns and the role of meteorology," *Science of The Total Environment*, vol. 803, p. 149931, Jan. 2022.
- [269] "Modern cities become less dense as they grow," *The Economist*, Oct. 2019.
- [270] O. Amponsah, D. K. Blija, R. A. Ayambire, S. A. Takyi, H. Mensah, and I. Braimah, "Global urban sprawl containment strategies and their implications for rapidly urbanising cities in Ghana," *Land Use Policy*, vol. 114, p. 105979, Mar. 2022.
- [271] G. Owusu, "Coping with Urban Sprawl: A Critical Discussion of the Urban Containment Strategy in a Developing Country City, Accra," in *Cities to be tamed? Standards and alternatives in the transformation of the urban South*, vol. 1, (Milan), p. 17, Planum. The Journal of Urbanism, 2013.
- [272] H. Wellington, "In Gated cages, glazed boxes and dashed housing hopes—In remembrance of the dicey future of Ghanaian housing," in *Proceedings of the 2009 National Housing Conference, Accra, Ghana*, pp. 7–8, 2009.
- [273] G. Owusu, M. Oteng-Ababio, and R. L. Afutu-Kotey, "Conflicts and governance of landfills in a developing country city, Accra," *Landscape and Urban Planning*, vol. 104, pp. 105–113, Jan. 2012.
- [274] E. J. Adugbila, J. A. Martinez, and K. Pfeffer, "Road Infrastructure Expansion and Socio-Spatial Fragmentation in the Peri-Urban Zone in Accra, Ghana," SSRN Scholarly Paper ID 4034016, Social Science Research Network, Rochester, NY, Feb. 2022.
- [275] L. Møller-Jensen, A. N. Allotey, R. Y. Kofie, and P. W. Yankson, "A comparison of satellite-based estimates of urban agglomeration size for the Accra area," *ISPRS International Journal of Geo-Information*, vol. 9, no. 2, 2020. Publisher: MDPIAG.

- [276] F. A. Armah, D. O. Yawson, and A. A. N. M. Pappoe, “A Systems Dynamics Approach to Explore Traffic Congestion and Air Pollution Link in the City of Accra, Ghana,” *Sustainability*, vol. 2, pp. 252–265, Jan. 2010. Number: 1 Publisher: Molecular Diversity Preservation International.
- [277] Z. Zhou, K. L. Dionisio, R. E. Arku, A. Quaye, A. F. Hughes, J. Vallarino, J. D. Spengler, A. Hill, S. Agyei-Mensah, and M. Ezzati, “Household and community poverty, biomass use, and air pollution in Accra, Ghana,” *Proceedings of the National Academy of Sciences*, vol. 108, pp. 11028–11033, July 2011. Publisher: Proceedings of the National Academy of Sciences.
- [278] S. Openshaw, “The modifiable areal unit problem,” *Concepts and techniques in modern geography*, 1984. Publisher: GeoBooks.
- [279] H. S. He, S. J. Ventura, and D. J. Mladenoff, “Effects of spatial aggregation approaches on classified satellite imagery,” *International Journal of Geographical Information Science*, vol. 16, pp. 93–109, Jan. 2002. Publisher: Taylor & Francis .eprint: <https://doi.org/10.1080/13658810110075978>.
- [280] G. A. Reina, R. Panchumorthy, S. P. Thakur, A. Bastidas, and S. Bakas, “Systematic Evaluation of Image Tiling Adverse Effects on Deep Learning Semantic Segmentation,” *Frontiers in Neuroscience*, vol. 14, p. 65, Feb. 2020.
- [281] F. Isensee, P. F. Jäger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, “Automated Design of Deep Learning Methods for Biomedical Image Segmentation,” *Nature Methods*, vol. 18, pp. 203–211, Feb. 2021. arXiv:1904.08128 [cs].
- [282] T. Chen, B. Xu, C. Zhang, and C. Guestrin, “Training Deep Nets with Sublinear Memory Cost,” Apr. 2016. arXiv:1604.06174 [cs].
- [283] T. Andermann, A. Antonelli, R. L. Barrett, and D. Silvestro, “Estimating Alpha, Beta, and Gamma Diversity Through Deep Learning,” *Frontiers in Plant Science*, vol. 13, 2022.
- [284] A. Ojo, D. Vickers, and D. Ballas, “Creating a small scale area classification for understanding the economic, social and housing characteristics of small geographical areas in the Philippines,” *Regional Science Policy & Practice*, vol. 5, no. 1, pp. 1–24, 2013. .eprint: <https://rsaiconnect.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1757-7802.2012.01076.x>.
- [285] A. Skupin and A. Esperbé, “An alternative map of the United States based on an n-dimensional model of geographic space,” *Journal of Visual Languages & Computing*, vol. 22, pp. 290–304, Aug. 2011.
- [286] A. D. Singleton and P. Longley, “The internal structure of Greater London: a comparison of national and regional geodemographic models,” *Geo: Geography and Environment*, vol. 2, no. 1, pp. 69–87, 2015. .eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/geo2.7>.

- [287] M. Schonlau, “The Clustergram: A Graph for Visualizing Hierarchical and Non-hierarchical Cluster Analyses,” *The Stata Journal*, vol. 2, pp. 391–402, Dec. 2002. Publisher: SAGE Publications.
- [288] R. Maclean and F. O’Reilly, “‘Everyone’s Looking for Plastic.’ As Waste Rises, So Does Recycling,” *The New York Times*, Jan. 2022.
- [289] J. Lewis, “The Future of Mud,” July 2022. Section: Science.
- [290] J. Adounkpè, A. Lawin, C. Ahouannou, R. Akiyo, and B. Sinsin, “Modeling Solar Energy Transfer through Roof Material in Africa Sub-Saharan Regions,” *ISRN Renewable Energy*, vol. 2013, Jan. 2013.
- [291] V. Zoma and I. Dahani, “Main characteristics of urban markets in Africa,” *International Journal of Scientific Research and Management*, vol. 10, pp. 25–31, June 2022. Publisher: International Journal of scientific research and management.
- [292] A. Kumar and F. Barrett, “Stuck in traffic: Urban transport in Africa,” *World Bank Group*, 2008.
- [293] R. Sietchiping, M. J. Permezel, and C. Ngomsi, “Transport and mobility in sub-Saharan African cities: An overview of practices, lessons and options for improvements,” *Cities*, vol. 29, pp. 183–189, June 2012.
- [294] S. E. Antos, S. V. Lall, and N. Lozano-Gracia, “The Morphology of African Cities,” *World Bank Group*, Dec. 2016.
- [295] A. Zerbo, R. C. Delgado, and P. A. González, “Vulnerability and everyday health risks of urban informal settlements in Sub-Saharan Africa,” *Global Health Journal*, vol. 4, pp. 46–50, June 2020.
- [296] O. Cissé and M. Sèye, “Flooding in the suburbs of Dakar: impacts on the assets and adaptation strategies of households or communities,” *Environment and Urbanization*, vol. 28, pp. 183–204, Apr. 2016. Publisher: SAGE Publications Ltd.
- [297] R. John, “Flooding in Informal Settlements: Potentials and Limits for Household Adaptation in Dar es Salaam City, Tanzania,” *American Journal of Climate Change*, vol. 9, pp. 68–86, Apr. 2020. Number: 2 Publisher: Scientific Research Publishing.
- [298] “Mapping Dar es Salaam’s trees through artificial intelligence,” May 2022.
- [299] G. L. Gaile, “The Rural-Urban Interface in Africa: Expansion and Adaptation,” *Economic Geography*, vol. 71, pp. 330–331, July 1995. Publisher: Routledge eprint: <https://www.tandfonline.com/doi/pdf/10.2307/144319>.
- [300] F. Schug, A. Bar-Massada, A. R. Carlson, H. Cox, T. J. Hawbaker, D. Helmers, P. Hostert, D. Kaim, N. K. Kasraee, S. Martinuzzi, M. H. Mockrin, K. A. Pfoch, and V. C. Radeloff, “The global wildland–urban interface,” *Nature*, pp. 1–6, July 2023. Publisher: Nature Publishing Group.

- [301] D. Simon, D. McGregor, and K. Nsiah-Gyabaah, “The changing urban–rural interface of African cities: definitional issues and an application to Kumasi, Ghana,” 2004.
- [302] D. Lee-Smith and D. Lamba, “Nutrition and urban agriculture in Sub-Saharan African cities,” 2015.
- [303] “Why are there so many unfinished buildings in Africa?,” *The Economist*, Apr. 2021.
- [304] R. Prieto-Curiel, J. E. Patino, and B. Anderson, “Scaling of the morphology of African cities,” *Proceedings of the National Academy of Sciences*, vol. 120, p. e2214254120, Feb. 2023. Publisher: Proceedings of the National Academy of Sciences.
- [305] G. Nakweya, “Rising air pollution in African cities,” *Nature Africa*, Apr. 2022. Bandiera.abtest: a Cg_type: Research Highlight Publisher: Nature Publishing Group.
- [306] W. J. Kombe, “The State of inequalities in Sub-Saharan African and Asian cities,” 2022.
- [307] A. Storeygard, “Farther on down the Road: Transport Costs, Trade and Urban Growth in Sub-Saharan Africa,” *The Review of Economic Studies*, vol. 83, no. 3 (296), pp. 1263–1295, 2016. Publisher: [Oxford University Press, The Review of Economic Studies, Ltd.].
- [308] T. Bernatzki, M. Busse, and R. Hoekstra, “Promoting Rwanda’s business environment: Impact of reforms and drivers of change,” *Development Policy Review*, vol. 40, no. 2, p. e12578, 2022. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/dpr.12578>.
- [309] E. Ashimwe, “Kigali Special Economic Zone to be doubled in size,” Mar. 2023. Section: News.
- [310] E. E. Rwakarehe, “Review of Strategies for Curbing Traffic Congestion in Sub-Saharan Africa Cities: Technical and Policy Perspectives | Tanzania Journal of Engineering and Technology,” 2021.
- [311] R. R. Gyimah, “The hot zones are cities: Methodological outcomes and synthesis of surface urban heat island effect in Africa,” *Cogent Social Sciences*, vol. 9, p. 2165651, Dec. 2023. Publisher: Cogent OA _eprint: <https://doi.org/10.1080/23311886.2023.2165651>.
- [312] J. P. Bizimana and M. Schilling, “Geo-Information Technology for Infrastructural Flood Risk Analysis in Unplanned Settlements: A Case Study of Informal Settlement Flood Risk in the Nyabugogo Flood Plain, Kigali City, Rwanda,” in *Geospatial Techniques in Urban Hazard and Disaster Analysis* (P. S. Showalter and Y. Lu, eds.), Geotechnologies and the Environment, pp. 99–124, Dordrecht: Springer Netherlands, 2010.

- [313] C. A. Mensah, “Sustaining urban green spaces in Africa: a case study of Kumasi Metropolis, Ghana,” 2015.
- [314] P. Drechsel and S. Dongus, “Dynamics and sustainability of urban agriculture: examples from sub-Saharan Africa,” *Sustainability Science*, vol. 5, pp. 69–78, Jan. 2010.
- [315] B. Güneralp, S. Lwasa, H. Masundire, S. Parnell, and K. C. Seto, “Urbanization in Africa: challenges and opportunities for conservation,” *Environmental Research Letters*, vol. 13, p. 015002, Dec. 2017. Publisher: IOP Publishing.
- [316] B. B. Lin, S. M. Philpott, and S. Jha, “The future of urban agriculture and biodiversity-ecosystem services: Challenges and next steps,” *Basic and Applied Ecology*, vol. 16, pp. 189–201, May 2015.
- [317] M. Thondoo, O. Marquet, S. Márquez, and M. J. Nieuwenhuijsen, “Small cities, big needs: Urban transport planning in cities of developing countries,” *Journal of Transport & Health*, vol. 19, p. 100944, Dec. 2020.
- [318] M. Schmitt, S. A. Ahmadi, and R. Hänsch, “There is no data like more data – current status of machine learning datasets in remote sensing,” June 2021. arXiv:2105.11726 [cs, eess].
- [319] S. A. Ahmadi, “Awesome Satellite Benchmark Datasets,” Aug. 2023. original-date: 2021-07-03T17:18:50Z.
- [320] D. M. Kammen and D. A. Sunter, “City-integrated renewable energy for urban sustainability,” *Science*, vol. 352, pp. 922–928, May 2016. Publisher: American Association for the Advancement of Science.
- [321] A. Ramaswami, A. G. Russell, P. J. Culligan, K. R. Sharma, and E. Kumar, “Meta-principles for developing smart, sustainable, and healthy cities,” *Science*, May 2016. Publisher: American Association for the Advancement of Science.
- [322] M. Ravallion, S. Chen, and P. Sangraula, “New evidence on the urbanization of global poverty,” *Population and development review*, vol. 33, no. 4, pp. 667–701, 2007. Publisher: Wiley Online Library.
- [323] “UC Merced Land Use Dataset.” <http://weegee.vision.ucmerced.edu/datasets/landuse.html>.
- [324] “Dstl Satellite Imagery Feature Detection.” <https://kaggle.com/c/dstl-satellite-imagery-feature-detection>.
- [325] “SAT-4 and SAT-6 airborne datasets.” <https://csc.lsu.edu/~saikat/deepsat/>.
- [326] “SpaceNet Imagery.” <https://spacenetchallenge.github.io/datasets/spacenetImagery-summary.html>.
- [327] “BigEarthNet.” <http://bigearth.net/>.
- [328] “DEEPGLOBE - CVPR18 - Home.” <http://deepglobe.org>.
- [329] “Functional Map of the World Challenge.” <https://www.iarpa.gov/challenges/fmow.html>.

Appendix A

A.1 Maxar Open Data

Country	City	Date
Angola	Luanda	2017-08-25
Burkina Faso	Ouagadougou	2020-01-10
DR Congo	Kinshasa	2019-10-19
Ethiopia	Addis Ababa	2020-04-15
Ghana	Accra	2019-02-11
Ivory Coast	Abidjan	2019-06-25
Kenya	Nairobi	2019-05-15
Malawi	Blantyre	2019-10-03
Malawi	Lilongwe	2019-06-25
Malawi	Zomba	2019-05-07
Nigeria	Kano	2020-03-12
Nigeria	Lagos	2019-06-06
Nigeria	Ibadan	2020-01-22
Senegal	Dakar	2020-03-09
Sierra Leone	Freetown	2020-02-03
Sierra Leone	Bo	2020-01-27
Sierra Leone	Kenema	2020-01-15
Sierra Leone	Makeni	2020-01-14
Somalia	Mogadishu	2019-06-21
Tanzania	Dodoma	2019-05-21
Tanzania	Dar Es Salaam	2018-09-17

Table A.1: List of available satellite basemap imagery by Maxar Open Data [2]

A.2 Maxar internal use license

END USER LICENSE TERMS INTERNAL USE LICENSE

Version F4-21-21

These Internal Use License terms (the “License Terms”) apply to your use of a Product if you have licensed the Product subject to an Internal Use License either from an Affiliate of Maxar Technologies Inc., a Delaware corporation with offices located at 1300 W. 120th Avenue, Westminster, Colorado 80234 USA (“Maxar”) directly or from a Certified Reseller of Maxar. These License Terms are entered into by Maxar and Customer. These License Terms contain the general terms relating to Customer’s access to and use of the Product. The applicable Customer Agreement sets forth the terms pursuant to which Customer purchased the Internal Use License.

By signing or otherwise indicating acceptance of a Customer Agreement or downloading, accessing or using any Product licensed under an Internal Use License, you, on behalf of Customer, are accepting and agreeing to be bound by these License Terms. If you are entering into these License Terms on behalf of a company or other legal entity or government agency, you represent that you have the authority to bind that entity and its Affiliates to the terms and conditions of these License Terms. If you do not agree to the terms in these License Terms, do not download, access or use any Maxar Product. Capitalized terms used in these License Terms are defined in Section 14 of these License Terms.

1. GRANT OF LICENSE. Subject to Customer’s compliance with these License Terms and the applicable Customer Agreement, including, without limitation, payment of all applicable fees, during the Term, Maxar grants to Customer a non-exclusive, non-transferable, limited license to allow an unlimited number of its Authorized Users to:

- (a) store, access, evaluate, use and reproduce the Product solely for Customer’s Internal Use;
- (b) process, modify, enhance, adapt and create Derivatives of the Product (including Data Derivatives as defined in Section 2.3 of these License Terms depending on the type of Product) via formatting, editing, digitization, data combination; or via extraction of geographic features, human-made features, persons or animals and related data via identification, measurement and/or analysis, and store, access, evaluate, use and reproduce those Derivatives solely for Customer’s Internal Use; provided, however, that Data Derivatives can be used for any and all purposes, subject to the attribution requirements set forth in Section 7 of these License Terms; and
- (c) display an extract of the Product (excluding Information Products) or a Derivative thereof on a public website in a secure, non-extractable, and non-downloadable format that prohibits manipulation of any pixel or metadata contained in the Product or Derivative thereof and does not result in any revenue to Customer, as follows:
 - (i) on one domain name;
 - (ii) 2048 x 2048 pixels;
 - (iii) at a resolution no better than resolution of imagery in the Product;
 - (iv) .png, .gif, .jpg, .jpeg, .jpe, .jiff, .bmp, .pdf; or any format without geo-referencing information (TIFF, NITF, GeoPDF, JP2 and JPEG2000 are not permitted);
 - (v) only collection date/time, vehicle, and band combination metadata can be published with the extract; and
 - (vi) properly attribute the imagery as required in Section 7 of these License Terms.

2. SUBLICENSE RIGHTS.

2.1 GENERALLY. Customer may not sublicense to any Third Party the rights granted to Customer in Section 1 of these License Terms. Customer will ensure that each Authorized User complies with these License Terms. Customer will be liable for all acts and omissions of its Authorized Users relating to the Product or any violation of these License Terms. A breach of these License Terms by an Authorized User is deemed to be a breach by Customer. If Customer wants to sublicense the Product to an Affiliate of Customer, Customer must purchase a Group License.

2.2 RESERVED.

2.3 DATA DERIVATIVES. “Data Derivatives” are those Derivatives created by Customer that Customer is permitted to use for any and all purposes. With respect to each Product type, the following Derivatives are deemed to be “Data Derivatives”:

- (a) Elevation Products. With respect to Elevation Products, a Data Derivative is a Derivative of the Elevation Product that does not contain any elevation, depth or height information from the source Elevation Product.
- (b) Imagery Products. With respect to Imagery Products, a Data Derivative is a Derivative of the Imagery Product that does not contain any imagery data or metadata from the source Imagery Product and is irreversible and uncoupled from the imagery

data in the source Imagery Product. However, Data Derivatives specifically exclude the following Derivatives: orthorectified imagery; PAN, MS and PAN-Sharpener imagery; and Digital Elevation Models (DEM), Digital Surface Models (DSM), Digital Terrain Models (DTM), Triangulated Irregular Network (TIN) and Point Cloud elevation models (including, without limitation, 3D and building models, bathymetry and sea floor mapping).

- (c) Information Products. With respect to Information Products, there are no Data Derivatives.
- (d) Tasking Products. With respect to Tasking Products, a Data Derivative will be determined based upon the type of Product collected. For example, if the Tasking Product results in the generation of an Imagery Product, then the permitted Data Derivatives for the Tasking Product are the same as those for the resulting Imagery Product.

3. ADDITIONAL LICENSE RIGHTS.

- 3.1 **EDUCATION LICENSE RIGHTS.** If Customer licenses a Product pursuant to the Education Discount, in addition to the rights set forth in Section 1 of these License Terms, during the Term, Maxar grants to Customer a non-exclusive, non-transferable, limited license to use the Product for research purposes. However, Customer must obtain Maxar's prior written consent before publishing any Product or Derivative for research purposes.
- 3.2 **NGO/GDO LICENSE RIGHTS.** If Customer licenses a Product pursuant to the NGO/GDO Discount, in addition to the display rights set forth in Section 1 of these License Terms, during the Term, Maxar grants to Customer a non-exclusive, non-transferable, limited license to publish extracts of the Product or a Data Derivative to supplement text or newsworthy events in a hardcopy, broadcast or electronic format in publications, on a website, in a video/movie or other similar media, subject to the same formatting restrictions and requirements set forth in Section 1 of these License Terms. However, Customer must obtain Maxar's prior written consent before publishing any extracts of the Products or Derivative.
- 4. **LICENSE TERM.** The Term of the Internal Use License for each Product will begin upon delivery of the Product to Customer or Certified Reseller, whichever occurs first, and will continue perpetually or for a one-year term, as set forth in the Customer Agreement, unless terminated as set forth in Section 12 of these License Terms. If the License is for a one-year term, then upon expiration of each one-year term, the license Term will automatically renew and Customer will be invoiced for the applicable license fees under the Customer Agreement, unless Customer notifies Maxar or Certified Reseller at least thirty (30) days prior to the end of the then current Term that Customer does not want to renew the license.
- 5. **USE RESTRICTIONS.** Customer recognizes and agrees that the Product is the property of Maxar and contains valuable assets and proprietary information of Maxar and its suppliers, as applicable. Any Third Party Content included in a Product is subject to the terms and conditions of any end user license agreement or additional terms accompanying the Product, specified in the Product Specification, and/or posted on Maxar's website (in each instance, "Third Party Terms"). Third Party Terms are incorporated by referenced into this Agreement and a breach of Third Party Terms will be deemed a breach of this Agreement. In addition, except as expressly permitted in Sections 1, 2, and 3 of these License Terms, Customer will not, and will not permit any Authorized User or other Third Party to:
 - (a) distribute, sublicense, transfer, assign, rent, sell, lease, loan, make publicly available, publish or otherwise convey access to or use of the Product or Derivatives (other than Data Derivatives) to anyone other than Authorized Users;
 - (b) use the Product or Derivatives (other than Data Derivatives) for the business needs of any Third Party, including without limitation, providing any services to any Third Parties;
 - (c) store, post or process the Product or Derivatives in a system or platform that is publicly-accessible;
 - (d) use the Product or Derivatives to improve the accuracy of any other satellite imagery via algorithmic processing or any other method;
 - (e) with respect to Information Products, use or attempt to use the Information Product or Derivatives thereof to train machine learning algorithms for feature extraction or any other purpose;
 - (f) remove, bypass or circumvent any electronic or other forms of protection included on or with the Product;
 - (g) alter, obscure or remove any copyright notice, copyright management information or proprietary legend contained in or on the Product;
 - (h) modify or use the Product and Derivatives in any manner that infringes upon the Intellectual Property Rights of another entity or individual, or violates any applicable laws;
 - (i) reverse engineer, disassemble, decompile, adapt or otherwise attempt to derive the algorithms, source code, databases or data structures upon which the Product is based, but only to the extent this restriction is permitted by law; or

- (j) otherwise use or access the Product or any Derivative for any purpose not expressly permitted under these License Terms, including, without limitation, for Commercial Purposes.

6. **OWNERSHIP.** All right, title and interest in and to the Product and all corrections, enhancements, or other modifications to the Product made by Maxar or any Third Party at Maxar's direction, and all Intellectual Property Rights therein are the sole and exclusive property of Maxar or its suppliers, as applicable. All right, title and interest, including all Intellectual Property Rights, in and to enhancements or modifications made by Customer in the creation of a Derivative and any new material contributed by Customer in the creation of a Derivative, but specifically excluding materials owned by Maxar or its suppliers (including, without limitation, Products integrated, referenced, recast, transformed or adapted in the Derivative) are the exclusive property of Customer. However, notwithstanding the ownership rights of Customer in the enhancements, modifications and contributed materials, use of a Derivative by Customer is subject to the license and use restrictions set forth in Sections 1, 2, 3 and 5 of these License Terms. All rights not expressly granted to Customer in these License Terms are reserved by Maxar.
7. **ATTRIBUTION.** Customer will not delete, alter, cover or distort any copyright, trademark or other proprietary rights notice placed on or in the Product and will ensure that all notices are reproduced on all copies. All Derivatives (including Data Derivatives) must include the following copyright notice on or adjacent to the Derivative:
 - (a) Maxar Products. [Product] © [YEAR] Maxar Technologies.
 - (b) AW3D Enhanced, AW3D Metro, AW3D Standard and AW3D Telecom Products. [Product] © [YEAR] NTT Data Corporation.
 - (c) Ecopia Building Footprints Powered by Maxar. Ecopia Building Footprints © [YEAR] Ecopia Tech Corporation. Imagery © [YEAR] Maxar Technologies.
8. **CERTIFICATION AND AUDIT.** Upon Maxar's written request, and not more than once per calendar year, Customer will certify its compliance with the Customer Agreement and these License Terms. Upon thirty (30) days written notice and no more than once every twelve (12) months, Maxar may audit Customer's compliance with the Customer Agreement and these License Terms. Customer will cooperate with Maxar's audit and provide reasonable assistance and access to information and records related to this Agreement and Customer's use and/or distribution of the Product. Audits will not unreasonably interfere with Customer's normal business operations and will be subject to reasonable confidentiality requirements. If an audit results in a finding of non-compliance, Maxar may, at its discretion: (a) invoice any additional fees due with interest as set forth herein and recover the cost of the audit if additional fees exceed five percent (5%) of the fees paid during the audit period; and (b) terminate the Customer Agreement and these License Terms in accordance with Section 12.1 below. Customer must pay the invoices issued under this Section within thirty (30) days following the date of invoice.
9. **INDEMNIFICATION BY CUSTOMER.** Customer will defend, indemnify and hold Maxar, its Affiliates, its suppliers and Certified Reseller harmless from and against any claims that may arise against Maxar, its Affiliates, its suppliers or Certified Reseller out of Customer's use of the Product, including, without limitation, a violation by Customer of Section 13.6, 13.7 or 13.8 of these License Terms.
10. **LIMITED WARRANTY AND DISCLAIMER.**
 - 10.1 **LIMITED WARRANTY.** Maxar warrants to Customer only that the Product, as delivered by Maxar, will (a) be of the area of interest set forth in the applicable Customer Agreement; and (b) comply in all material respects with the applicable Product Specification. Maxar's sole obligation and Customer's exclusive remedy for a breach of this warranty is for Maxar, at its option and expense, to: (i) repair or replace the non-conforming Product; or (ii) terminate the applicable license and refund all fees paid by Customer for the non-compliant Product. Any claim under this warranty must be made within thirty (30) days following the initial delivery of the Product. This limited warranty is void if any non-conformity has resulted from any accident, abuse, misuse, misapplication or modification of or to the Product by anyone other than Maxar or any breach by Customer of these License Terms.
 - 10.2 **DISCLAIMER OF WARRANTIES.** EXCEPT AS EXPRESSLY WARRANTED IN SECTION 10.1, THE PRODUCT IS PROVIDED "AS IS," WITHOUT ANY WARRANTY OF ANY KIND, WHETHER EXPRESS, IMPLIED OR STATUTORY, INCLUDING, WITHOUT LIMITATION, WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE, TITLE, NON-INFRINGEMENT OR NON-MISAPPROPRIATION OF INTELLECTUAL PROPERTY RIGHTS OF A THIRD PARTY, CUSTOM, TRADE, QUIET ENJOYMENT, ACCURACY OF INFORMATION, CONTENT OR RESULTS, OR CONDITIONS ARISING UNDER ANY OTHER LEGAL REQUIREMENT. NEITHER MAXAR NOR ITS AFFILIATES, OR SUPPLIERS WARRANT THAT THE PRODUCT WILL BE ACCURATE, CURRENT OR COMPLETE, THAT THE PRODUCT WILL MEET CUSTOMER'S NEEDS OR EXPECTATIONS OR THAT THE OPERATION OF THE PRODUCT WILL BE ERROR FREE OR UNINTERRUPTED. FURTHER, SPATIAL, SPECTRAL AND TEMPORAL ACCURACY IS NOT GUARANTEED.
11. **LIMITATION OF LIABILITY.** IN NO EVENT WILL MAXAR, ITS AFFILIATES OR ITS SUPPLIERS BE LIABLE FOR ANY SPECIAL, INDIRECT, INCIDENTAL, EXEMPLARY, PUNITIVE OR CONSEQUENTIAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOSS OR DAMAGE TO DATA, INACCURACY OF DATA, LOSS OF ANTICIPATED REVENUE OR PROFITS, WORK STOPPAGE OR IMPAIRMENT OF OTHER ASSETS OR LOSS OF GOOD WILL, WHETHER OR NOT FORESEEABLE AND WHETHER OR NOT A PARTY HAS BEEN ADVISED OF THE POSSIBILITY OF THE DAMAGES AND

NOTWITHSTANDING ANY FAILURE OF THE ESSENTIAL PURPOSE OF THESE LICENSE TERMS OR ANY LIMITED REMEDY HEREUNDER. IN NO EVENT WILL THE TOTAL LIABILITY OF MAXAR, ITS AFFILIATES AND ITS SUPPLIERS ARISING OUT OF OR IN CONNECTION WITH THE PRODUCT(S) EXCEED THE FEES PAID BY CUSTOMER FOR THE PRODUCT(S) GIVING RISE TO THE CLAIM. THE FOREGOING LIMITATIONS APPLY TO ALL CAUSES OF ACTION IN THE AGGREGATE, INCLUDING WITHOUT LIMITATION, BREACH OF CONTRACT, BREACH OF WARRANTY, INDEMNIFICATION, NEGLIGENCE, STRICT LIABILITY, MISREPRESENTATION AND OTHER TORTS AND STATUTORY CLAIMS.

12. TERMINATION.

- 12.1 **BY MAXAR.** In addition to those termination rights set forth in Section 10, Maxar may terminate these License Terms and associated license rights upon written notice to Customer if Customer (a) breaches Section 1, 2, 3, 5, 6, 8, 9, 13.6, 13.7, or 13.8 of these License Terms; or (b) materially breaches any other provision of these License Terms or the Customer Agreement and fails to cure the breach within thirty (30) days after receiving written notice to do so.
- 12.2 **BY CUSTOMER.** Customer may terminate these License Terms and associated license rights at any time by (a) permanently deleting the Product and Derivatives (excluding Data Derivatives, which are owned by Customer) from all devices and systems and destroying any copies on disk; and (b) certifying to Maxar in writing that all copies of the Product and Derivatives (excluding Data Derivatives, which are owned by Customer) have been deleted or destroyed; however, Customer is still responsible for paying all license fees in full.
- 12.3 **OBLIGATIONS UPON TERMINATION.** Upon termination or expiration of these License Terms, all rights to use the Product granted to Customer under these License Terms will immediately cease and Customer will (and will cause all Authorized Users to) (a) stop all use of the Product and Derivatives (excluding Data Derivatives) and (b) permanently delete the Product and Derivatives (excluding Data Derivatives) from all devices and systems and destroy any copies on disk. Within ten (10) days following termination or expiration, Customer will certify to Maxar in writing that all copies of the Product and Derivatives (excluding Data Derivatives) licensed under these License Terms have been deleted or destroyed. The expiration or termination of these License Terms does not relieve either party of any obligations that have accrued on or before the effective date of the termination or expiration.
- 12.4 **SURVIVAL.** The duties and obligations of the parties under Sections 2.3 (Data Derivatives), 5 (Use Restrictions), 6 (Ownership), 7 (Attribution), 8 (Certification and Audit), 9 (Indemnification), 11 (Limitation of Liability), 12.3 (Obligations upon Termination), 12.4 (Survival), and 13 (General Terms) of these License Terms will survive expiration or termination of these License Terms.

13. GENERAL TERMS.

- 13.1 **ENTIRE AGREEMENT.** These License Terms, together with the Customer Agreement, constitute the entire agreement between the parties with respect to use of the Product and supersede all previous and contemporaneous agreements, understandings and arrangements, whether oral or written.
- 13.2 **ASSIGNMENT.** Customer may not transfer or assign any of its rights or delegate any of its obligations under these License Terms, in whole or in part and including any transfers by operation of law, without the prior written consent of Maxar. Any attempted assignment or transfer in violation of this Section will be null and void. These License Terms will be binding on and inure to the benefit of the parties and their respective permitted successors and assigns.
- 13.3 **AMENDMENT.** These License Terms may be amended or supplemented only by a writing that refers to these License Terms and that is signed by both parties.
- 13.4 **WAIVER.** All waivers must be in writing and signed by the party granting the waiver. A waiver of any of rights or remedies in a particular instance will not be construed as a waiver of the same or a different right or remedy in a subsequent instance. The failure or delay by a party to require performance of any provision of these License Terms does not constitute a waiver.
- 13.5 **SEVERABILITY.** If any provision of these License Terms is invalid, illegal or unenforceable, that provision will be deemed to be restated so that it is enforceable to the maximum extent permissible under law and is consistent with the original intent and economic terms of the invalid provision.
- 13.6 **COMPLIANCE WITH LAWS.** Customer is responsible for its own compliance with laws, regulations and other legal requirements applicable to the conduct of its business, use of the Product and these License Terms, and agrees to comply with all these laws, regulations and other legal requirements including, without limitation, the Foreign Corrupt Practices Act of the United States of America and the Convention on Combating Bribery of Foreign Government Officials.
- 13.7 **INTERNATIONAL TRADE COMPLIANCE.** The Products are subject to the customs and export control laws and regulations of the United States and any country in which the Products are manufactured, received or used, including, without limitation, the Export Administration Regulations and the International Traffic in Arms Regulations. Customer will comply with these laws, regulations and rules in the performance of its obligations under these License Terms. Further, Customer will not provide the Products to blocked, prohibited or restricted individuals and entities as required by the U.S. Department of Treasury Office of Foreign Assets Control ("OFAC"), including, without limitation, the Denied Persons List, Unverified List and Entity List. Customer

will not knowingly do business with criminal organizations, terrorist organizations or other people or groups, either directly or indirectly, that are likely to use the Products for purposes that are illegal or adverse to the interests of the United States Government or Maxar generally. Customer will cooperate with Maxar to ensure ongoing compliance with all laws, regulations and other legal requirements applicable to the conduct of its business and these License Terms and will provide Maxar with the assurances and official documents that Maxar may request periodically to verify Customer's compliance with these License Terms.

13.8 DATA PROTECTION. If delivery, creation, or Customer's use of the Product will involve the Processing of Personal Data, Customer is solely responsible for its compliance at all times with applicable laws, regulations and other legal requirements related to such use, including, without limitation, obtaining any and all applicable approvals and consents necessary for such use from any Data Subjects and regulatory authorities. Without limiting the foregoing, Customer will ensure that it has in place a privacy policy that provides transparent communication of the Processing activities and the rights of Data Subjects. Additionally, Customer will employ adequate technical and organizational security measures to protect Personal Data against a Personal Data breach.

13.9 GOVERNING LAW AND DISPUTE RESOLUTION. All matters and disputes arising out of or in connection with this License Agreement will be governed by and construed under the laws and using the method of dispute resolution indicated below, based upon where Customer is domiciled:

If Customer is domiciled in:	The governing law is:	Any suit, action or proceeding arising out of or relating to this Agreement must be:
United States of America, Canada, or Mexico	New York and controlling United States federal law	Instituted in the United States District Court for the Southern District of New York or the state courts located in New York, New York
Any other country	Laws of England & Wales	Finally settled under the Rules of Arbitration of the International Chamber of Commerce by arbitrator(s) appointed in accordance with those Rules. The place of arbitration shall be New York, New York, and any and all awards and other decisions shall be deemed to have been made there, without prejudice to the right of the arbitral tribunal to hold hearings, meetings, or sessions any place it deems appropriate.

Each party irrevocably submits to the exclusive jurisdiction of the applicable court set forth above. However, each party will have the right at any time to seek a temporary or permanent injunction or other equitable remedy or relief in any court having subject matter jurisdiction anywhere in the world. With respect to arbitration, there will be one arbitrator and the language of arbitration will be English. The party or parties prevailing in any legal actions, arbitration or other proceeding relating to this License Agreement, whether in arbitration, at trial or upon appeal, will be entitled to recover reasonable attorneys' fees and other costs and expenses incurred, in addition to any other relief to which it may be entitled. The parties agree that the United Nations Convention on Contracts for the International Sale of Goods does not apply to this License Agreement.

13.10 NOTICES. Except for routine operational correspondence, all notices made or given pursuant to these License Terms must be in writing and will be considered delivered and effective upon receipt (or when delivery is refused) when: (a) personally delivered; (b) sent by registered or certified mail (postage prepaid, return receipt requested); (c) sent by nationally-recognized private carrier (e.g., Federal Express, DHL, etc.) (with signature required and all fees prepaid); or (d) sent by email with confirmation of transmission. Notices will be sent to Customer at the address set forth in the Order Confirmation (or if none is specified, the address to which Maxar sends invoices). Notices to Maxar must be sent to (a) legalservices@maxar.com, if electronic; and (b) to 1300 W. 120th Avenue, Westminster, Colorado 80234, USA, attention Legal Department, if sent in hard copy.

13.11 CONTROLLING LANGUAGE. These License Terms are drafted in the English language only. English will be the controlling language in all respects, and all versions of these License Terms in any other language are for accommodation only and will not be binding on the parties. All communications and notices to be made or given pursuant to these License Terms must be in the English language. The titles and headings herein are for reference purposes only.

13.12 FORCE MAJEURE. Except for Customer's obligation to make payment under the Customer Agreement or these License Terms, neither party will be liable for any failure or delay in fulfilling or performing any term of these License Terms when and to the extent the failure or delay is caused by or results from acts or events beyond that party's reasonable control, including, without limitation: acts of God; fire; water damage; natural disaster (including earthquakes, storms, and floods); power or utility outages; strikes; war, military action, or act of terrorism; medical crisis, pandemic or epidemic; a total or partial loss, malfunction, or failure of a satellite, ground station, or communications network, whether temporary or permanent; a change in law or regulation (including export control regulations); acts, directives and orders of government and health authorities; or an order or judgment of a court (not arising out of breach by the party of these License Terms). The party suffering a force majeure event will promptly give notice to the other party, stating the period of time the occurrence is expected to continue.

13.13 EQUITABLE REMEDIES. The parties agree that a breach or threatened breach by Customer of its obligations under these License Terms would give rise to irreparable harm to Maxar and that Maxar will be entitled to seek equitable relief (without any

requirement to post bond), including injunctive relief or specific performance of the terms, in addition to any other remedy to which it is entitled at law or in equity.

14. DEFINITIONS.

"Affiliate" means any legal entity controlling, controlled by or under common control with a party, where "control" means (a) the ownership of at least fifty percent (50%) of the equity or beneficial interest of the entity; (b) the right to vote for or appoint a majority of the board of directors or other governing body of the entity; or (c) the power to direct or cause the direction of the management and policies of such party by any means.

"Authorized User" means an employee or Contractor that is authorized by Customer to use the Product on behalf of Customer, or if the Customer is a university or school, then Authorized Users may include Customer's currently-enrolled students.

"Certified Reseller" means a reseller authorized by Maxar to resell licenses to use the Products.

"Commercial Purpose" means redistribution, retransmission or publication in exchange for a fee or other consideration, which may include, without limitation: (a) advertising; (b) use in marketing and promotional materials and services on behalf of a customer, client, employer, employee or for Customer's benefit; (c) use in any materials or services for sale or for which fees or charges are paid or received; and (d) use in any books, news publication or journal.

"Contractor" means an individual contracted by Customer, either directly or through a consulting company or other entity, to provide services on behalf of or for the benefit of Customer. For avoidance of doubt, "Contractor" includes a volunteer who is under contract with Customer to provide services on behalf of Customer, provided that the Customer maintains direct control over all access to, and copies of, Products and/or Derivatives, and the volunteer is not permitted to use the Products or Derivatives for any purpose other than Customer's permitted uses.

"Customer Agreement" means (a) with respect to a Customer that purchases a license to use the Product from Maxar directly, that agreement consisting of the applicable Order Confirmation and Product Terms and Conditions, which reference these License Terms; and (b) with respect to a Customer that purchases a license to use the Product from a Certified Reseller, that agreement between the Certified Reseller and Customer pursuant to which Customer receives a license to the Product.

"Customer" means that individual, legal entity or government agency that has purchased a license to use the applicable Product either directly from Maxar or from a Certified Reseller.

"Data Derivative" means, generally, a Derivative of the Product that may be used for any and all purposes in accordance with Section 1(b) of these License Terms. With respect to each type of Product, the Data Derivatives are described in Section 2.3 of these License Terms.

"Data Subject" means an identifiable natural person.

"Derivative" means any addition, improvement, update, modification, transformation, adaptation or derivative work of or to a Product, including, without limitation, reformatting of the Product into a different format or media from which it is delivered to Customer; any addition or extraction of data, information or other content to or from the Product; or any copy or reproduction of the Product.

"Education Discount" means that discount that is extended to a Customer that is a university, college, technical training institute or school utilizing the Product solely for educational purposes.

"Elevation Product" means a Product that consists of data sets, vectors, or other information that depict or indicate elevation, depth, building heights, 3D renderings of objects on the surface of the earth, or other three dimensional features. Elevation Products include Digital Elevation Models (DEM), Digital Surface Models (DSM), Digital Terrain Models (DTM), Triangulated Irregular Network (TIN) and Point Cloud elevation models (including, without limitation, 3D and building models, bathymetry and sea floor mapping).

"Imagery Product" means a Product that consists of aerial photography, satellite imagery and other imagery data owned or controlled by Maxar, including all metadata. "Imagery Product" does not include satellite or aerial imagery owned by Third Party suppliers and licensed under Third Party Terms.

"Information Product" means a Product that consists of extracted data layers, shapefiles, vectors, summary, analysis or other report, dataset or other information that may include excerpts of satellite imagery or aerial photography.

"Intellectual Property Rights" means all past, present, and future trade secret rights, patent rights, copyrights, moral rights, contract rights, trademark rights, service marks, and other proprietary rights in any jurisdiction, including those rights in inventions, software, domain names, know-how, methods, processes, information and technology.

"Internal Use" means use of the Product and permitted Derivatives solely for the internal business purposes of Customer (subject to those restrictions set forth in Section 5 of these License Terms) and not for any Commercial Purpose.

"License Terms" has the meaning set forth in the Preamble.

"NGO/GDO Discount" means that discount that is extended to a Customer that (a) is a non-governmental, non-profit organization or a global development organization, and (b) contributes to or participates in cooperation projects, education, training or other humanitarian, progressive or watchdog activities.

"Order Confirmation" means that agreement or other document prepared by Maxar that sets forth the Product(s) Maxar offers to license to Customer and related terms and that is presented to Customer for acceptance. A quotation that includes an estimated fee is not an Order Confirmation.

"Personal Data" means any information that directly or indirectly identifies a Data Subject, such as a name, an identification number, location data, an online identifier or one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

"Processing" means any operation that is performed on Personal Data, whether or not by automated means, such as collection, storage, alteration, use, dissemination or destruction.

"Product Specification" means, with respect to each Product, the description and specification published by Maxar and available upon request or at <https://www.maxar.com/legal>.

"Product Terms and Conditions" means the Product Terms and Conditions pursuant to which Maxar provides the Product to Customer, available at <https://www.maxar.com/legal>.

"Product(s)" means those product(s) and/or service(s) licensed by Customer, as described in the Customer Agreement and Product Specification, which may include, without limitation, Elevation Products, Imagery Products, Information Products, Subscription Services Products, and Tasking Products.

"Tasking Product" means a Product that allows Customer to designate when and where imagery or other data should be collected by a constellation of satellites.

"Term" means that period of time that Customer is entitled to use the Product as set forth in the Customer Agreement and further defined in Section 4 of these License Terms.

"Third Party" means any individual, legal entity, corporation, limited liability company, partnership, other organization or government agency that is not a party to this Agreement and is not an Affiliate of Maxar.

"Third Party Content" means any content, software or other data that is owned by a Third Party and not owned by Maxar or its Affiliates.

"Third Party Terms" has the meaning set forth in Section 5 above.

-- END OF INTERNAL USE LICENSE --

A.3 Satellite image datasets

Dataset List				
Dataset name	Year	Type	Description	Ref
UC Merced	2010	Scene classification (patch)	Satellite imagery tiles (256x256 pixels), 21 land use classes	[323]
Land Use				
DSTL	2017	Semantic segmentation	10 land cover categories from crops to vehicle small, 57 1x1km images	[324]
Deepsat	2015	Scene classification (patch)	6 land cover categories, 400k 28 x 28 pixel chips	[325]
SpaceNet	2017/18/20	Instance segmentation	Different challenges including 27 WorldView 2 images	[326]
BigEarthNet	2019	Scene classification (patch)	Multiple land cover labels per chip, 590.326 chips	[327]
DeepGlobe	2018	Semantic segmentation	and Cover challenge: 3 different cities with images between 6,000-10,000 pixels	[328]
Functional Map of the World	2017	Scene classification (patch)	Over 1 million images from over 200 countries (256 x 256 x 4/8 bandas)	[329]

Figure A.1: List of labelled satellite image datasets considered for transfer learning. The table was created in 2020 and some datasets have been updated since.

A.4 XGBoost hyperparameters, Accra model

Hyperparameter	Value
Percentage of features used for building each tree (colsample_bytree)	0.6710635126567753
Min split loss (gamma)	2.9237052184478376
Max depth of tree	6
Minimum sum of instance weight (hessian) needed in a child	3.0
L1 regularisation term (reg_alpha)	48.0
L2 regularisation term (reg_lambda)	0.5286917420754508

Table A.2: Best hyperparameters for XGBoost model predicting cluster membership in Accra.

A.5 Hyperparameter experiments for multi-city clustering

A.5.1 Approach A

LR	0.01	1e-2	1e-3	1e-4	0.01	1e-2	1e-3	1e-4	0.01	1e-2	1e-3	1e-4
K	8	8	8	8	12	12	12	12	16	16	16	16
B.	0.1954	0.2183	0.2134	0.2118	0.1916	0.2289	0.2133	0.2074	0.1922	0.1617	0.1428	0.1818
area												
B.	0.1987	0.2287	0.1987	0.1987	0.2102	0.2232	0.2102	0.2072	0.2097	0.1772	0.1652	0.1907
count												
Avg.	0.2293	0.2253	0.2308	0.2269	0.2239	0.2397	0.2083	0.2237	0.2112	0.2058	0.1959	0.2252
b.												
size												
B.	0.2226	0.2397	0.2007	0.1989	0.2313	0.2338	0.2134	0.1967	0.2418	0.2455	0.2476	0.2101
ori-												
ent.												
All	0.2244	0.2143	0.2016	0.2013	0.2225	0.2262	0.1822	0.2015	0.2135	0.2009	0.1970	0.2172
roads												
M.	0.2431	0.2471	0.2414	0.2407	0.2391	0.2468	0.2391	0.2405	0.2458	0.2431	0.2432	0.2445
roads												
Dist.	0.2186	0.2157	0.2032	0.2016	0.2127	0.2257	0.1814	0.1977	0.1881	0.1912	0.1844	0.2047
all												
roads												
Dist.	0.2322	0.2366	0.2352	0.2366	0.2416	0.2368	0.2343	0.2344	0.2327	0.2303	0.2363	0.2272
m												
roads												
Pop.	0.1844	0.2183	0.2024	0.2104	0.1980	0.2105	0.1990	0.1982	0.1861	0.1742	0.1543	0.1900
den.												
NDVI	0.1940	0.1831	0.1671	0.1618	0.1969	0.2007	0.1818	0.1654	0.1604	0.1431	0.1448	0.1637
Total	2.1427*	2.2271*	2.0945	2.0887	2.1678*	2.2723*	2.0630	2.0727	2.0815	1.9730	1.9115	2.0551*

Table A.3: Approach A, across all cities. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics. Analyses with a * return inherent image stitching of the Accra image that is part of the commercial pre-processing. Running the model with K=16 and LR of 0.1 resulted in an error.

A.5.2 Approach B

Learning rate	0.1	0.0001	0.00001	0.000001
B. area	0.1871	0.1611	0.1427	0.1381
B. count	0.1872	0.1627	0.1401	0.1401
Avg. b. size	0.2338	0.2330	0.2189	0.2171
B. orient.	0.1682	0.2122	0.2134	0.2029
All roads	0.2249	0.2017	0.1895	0.1870
M. roads	0.2253	0.2356	0.2350	0.2335
Dist. all roads	0.2086	0.2011	0.1936	0.1919
Dist. m. roads	0.1912	0.1744	0.1623	0.1585
Pop. den.	0.2013	0.1878	0.1785	0.1817
NDVI	0.2036	0.1845	0.1706	0.1610
Total	2.0312	1.9541	1.8446	1.8118

Table A.4: Approach B, Dakar. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.

Learning rate	0.1	0.01	0.001	0.0001	0.00001
B. area	0.1929	0.1772	0.1992	0.2074	0.1974
B. count	0.2002	0.1822	0.1927	0.2212	0.1992
Avg. b. size	0.2448	0.2425	0.2450	0.2461	0.2467
B. orient.	0.2504	0.2493	0.2478	0.2489	0.2510
All roads	0.2491	0.2485	0.2490	0.2491	0.2492
M. roads	0.2480	0.2500	0.2484	0.2468	0.2537
Dist. all roads	0.2244	0.2134	0.2173	0.2280	0.2250
Dist. m. roads	0.2325	0.2256	0.2316	0.2375	0.2322
Pop. den.	0.1878	0.1696	0.1941	0.2056	0.1918
NDVI	0.1943	0.1804	0.1950	0.2016	0.1898
Total	2.2244	2.1387	2.2201	2.2922	2.2360

Table A.5: Approach B, Des. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.

Learning rate	0.1	0.01	0.001	0.0001	0.00001
B. area	0.1651	0.1795	0.1695	0.1746	0.1477
B. count	0.1957	0.1947	0.1917	0.2062	0.1637
Avg. b. size	0.2063	0.2290	0.2144	0.2171	0.2077
B. orient.	0.2350	0.2442	0.2167	0.2473	0.2031
All roads	0.2293	0.2379	0.2289	0.2397	0.2354
M. roads	0.2387	0.2457	0.2332	0.2475	0.2383
Dist. all roads	0.2038	0.2010	0.2029	0.1804	0.1734
Dist. m. roads	0.2290	0.2355	0.2294	0.2328	0.2228
Pop. den.	0.2059	0.2069	0.1984	0.1994	0.1907
NDVI	0.2003	0.2104	0.1956	0.2020	0.1873
Total	2.1091	2.1848	2.0807	2.1470	1.9701

Table A.6: Approach B, Kigali. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.

A.5.3 Approach C

Learning rate	0.1	0.01	0.001	0.0001	0.00001
B. area	0.1729	0.1624	0.1576	0.1520	0.1474
B. count	0.1802	0.1682	0.1652	0.1577	0.1532
Avg. b. size	0.1745	0.1757	0.1761	0.1730	0.1745
B. orient.	0.2301	0.1968	0.1864	0.1804	0.1495
All roads	0.1505	0.1515	0.1514	0.1521	0.1509
M. roads	0.2291	0.2334	0.2364	0.2392	0.2330
Dist. all roads	0.1445	0.1524	0.1476	0.1431	0.1389
Dist. m. roads	0.1729	0.1697	0.1673	0.1684	0.1658
Pop. den.	0.1660	0.1616	0.1600	0.1571	0.1512
NDVI	0.1070	0.1063	0.1050	0.1007	0.0988
Total	1.7277	1.6780	1.6530	1.6237	1.5632

Table A.7: Approach C, across all cities. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.

Learning rate	0.1	0.01	0.001	0.0001	0.00001
B. area	0.2149	0.2052	0.2181	0.1978	0.1964
B. count	0.2147	0.2032	0.2237	0.1967	0.1967
Avg. b. size	0.2406	0.2379	0.2369	0.2355	0.2319
B. orient.	0.2439	0.1926	0.1783	0.1235	0.1064
All roads	0.2431	0.2457	0.2459	0.2456	0.2450
M. roads	0.2489	0.2496	0.2513	0.2389	0.2337
Dist. all roads	0.1822	0.1957	0.1837	0.1802	0.1717
Dist. m. roads	0.2266	0.2192	0.2242	0.2268	0.2267
Pop. den.	0.2151	0.2081	0.2106	0.1966	0.2035
NDVI	0.1659	0.1707	0.1786	0.1539	0.1599
Total	2.1959	2.1279	2.1513	1.9955	1.9719

Table A.8: Approach C, Accra. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.

Learning rate	0.1	0.01	0.001	0.0001	0.00001
B. area	0.1375	0.1340	0.1419	0.1385	0.1327
B. count	0.1416	0.1416	0.1416	0.1361	0.1361
Avg. b. size	0.2267	0.2329	0.2319	0.2236	0.2240
B. orient.	0.2318	0.1847	0.1771	0.1585	0.1564
All roads	0.1962	0.1903	0.1913	0.1874	0.1831
M. roads	0.2201	0.2227	0.2272	0.2412	0.2414
Dist. all roads	0.1909	0.1783	0.1846	0.1807	0.1828
Dist. m. roads	0.1733	0.1653	0.1678	0.1624	0.1582
Pop. den.	0.1536	0.1474	0.1724	0.1771	0.1683
NDVI	0.1843	0.1492	0.1732	0.1741	0.1527
Total	1.8560	1.7464	1.8090	1.7796	1.7357

Table A.9: Approach C, Dakar. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.

Learning rate	0.1	0.01	0.001	0.0001	0.00001
B. area	0.1846	0.1702	0.1494	0.1507	0.1510
B. count	0.1927	0.1632	0.1577	0.1416	0.1416
Avg. b. size	0.2364	0.2368	0.2412	0.2337	0.2355
B. orient.	0.2129	0.1963	0.1860	0.2488	0.1783
All roads	0.2447	0.2465	0.2446	0.2463	0.2450
M. roads	0.2190	0.2276	0.2335	0.2528	0.2396
Dist. all roads	0.1958	0.2103	0.2076	0.2039	0.2004
Dist. m. roads	0.2209	0.2173	0.2173	0.2180	0.2144
Pop. den.	0.1691	0.1615	0.1505	0.1530	0.1518
NDVI	0.1619	0.1666	0.1486	0.1454	0.1491
Total	2.0380	1.9963	1.9364	1.9942	1.9067

Table A.10: Approach C, Dar es Salaam. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.

Learning rate	0.1	0.01	0.001	0.0001	0.00001
B. area	0.1845	0.1701	0.1612	0.1602	0.1426
B. count	0.2147	0.1912	0.1707	0.1707	0.1637
Avg. b. size	0.2129	0.2096	0.2125	0.2087	0.1874
B. orient.	0.2488	0.2455	0.2396	0.2470	0.2473
All roads	0.2336	0.2276	0.2365	0.2408	0.2349
M. roads	0.2427	0.2389	0.2416	0.2495	0.2491
Dist. all roads	0.2071	0.1980	0.1887	0.1914	0.1833
Dist. m. roads	0.2357	0.2321	0.2257	0.2235	0.2154
Pop. den.	0.2008	0.2000	0.1902	0.1881	0.1730
NDVI	0.2067	0.1923	0.1922	0.1846	0.1687
Total	2.1875	2.1053	2.0589	2.0645	1.9654

Table A.11: Approach C, Kigali. Comparison of learning rates (LR). Bold text indicates the best scores across urban characteristics.

A.5.4 Cluster names, Approach A with K=16.

Cluster number	Accra	Dakar	Dar es Salaam	Kigali
0		Mixed environments	Mixed environments	Mixed environments
1	Water and bare land	Water and bare land		Water and bare land
2		Dark vegetation	Dark vegetation	Dark vegetation
3		Shrubs and agriculture	Shrubs and agriculture	Shrubs and agriculture
4	Densely populated			
5		Bare land and shrubs	Bare land and shrubs	Bare land and shrubs
6	Dark vegetation			
7	Mixed environments			
8	Mixed environments			
9		Densely populated	Densely populated	Densely populated
10		Agriculture and bushes	Agriculture and bushes	Agriculture and bushes
11		Med. vegetation	Med. vegetation	Med. vegetation
12	Mixed environments	Mixed environments	Mixed environments	Mixed environments
13		Agriculture	Agriculture	Agriculture
14	Vegetation			
15	Vegetation			

A.6 XGBoost hyperparameters, combined approach

Hyperparameter	Value
Percentage of features used for building each tree	0.7503753049475365
Min split loss (gamma)	1.5913323626041533
Max depth of tree	14
Minimum sum of instance weight (hessian) needed in a child	6
L1 regularisation term	70
L2 regularisation term	0.4037776702619792

Table A.12: Best hyperparameter for XGBoost model predicting cluster membership of four cities.

A.7 ESA land cover co-occurrence

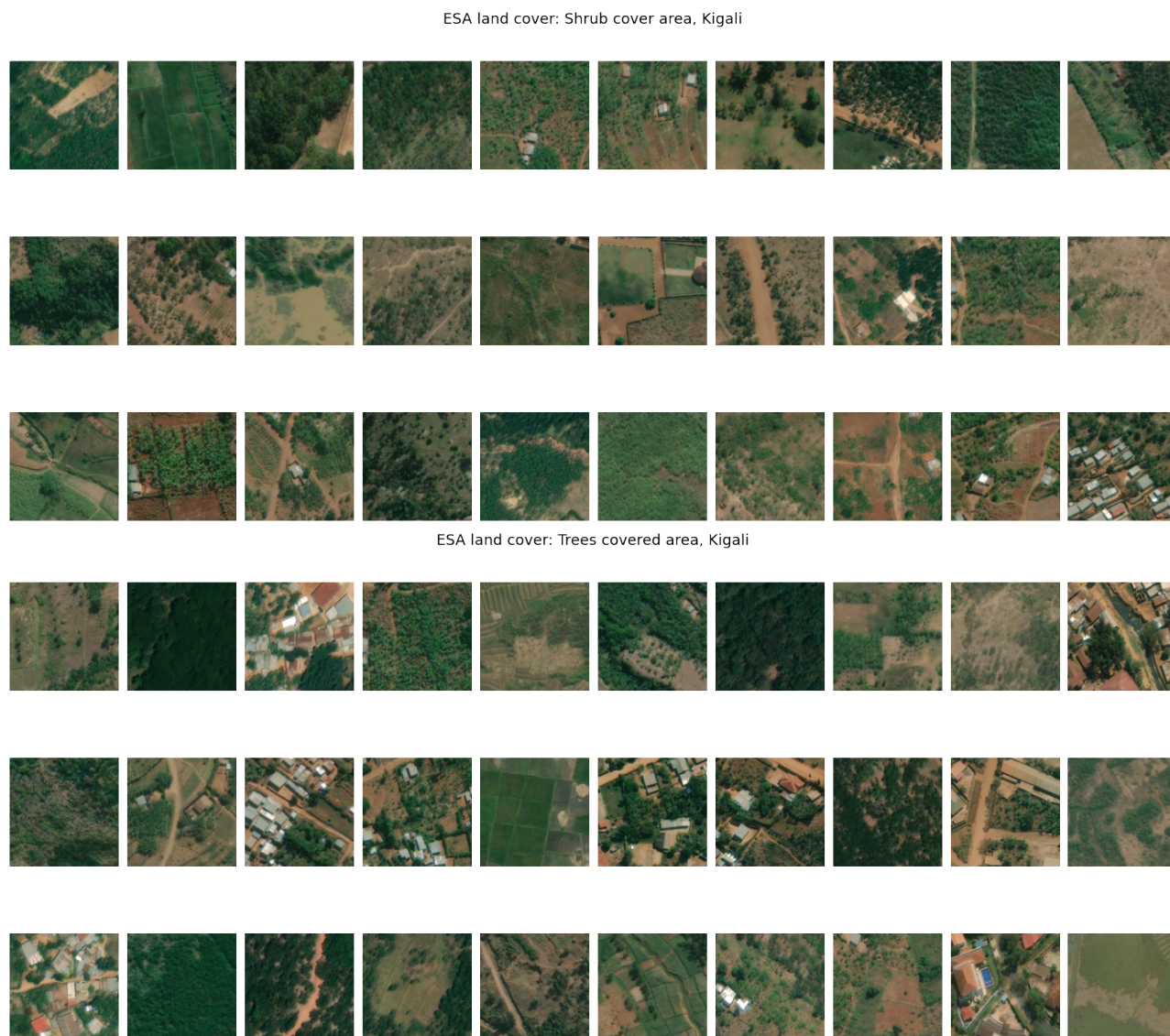


Figure A.2: Tiles of two different ESA land cover classes of Kigali that are visually very similar. Dynamic mosaic ©2020 Maxar Technologies.



Figure A.3: Tiles of two different clusters of Kigali. Maxar Products. Dynamic mosaic
©2020 Maxar Technologies



Figure A.4: Tiles labelled as *Dark dense vegetation* cluster, Accra.

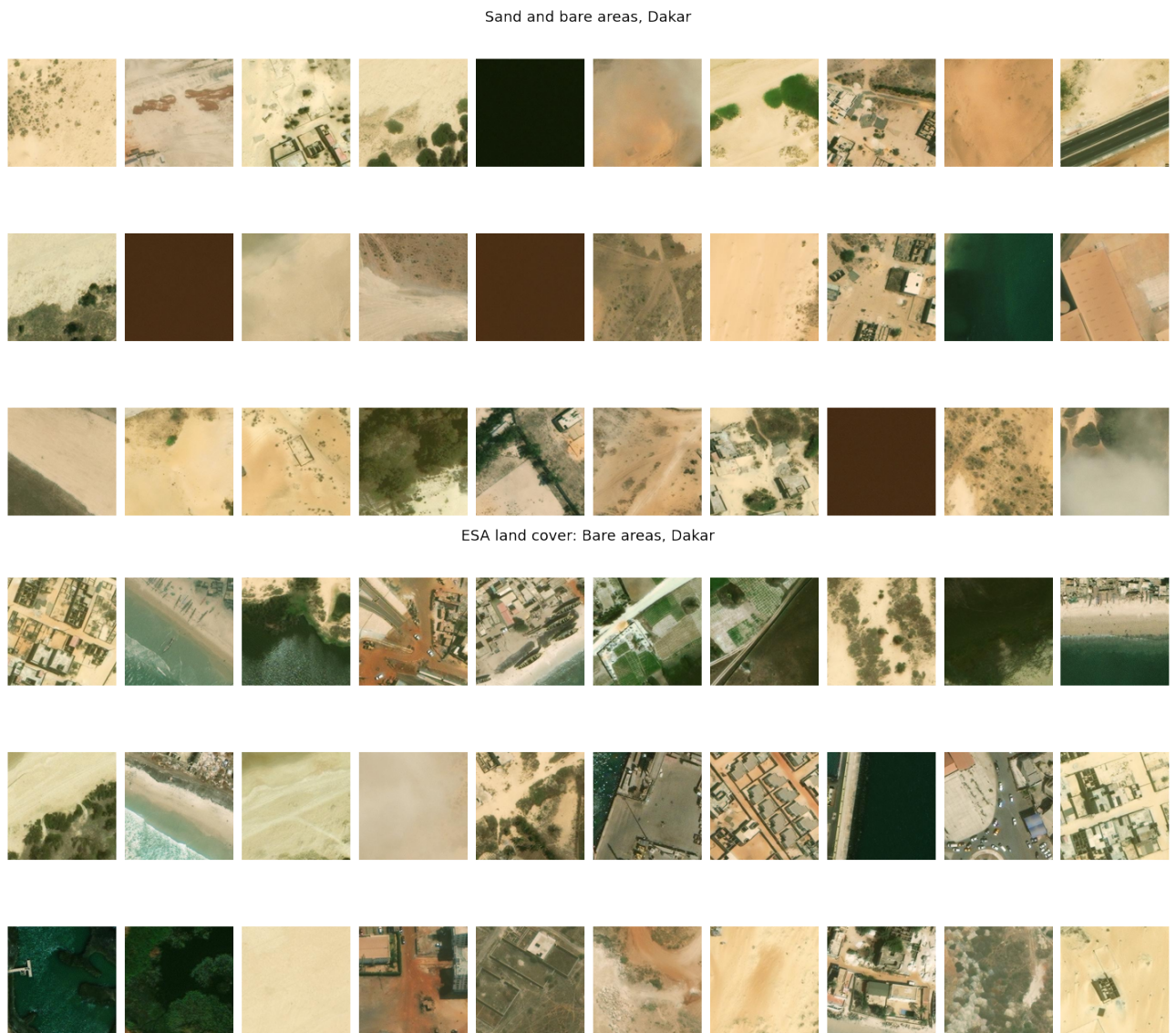


Figure A.5: Tiles labelled as *Sand/bare areas and water* cluster and as *Bare areas* by the ESA land cover dataset, Dakar. Dynamic mosaic ©2021/2 Maxar Technologies.

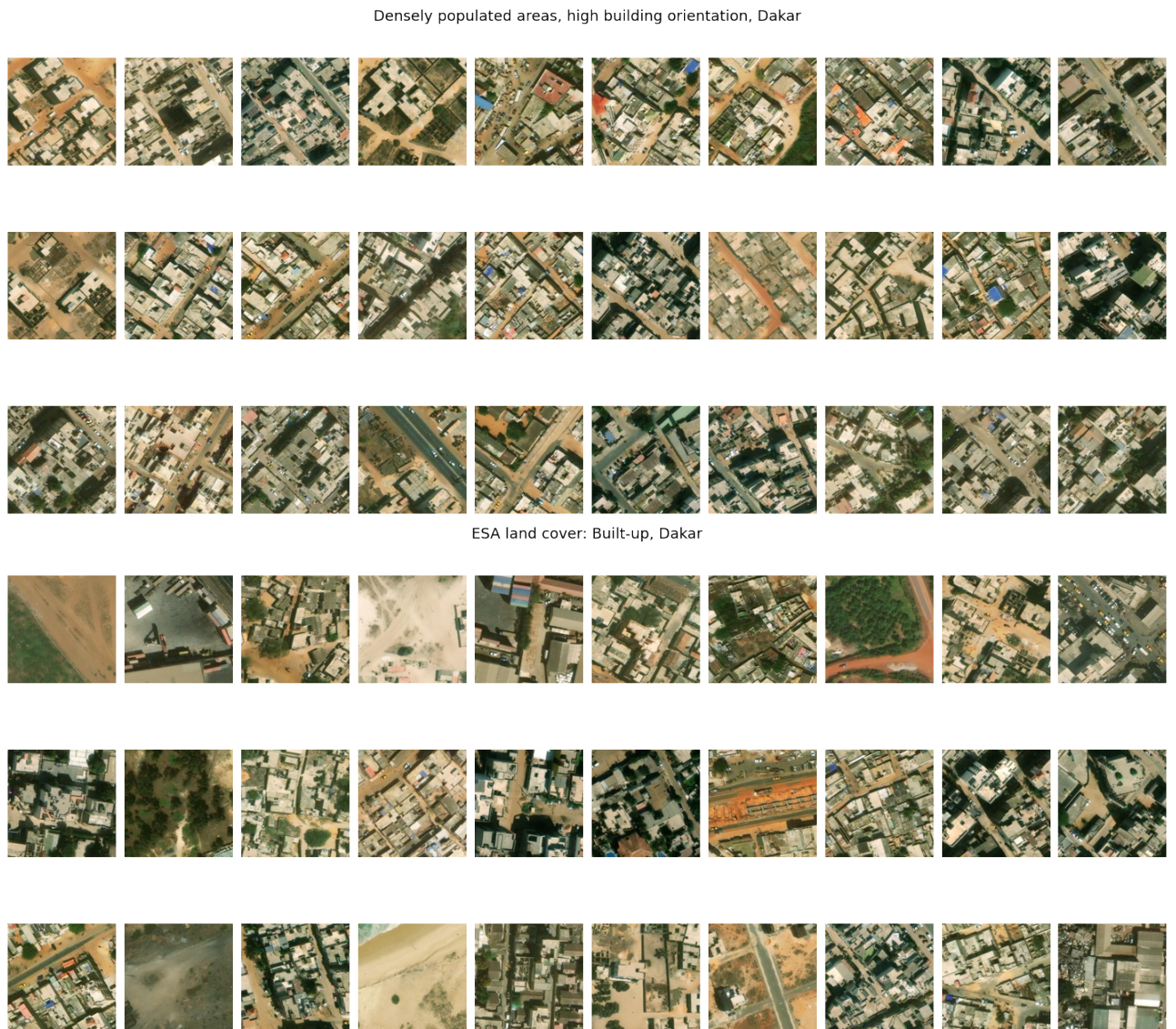


Figure A.6: Tiles labelled as *Densely populated areas with high building orientation* cluster and as *Built-up* by the ESA land cover dataset, Dakar. Dynamic mosaic ©2021/2 Maxar Technologies.

A.8 Sensitivity to cluster number, multi-city analysis

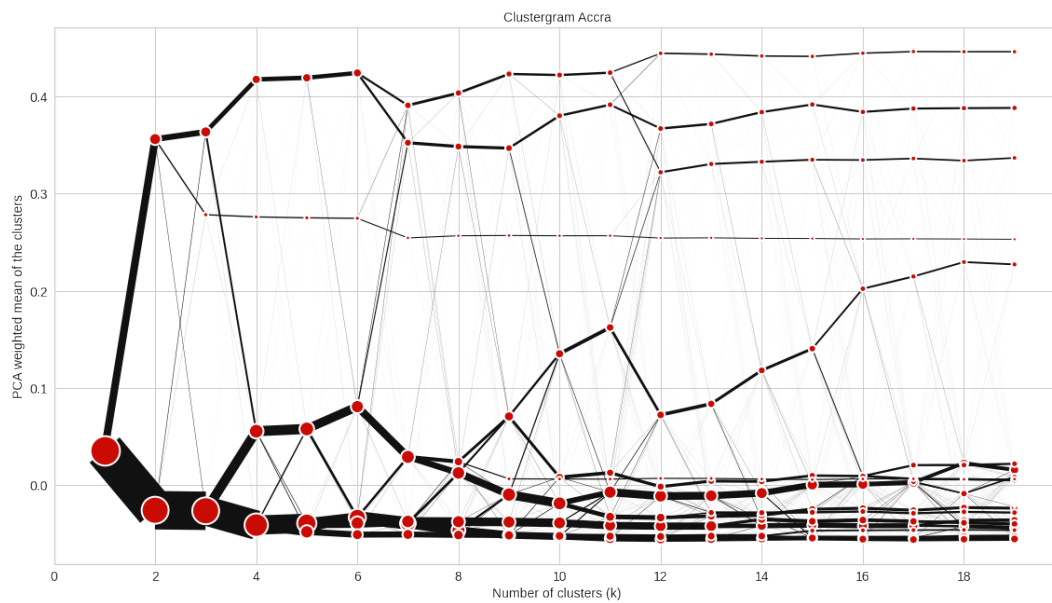


Figure A.7: Clustergram of image representations of Accra.

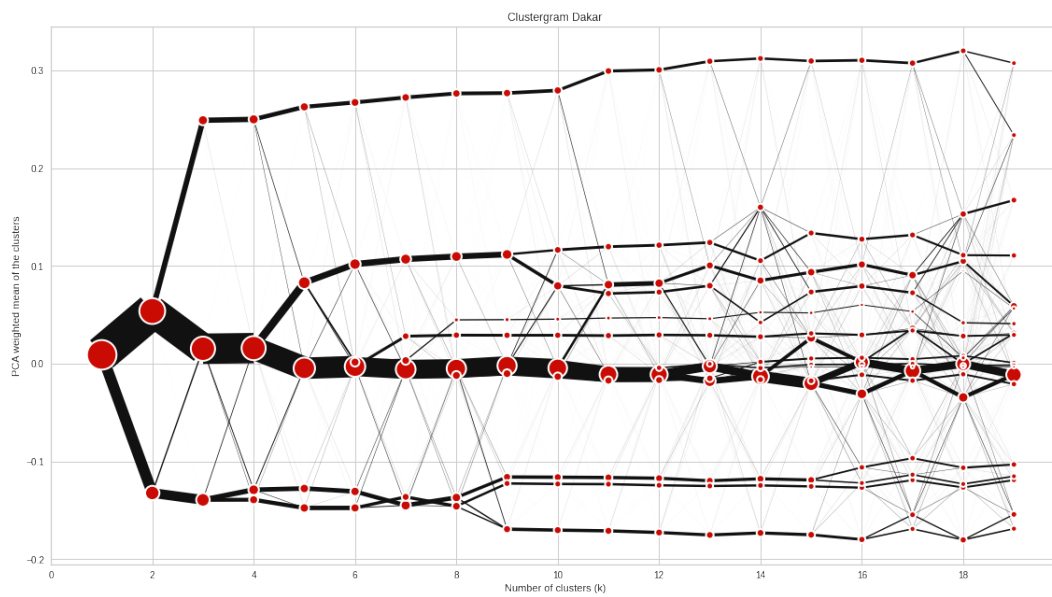


Figure A.8: Clustergram of image representations of Dakar.

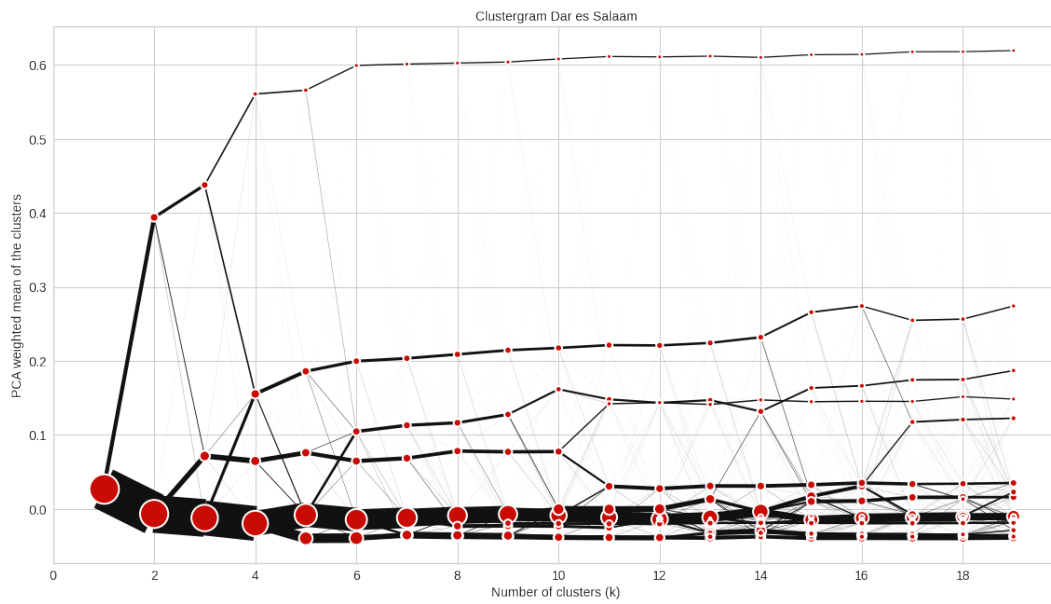


Figure A.9: Clustergram of image representations of Dar es Salaam.

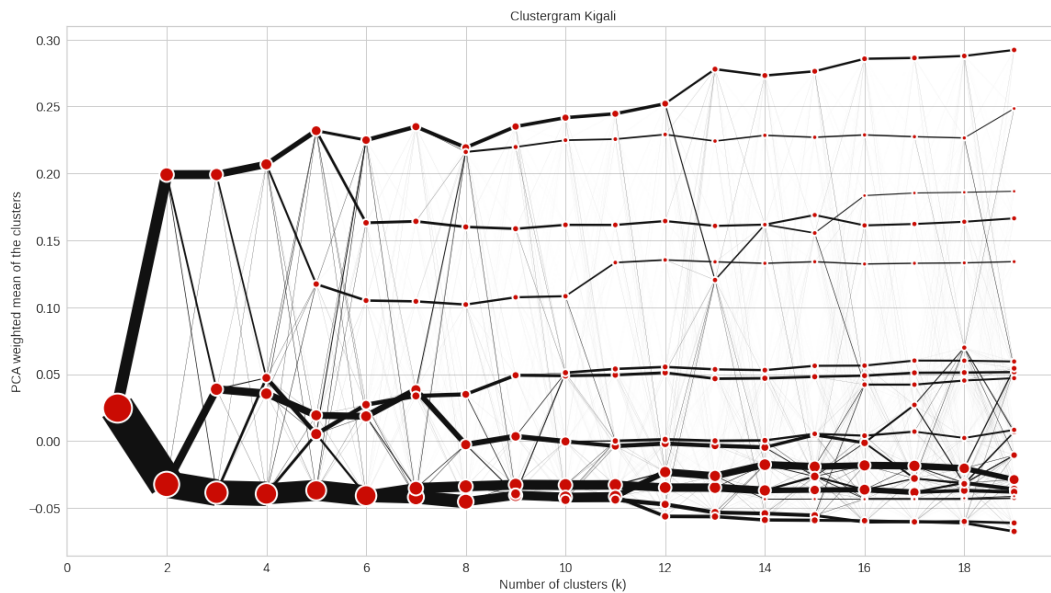


Figure A.10: Clustergram of image representations of Kigali.

A.9 Inter- and intra-cluster distances

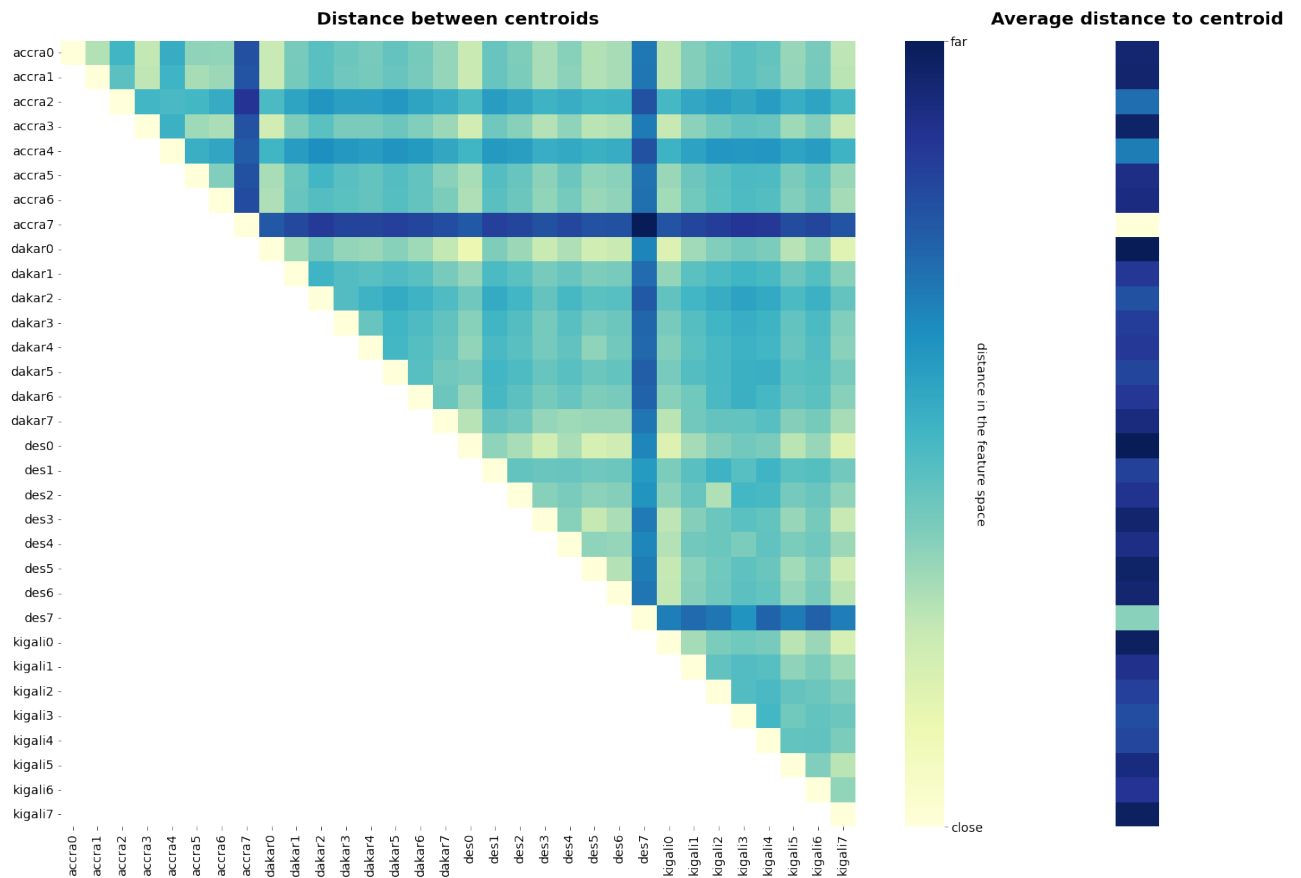


Figure A.11: The figure shows the intra- and inter-cluster similarity in the feature space. The intra-cluster similarity is calculated as average distance of tiles to cluster centroid. The distance between centroids is a measure of similarity across clusters. The smaller the distance between the clusters, the more similar they are, and vice versa. For computational efficiency, within- and between-cluster distances were calculated using the first two PCs.