

AUTOMATIC DELIRIUM CLASSIFICATION IN
INTENSIVE CARE USING NON-INVASIVE EYE
TRACKING

DR AHMED AL-HINDAWI

Thesis submitted for the degree of Doctor of Philosophy

Supervised by

PROFESSOR YIANNIS DEMIRIS

&

DR MARCELA VIZCAYCHIPI

Imperial College London

June 2022

COPYRIGHT DECLARATION

The copyright of this thesis rests with the author and is made available under a **Creative Commons Attribution Non-Commercial No Derivatives** licence. Researchers are free to copy, distribute or transmit the thesis on the condition that they attribute it, that they do not use it for commercial purposes and that they do not alter, transform or build upon it. For any reuse or redistribution, researchers must make clear to others the licence terms of this work.

ORIGINALITY DECLARATION

I hereby declare that this thesis and the work herein detailed, was composed and originated by myself, except where appropriately referenced and credited.

London, June 2022

Dr Ahmed Al-Hindawi

ABSTRACT

Delirium, an acute confusional state, is unfortunately common in Adult Intensive Care Units (AICUs). Its development heralds increased risk of morbidity, mortality, increased length of hospital stay, and worse cognitive outcomes following discharge with cognitive scores not too dissimilar to Alzheimer's dementia. In this thesis, we explore the use of machine learning techniques for the classification of delirium from Electronic Patient Records (EPR) data routinely collected as part of the patient's clinical care. We hypothesise that delirium can be classified by eye-movements as a measurable biological signal of *overt* attention as a surrogate marker for *covert* cognitive attention.

We thus developed, and validated, an eye tracking platform that exceeds empiric requirements of invasiveness, performance, accuracy, and precision. Following appropriate approvals, we then deployed the platform across two hospitals in London. The data gathered facilitated the training of machine learning algorithms that can classify delirium based on eye-movements and scene-information with an Area Under Receiver Operator Curve (AUROC) of 0.69 and 0.78 respectively, and an Area Under the Precision Recall Curve (AUPRC) of 0.8 and 0.83 respectively. These classifiers also provided novel insights into the visual attention of delirium patients.

To explain these classifier's findings, we further develop a novel, theory of mind inspired, architecture composed of multiple hierarchical competing forward models. Using this architecture, we demonstrate that patients with delirium have globally reduced levels of visual attention throughout the visual processing hierarchy to a statistically significant degree ($p = 0.044$).

ACKNOWLEDGEMENTS

I cannot start an acknowledgements section without firstly thanking both of my supervisors. I am grateful to Professor Yiannis Demiris and Dr Marcela Vizcaychipi. Without Marcela, I would not have had the opportunity to pursue what amounted to 6 years worth of work – thank you for believing in me, for giving me a job, and for lighting a fire in me for Intensive Care Medicine. I will always remember my time at Chelsea and Westminster Hospital fondly.

I am eternally grateful to Yiannis – his warmth, patience, and contagious enthusiasm facilitated scientifically fruitful discussions that are unforgettable. He gave me the space to think freely and pursue what is an unusual topic given his laboratory's focus. I have learnt a huge amount about scientific rigour, leadership, and humanity. I reminisce about the times I would knock on his door, discuss an issue, and leave an hour or two later (and on occasions, significantly more than that), having not only solved the problem, but also feeling that I had caught up with a friend. I hope to replicate a fraction of his qualities.

I would like to thank past and present members of the Personal Robotics Laboratory – specifically Dr Tobias Fischer, Dr Rodrigo Chacón Quesada, Dr Joshua Elsdon, Dr Fan Zhang, and Mr (hopefully soon to be Dr) Cedric Goubard. The discussions we have had have been a constant supply of enthusiasm and inspiration and for that I'll be forever grateful. I owe them all a significant amount of coffee (or energy drinks).

I would like to thank Dr Alice Sisson and Dr Michelle Hayes – your pastoral supervision was appreciated more than I often let on. Even during the COVID-19 pandemic, you had the emotional bandwidth to put up with me and I'm forever grateful. I don't think I've drank more tea in my life than

with you two. I would also like to thank Dr Luke Moore, Dr Roger Davies, and Professor Suveer Singh: you have all supported me significantly in ways that I cannot fully express.

I would like to extend a massive thank you to my friends for listening me endlessly rant about my PhD – Dr Pramuditha Fernando, Mr (he's a Surgeon) James McGhee, Dr Hitesh Patel, and Dr Khalid Zalmay who have kept me laughing until my cheeks hurt.

My deepest gratitude goes to my family, my father Dr Muhaned Al-Hindawi, my mother Mrs Khilod Ali, and my two sisters Ms Rand Al-Hindawi, and Ms Asma'a Al-Hindawi. They are the rock upon which I stand and have been my unwavering support in the various endeavours I found myself in. They are my constant source of inspiration, instilling a sense of hard-work and a desire to do better. Whenever I was stressed, I would cycle to my parent's house and be lavished with love, support, delicious food, and endless amounts of chocolate.

Lastly, my eternal gratefulness and love goes to my dear wife, Dr Samantha Baillie. You have been a constant source of love, affection, and happiness. Thank for choosing me to spend your life with, I can't wait to see what the future holds.

CONTENTS

1	Introduction	1
1.1	Research Objectives	2
1.1.1	Time-Series Analysis	3
1.1.2	Biological Signal Processing	4
1.2	Contributions	4
1.2.1	Delirium Classification from Electronic Healthcare Records Data	5
1.2.2	Development and Validation of Eye Tracking Platform .	5
1.2.3	Feasibility Study	6
1.2.4	Delirium Classification from Eye Movements	6
1.2.5	Competing Forward Models for Visual Attention Un- derstanding in Delirium	6
1.3	Thesis Outline	7
1.4	Appendices Outline	8
2	Background	11
2.1	Operational Diagnosis of Delirium	12
2.2	Prediction of Delirium	13
2.3	Outcomes of Patients who Develop Delirium	14
2.4	Eye-Movements in Disease	17
2.4.1	Eye-Movements in Dementia	18
2.4.2	Eye-Movements in Psychiatric Diseases	20
2.5	Time-series classification	20
2.5.1	Long Short-Term Memory Modules	21
2.5.2	Temporal Convolutional Neural Networks	22
2.6	Performance of Classifiers	23

2.7	Handling Missing Data	24
3	Delirium Classification using Routinely Collected Healthcare Data	27
3.1	Methodology	28
3.1.1	Data Collection and Transformation	29
3.1.2	Analysis and Model	32
3.2	Results	33
3.3	Discussion	36
3.4	Conclusion	38
4	An Eye-Tracking Platform for Patients in Intensive Care	41
4.1	Empiric System Requirements	41
4.2	Design and Performance Analysis	42
4.3	Measurements of Accuracy and Precision	47
4.4	Results	49
5	Improving Blinking Classification by Curriculum Learning by Augmentation	55
5.1	Related Works	57
5.2	Method	58
5.2.1	Blink Datasets	58
5.2.2	Curriculum by Augmentation	59
5.2.3	Pacing	61
5.2.4	Neural Network Architecture and Training	63
5.2.5	Assessment of Model Performance	63
5.3	Results	64
5.3.1	Ablation Studies	69
5.4	Conclusion	69
6	What is the Patient looking at? Robust Gaze-Scene Intersection	73
6.1	Method	78
6.1.1	Planar Viewpoint Generator	79
6.1.2	Intersection	81

6.2	Dataset	82
6.3	Results	84
6.4	Ablation Studies	85
6.5	Conclusion	85
7	Feasibility Study Details and Data Processing	87
7.1	Study Details	87
7.1.1	Ethics Approval and Consent to Participate	88
7.1.2	Inclusion Criteria	89
7.1.3	Exclusion Criteria	89
7.1.4	Sample Size	90
7.2	Data Processing	93
7.3	Fixation Classification	95
7.4	Conclusion	96
8	Delirium Classification using Eye Tracking Data	97
8.1	Delirium Classification from Eye Movements Model	99
8.2	Delirium Classification from Point of Regard Model	99
8.3	Training regime	99
8.4	Threshold Selection	100
8.5	Results	101
8.5.1	Eye movement differ between patients with and without delirium	101
8.5.2	Point of Regard differs between patients with and with- out delirium	104
8.6	Conclusion	105
9	Competing Forward Models as a Solution to the Inverse Yarbus Problem	107
9.1	The Inverse-Yarbus Problem	107
9.2	Overview of Architecture	110
9.2.1	Saliency Models	111

9.2.2	Saccade Biases	111
9.2.3	Inhibition of Return	111
9.2.4	Prediction Error	113
9.2.5	Competition	114
9.3	Validation	114
9.3.1	Experiments in Simulation	114
9.3.2	Visual Search Dataset	116
9.3.3	Saliency Models for Delirium Understanding	118
9.4	Forward Models Explain Underlying Attention Mechanisms in Delirium	119
9.5	Conclusion	125
10	Conclusions and Future Work	127
10.1	Overview and contributions of the thesis	127
10.2	Limitations	128
10.2.1	Hypothesis	129
10.2.2	Longitudinal Healthcare Data	129
10.2.3	Eye Tracking Platform	129
10.2.4	Delirium Classification using Temporal Convolutional Networks	130
10.2.5	Delirium Classification using Competing Forward Models	131
10.3	Future Directions	131
10.3.1	Prospective Evaluation	131
10.3.2	Neuro-computational models	132
10.4	Epilogue	132
11	Author's Publications	133
11.1	Pending	134
A	Scanpath Metrics	137
A.1	ScanMatch Algorithm	137
A.2	MultiMatch	138

A.3	Dynamic Time Warping	139
A.4	Combining DTW and MultiMatch	140
A.4.1	Validation of MultiMatch DTW	141
B	Trial Documents	145
B.1	Protocol	145
B.2	Declaration Forms	166
B.3	Consent Forms	171
B.4	Information Sheets	176
B.5	Patient Information Sheets	185
	Bibliography	195

LIST OF FIGURES

Figure 1.1	Outline of the thesis set to the ‘Confused Lady’ meme demonstrating how eye-movements as a <i>overt</i> signal of <i>covert</i> internal cognitive mechanisms has permeated popular culture.	7
Figure 2.1	Calendar plot of admission age of patients admitted to AICU over time starting from 2011 until 2018 demonstrating stable age of admission in the dataset. Grey indicates lack of new admissions and thus no data, Green indicates low average age and Red indicates high average age.	16
Figure 2.2	Calendar plot of the severity of disease as modelled by Acute Physiology and Chronic Health Evaluation II (APACHE-II) over the dataset’s duration of 2011 to 2018 demonstrating stable severity of disease in the dataset. Grey indicates lack of new admissions and thus no data, Green indicates low average acuity and Red is high average acuity.	17
Figure 2.3	Calendar plot of the Early PREdiction of DELIRium in ICu patients (E-PRE-DELIRIC) score over the dataset’s duration of 2011 to 2018 demonstrating that risk of delirium is increasing over the dataset. Grey indicates lack of new admissions and thus no data, Green indicates low risk of delirium and Red is high risk of delirium.	18

Figure 2.4	Violin plots of patient's outcomes and E-PRE-DELIRIC score demonstrating that patients who have died have a much higher E-PRE-DELIRIC score compared to those who were discharged alive.	19
Figure 3.1	CONSORT diagram of patient flow in the study looking at the use of Electronic Health Record (EHR) data for the classification of delirium. Patients in the first wave were used for model building and validation whereas patients in the second wave were used for testing.	30
Figure 3.2	Receiver Operator Characteristic (ROC) curves and Precision-Recall Curve (PRC) for classification of delirium from time-series data. The proposed model demonstrates an AUROC of 0.77 with tight confidence. The mean Average Precision (mAP) in the PRC is 0.80 demonstrates an adequate trade-off between sensitivity and positive predictive value. In the ROC curve, the lower triangle demonstrates an area where classification is worse than chance (flip of a coin) and the upper triangle demonstrates better than chance classification. . . .	34

Figure 3.3	The architecture of the neural network used for the analysis of EHR data to classify delirium presented pictorially. The inputs (X at time t) consist of the 17 variables outlined in Table 3.2 . Each layer's output then forms the input into the next layer. The first module, a causal Long-Short Term Memory (LSTM) extracts temporal features into a fixed vector. The Feature Embedding Module then embeds these features in higher dimensional space followed by a final linear transformation layer to output the probabilities that this time series corresponds to a Delirious or Non-Delirious patient.	35
Figure 3.4	The effect of time, considered as a fraction of the total data available, on the accuracy of the model at classifying delirium from EHR data. As more data is available, the model's accuracy monotonically increases from an AUROC of 0.67 to 0.77 and a mAP of 0.68 to 0.80. Gray lines are individual models with the Black line indicating the average accuracy gain. Blue arrows demonstrate the accuracy uplift.	36
Figure 4.1	An overview of the proposed system for eye tracking in AICU . Two cameras are used: a patient facing camera (head-camera, blue) and a patient-perspective facing depth camera (scene-camera, green) connected to a laptop (outside the viewpoint) running Robot Operating System (ROS). The flow of information through the system is demonstrated with the head-camera and scene-camera running concurrently.	43

Figure 4.2	Aggregated results for all participants demonstrating the accuracy and precision per marker where X and Y axes represent the ArUco marker location on screen in Pixels. The sensitivity of the camera's position relative to the viewer is evident where the highest accuracy and precision is when the user's head pose is directed at the camera.	50
Figure 4.3	Results of the empiric requirements of accuracy and precision compared to achieved accuracy and precision of our proposed eye tracker in AICU. Density plots demonstrating the required accuracy and precision from empirical measurements and the system's metrics as evaluated in the replicated environment in the laboratory.	51
Figure 4.4	Posterior density checks for our proposed eye tracking platform in AICU demonstrating similar ranges of data when compared to the training datasets. Density plots of aggregated data from 5 patients stratified by delirium status. Eye and Head refer to Eye and Head pose angles from the system while subscripts y , p and r refer to yaw, pitch and roll respectively. Blink is classified as per Real-Time Blink Estimation in Natural Environments (RT-BENE). Blink and Gaze statistics are of similar distributions between RT-GENE/BENE and that of non-delirious patients. Of note, no $Head_r$ data is present in Real-Time Eye Gaze Estimation in Natural Environments (RT-GENE)-RT-BENE datasets as faces processed to remove roll prior to inference.	53

Figure 5.1	The role of the δ as the strength operator for our proposed Curriculum Learning by Augmentation (CLbA) training regime. Original image patch ($\delta = 0$) and a sample of the possible data augmentations with increasing strength (and thus difficulty) of augmentation ($\delta = 0.5$ and $\delta = 1.0$ respectively).	61
Figure 5.2	Three possible pacing functions that increase the difficulty (δ) of the data-points as training progresses in a curriculum fashion as part of our proposed CLbA training regime.	62
Figure 5.3	The performance of the blinking classifier trained using our proposed CLbA training scheme. An ensemble of ResNet-18 models achieves state-of-the-art mAP of 0.971 compared to Curriculum Learning by Bootstrapping (CLbB) of 0.122. Ensemble CLbA using a trigonometric pacing function and one that uses an exponential pacing function are completely overlapping indicating identical performance.	66
Figure 5.4	The performance of the blinking classifier trained using our CLbA training regime. The AUROC of the CLbA approach is near maximal whilst the CLbB scheme performs poorly. Ensemble CLbA using a trigonometric pacing function and one that uses an exponential pacing function are completely overlapping indicating identical performance.	67

Figure 5.5	Normalised model calibration curve for blinking classification models trained using CLbA against baseline models. Models trained with CLbA produce the best calibrated models whilst CLbB produces poor calibration. Ensemble CLbA using a trigonometric pacing function and one that uses an exponential pacing function are completely overlapping indicating identical performance.	70
Figure 5.6	Probability thresholds versus F1-score of each blinking classification model. Models trained using our proposed CLbA approach produce appropriately high thresholds maximising the F1-Score whilst baseline models produce lower F1-Scores across all thresholds.	71
Figure 5.7	Example outputs of the model with respective probabilities in two subjects demonstrating an eyes open and blinking scenarios are classified with high discriminatory power.	71
Figure 6.1	A pictorial result of our proposed Point of Regard (PoR) estimation method. An in-painted and rectified depth image with head position and gaze direction indicated as the blue circle and arrow respectively. The red cross is the target label with the detected gaze-scene intersection point labelled as a yellow triangle (RGB image blurred to preserve privacy) with near 100% accuracy.	76

- Figure 6.2 The proposed architecture for PoR regression. Scene information including depth image (RGB image blurred) is provided where the Head Pose H and Gaze vector G are estimated. A planar cone is then rasterised starting at H with direction G with a central angle of θ and subsamples J . This rasterised image intersects the depth image. Gaze vector $\overrightarrow{HG}$ then intersects the set where candidate pixels are then classified whether they are inside the cone of gaze. The successful candidates are sorted and the intersection point selected. 77
- Figure 6.3 Our PoR estimation algorithm’s performance compared to Dual Attention Module with increasing noise and thus decreasing precision of the eye-tracker demonstrating that our proposed method surpasses the previous method at all noise levels. Fractions below each plot represent the failures in each technique as a fraction of the overall number of runs. Notches in the box plot that do not overlap indicate statistical significance. Plots are split according to reduction strategy where the Sampling strategy N was set to 5. 80
- Figure 6.4 The effect of depth in-painting on the performance of the proposed PoR estimation algorithm. Depth in-painting has negligible effect on accuracy but does reduce the number of failures. Noise is set to 0 in this experiment. 82

Figure 6.5	Ablation study for optimal number of closest samples for PoR estimation. The number of potential candidates to sample from set C_l relating to accuracy. The naive strategy of picking the voxel closest to the the ray has the highest accuracy under experimental conditions whereas the sample strategy ($N > 1$) decreases accuracy. 83
Figure 6.6	Ablation study for optimal number of samples of the planar-viewpoint generator for PoR estimation. The number of samples J tested against $\overrightarrow{HG}$ against normalised accuracy. As J increases, the number of successful queries increases. 84
Figure 7.1	Flow of patients through our pilot feasibility study titled Continuous Non-Invasive Eye Tracking for the Diagnosis of Delirium (CONfuSED). Upon arrival, patients are checked for eligibility and if they meet the criteria, appropriate consent/advice is sought from the patient if capacitous or from their next of kin or a nominated consultee if not capacitous. Upon inclusion into the study, at least daily measurement of delirium through Confusion Assessment Method in ICU (CAM-ICU) is performed alongside measurement of the patient's gaze as per Figure 7.2 88
Figure 7.2	An outline of the eye tracking system used in CONfuSED alongside the origin of the data used in the each experiment. 92
Figure 7.3	Data preparation performed in an offline stage following data acquisition. Raw un-ordered sequential messages are ordered into ordered temporal storage facilitating analysis. 94

Figure 7.4	Two scenarios presented illustrating fixation classification. Consecutive gaze angles are placed in a set and the maximum angle between successive gaze vectors is taken as the dispersion. Areas of low dispersion are deemed to be fixations.	96
Figure 8.1	The architecture for delirium classification from our eye tracking platform. Time-series inputs are encoded to a fixed-sized vector using a Temporal Convolutional Network (TCN) where a set of fully connected layers are used to convert the encoded vector to the probability of delirium. X_t represents the input at time point t . h_t represents the hidden latent fixed vector size that encodes previous time points. N and ℓ are the number of layers of the TCN and the fully connected layers respectively.	98
Figure 8.2	Receiver Operating Characteristic Curve as a performance indicator of the TCN for the binary classification task of diagnosing delirium. Two models are presented, an eye-movements only model, and a PoR model. The eye-movements model classifies delirium based on time-series eye movement data, namely the horizontal and vertical eye angles of each fixation with an AUROC of 0.69. Whereas the scene model classifies delirium from time-series PoR data to an AUROC of 0.78. Both classification models use the architecture in Figure 8.1.	101

- Figure 8.3 Precision-Recall Curve as a performance indicator of the TCN for the binary classification task of diagnosing delirium. Two models are presented, an eye-movements only model and a PoR model. The eye-movements model classifies delirium based on time-series eye movement data achieving an AUPRC of 0.80. Whereas the scene model, that contains time-series PoR data, achieves an AUPRC of 0.83. Both classification models use the architecture in Figure 8.1. 102
- Figure 8.4 The proposed classifier's recall split by patient's E-PRE-DELIRIC score. Both the eye movements model and the scene model demonstrate good re-call throughout the range of potential risk as calculated by E-PRE-DELIRIC. Probability threshold for recall calculation was set to maximise the F1-Score. 102
- Figure 8.5 The proposed classifier's precision split by patient's E-PRE-DELIRIC score. Both the eye movements model and the scene model demonstrate adequate precision throughout the range of potential risk as calculated by E-PRE-DELIRIC. Probability threshold for precision calculation was set to maximise the F1-Score. 103
- Figure 8.6 Normalised Calibration Plots of the mean predicted risk against the observed fractions for the two models tasked with delirium classification by eye movements or PoR time-series data. Numbers in brackets represent the Brier calibration score. 103

Figure 8.7	Threshold sweep against the F1 score at that threshold where the threshold at the maximum F1 Score is the optimal point to dichotomise the probabilistic output of the classification models. Numbers in brackets represents the F1 score and the thresholds at that F1-Score. 104
Figure 9.1	Eye-movement specific modules used in the forward models as a solution to the inverse Yarbus problem. Saccade Bias modelled in Figure 9.1a and Inhibition of Return in Figure 9.1b . Saccade Biases are modelled as a distribution with a baseline probability of 0.0 that extends to 1.0 within the standard deviation set by the field of view. Inhibition of Return is modelled as a cumulative distribution where the baseline inhibition is 1.0 indicating that no inhibition takes place following multiplication. At each fixation, a Gaussian is placed that inhibits return to that location where its probability decays back to 0.0. Three fixations are demonstrated in various states of decay. 112
Figure 9.2	Illustration of the scene of the simulated environment under for validation of the architecture using synthetic data. Two experiments were performed, the first is a <i>constant</i> task where fixations originated from the red circle, and a <i>switch</i> task where fixation started in the red circle and then switched to the blue. 115

Figure 9.3	Simulation results for the architecture as a solution to the inverse Yarbus problem. Two eye movement strategies are presented, a fixed strategy, and a switch strategy. Various β parameters are explored aiming to maximise the range of the softmax operator in Equation 9.4. Small values of β result in encroachment of the two simulations' probabilities while larger values maximise classification distance but lose sensitivity to small changes. 116
Figure 9.4	A box-plot of the last fixation scores across all subjects across all images. Outliers presented with an open circle. Last fixation of the Target is correctly classified across all images and all subjects leading to the collapse of the box-plot into a single line. 118
Figure 9.5	Forward model probabilities for the subjects viewing the same image. All fixations at the end of the search land on the target where the forward model for the target correctly classifies the viewing patterns. 119
Figure 9.6	Original scene image and corresponding saliency maps as calculated by Itti, SRA, and Object detection respectively. Note that the scene images are blurred for privacy but saliency calculations occur at full resolution without blurring. Note the finer saliency details of Itti's saliency model compared to SRA and Object respectively. SRA saliency produces proto-images, shapes that can be objects or part of objects whilst Object saliency produces only objects. 120

Figure 9.7	A schematic of the proposed architecture as a solution to the inverse Yabus problem. Multiple forward models are stacked, each containing a saliency, saccade bias, and inhibition of return modules which are parametrised by the current fixation at time t . Each forward model's correctness is then measured at the next fixation at time $t + 1$. The prediction error is then normalised using a modified soft-max as per Equation (9.4). Following normalisation, the forward model with the highest probability is deemed the winner and allocated a point.	122
Figure A.1	Multimatch Algorithm where Subfigure A.1a is the scan paths to compare. Figure A.1b is the comparison matrix according to a predefined metric. Subfigure A.1c is the adjacency matrix calculated from the comparator matrix in Figure A.1c. A graph is then constructed from top left to bottom right and Dijkstra's algorithm is used to find the shortest path as per Figure A.1d (Jarodzka, Holmqvist, and Nyström, 2010).	141
Figure A.2	Example of an interpolation between two scan-paths, $sp1$ (blue) and $sp2$ (orange). $sp3$ (green) is being interpolated with the corresponding α values from $sp1$ into $sp2$	142
Figure A.3	Validation result of combining MultiMatch with DTW. As $sp1$ is interpolated to $sp2$ as $sp3$ using the interpolation constant α the classification accuracy stays relatively accurate and only decreases steeply once the $sp3$ is ambiguous. ScanMatch fails linearly in comparison to MultiMatch.	143

LIST OF TABLES

Table 3.1	Patient Characteristics split across training and testing cohorts for the classification of delirium using routinely collected healthcare data. Testing cohort (second wave) had reduced severity of disease as demonstrated by a lower mean APACHE-II score and lower Rockwood Frailty score demonstrating the changing admission criteria landscape between the first wave and second.	31
Table 3.2	Parameters used in the time-series model for the classification of delirium from routinely collected healthcare data. The mean and standard deviations of each variable are used to centre continuous variables during training. At testing, the means and standard deviations are used to centre the values prior to inference.	39
Table 4.1	Demographics data of participants in the validation experiment and the preliminary patients in CONFUSED . .	48
Table 4.2	The questionnaire used assess the acceptability of the use of cameras on day to day activities of the nursing staff on AICU	49

Table 5.1	Performance of our proposed deep neural networks for blink detection and the previous state-of-the-art. ResNet-18 provides the fastest inference time. Previous work (namely Cortacero, Fischer, and Demiris, 2019) does not meet Nyquist criteria and is thus not suitable for real-time inference. Accuracy of ResNet-50 and VGG-16 are not presented in Figures 5.3 to 5.5 to avoid clutter.	68
Table 7.1	Patient Characteristic enrolled in the feasibility trial titled CONfuSED. Pertinent variables related to outcomes of interests as well as variables that predict the development of delirium are stratified by data collection centre.	91
Table 9.1	The tallied fixations for each forward model stratified by delirium and non-delirium episodes. Statistical significance testing using the χ^2 test between the tallied probabilities of delirium vs non-delirium with the null hypothesis being of no difference.	124
Table A.1	Distance metrics used in the evaluation of a scan path.	139

ACRONYMS

3DDFA	Face Alignment in Full Pose Range: A 3D Total Solution
AICU	Adult Intensive Care Unit
APACHE-II	Acute Physiology and Chronic Health Evaluation II
AUPRC	Area Under the Precision Recall Curve
AUROC	Area Under Receiver Operator Curve
CAM-ICU	Confusion Assessment Method in ICU
CLbA	Curriculum Learning by Augmentation
CLbB	Curriculum Learning by Bootstrapping
CLbT	Curriculum Learning by Transfer
CNN	Convolutional Neural Network
CONfused	Continuous Non-Invasive Eye Tracking for the Diagnosis of Delirium
CT	Computed Tomography
CWH	Chelsea & Westminster Hospital
DAM	Dual Attention Guided Gaze Target Detection in the Wild
DSM	Diagnostic and Statistical Manual of Mental Disorders
DTW	Dynamic Time Warping
E-PRE-DELIRIC	Early PREdiction of DELIRium in ICu patients
EEG	Electroencephalogram
EHR	Electronic Health Record
EPR	Electronic Patient Records

FOV	Field of View
HAMMER	Hierarchial Attentive Multiple Models for Execution and Recognition
HDF5	High-performance data management and storage suite
HMM	Hidden Markov Model
HMOSAIC	Hierarchical Modular Selection and Identification for Control
HopeNet	Fine-Grained Head Pose Estimation Without Keypoints
HRA	Health Research Authority
ICU	Intensive Care Unit
LSTM	Long-Short Term Memory
mAP	mean Average Precision
MICE	Multiple Imputation using Chained Equations
MRI	Magnetic Resonance Image
MTCNN	Multi-Task Cascaded Convolutional Networks for Joint Face Detection and Alignment
NIHR	National Institute of Health Research
PoR	Point of Regard
PoR	Point of Regard
PRC	Precision-Recall Curve
PVG	Planar Viewpoint Generator
RANSAC	Random Sample Consensus
RASS	Richmond Agitation Sedation Scale
REC	Research & Ethics Committee
ReLU	Rectified Linear Unit
RGB	Red-Green-Blue

RGBD	Red-Green-Blue-Depth
RNN	Recurrent Neural Network
ROC	Receiver Operator Characteristic
ROI	Regions Of Interest
ROS	Robot Operating System
RT-BENE	Real-Time Blink Estimation in Natural Environments
RT-GENE	Real-Time Eye Gaze Estimation in Natural Environments
S ³ FD	Single Shot Scale-invariant Face Detector
SRA	Spectral Residual Approach for Saliency Detection
SVM	Support Vector Machine
TCN	Temporal Convolutional Network
WMH	West Middelsex Hospital
WTA	Winner Takes All

INTRODUCTION

The use of technology in health care is increasing with some implementations even reaching the attention of mainstream news media outlets. This is in no small part due to the increasing amount of data available to health care professionals which had led to the development of sophisticated machine learning frameworks that have significant predictive power (Miotto et al., 2016). Adult Intensive Care Unit (AICU), as a medical speciality, is highly reliant on technology. Patients in AICU are continuously and invasively monitored throughout their stay leading to the generation of massive amounts of data. Technology has also led to the development of sophisticated treatment modalities such as ventilators and dialysis machines. AICU is also very welcoming to novel technology; for instance, recent work introduced the NAO robot into AICU (Sarabia and Demiris, 2013). One of the main challenges for the future of AICU is improving long-term outcomes following the survival of critical illness. Many patients suffer from psychological sequelae following their AICU stay that often results in a dramatic impact on their quality of life. The psychological sequelae often stem from the development of delirium (Jackson et al., 2004).

Delirium is an acute disorder of cognition and memory, often called confusion. It occurs commonly in acutely unwell patients with an estimated prevalence between 20 – 80% in AICU. Its development has been associated with negative outcomes including increased hospital length of stay, worse cognitive outcomes following discharge, and increased mortality (Ely et al., 2001b; European Delirium Association and American Delirium Society, 2014; Holt-Lunstad, Smith, and Layton, 2010; Lin et al., 2004; Lundström et al.,

2007; Meagher, 2001; Meier-Ruge, Hunziker, and Iwangoff, 1991; Palmu, 2011; Pandharipande et al., 2013; Pisani et al., 2006). Consequently, delirium has been included as a feature into other algorithms and scores that predict deterioration (e.g. National Early Warning Scores) and diagnostic tools (e.g. Sepsis-3). The development of delirium also prognosticates future outcomes (Lin et al., 2004; McCusker et al., 2002; Ouimet et al., 2007; Prytherch et al., 2010; Singer et al., 2016).

The formal diagnosis of delirium by Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria requires a thorough investigation and interview by trained clinicians (European Delirium Association and American Delirium Society, 2014). This stringent criterium has been operationalised into bedside scoring systems that aim to diagnose delirium based on heuristics and surrogate markers. One such scoring system is the Confusion Assessment Method in ICU (CAM-ICU). This scoring system aims to assess consciousness and attention as surrogate markers of cognition and memory (Ely et al., 2001a). However, it requires manual monitoring and awareness from the clinician to trigger the assessment. The scoring system itself is burdensome and can miss episodes of delirium owing to the intermittent nature of manual testing. This has the potential to lead to under-diagnosis (Gaudreau et al., 2005; Pisani et al., 2006).

In this thesis, we focus on exploring diagnostic avenues for delirium in AICU starting from time-series Electronic Health Record (EHR) data to incorporating cognitive models for classification using eye movements as a biological signal.

1.1 RESEARCH OBJECTIVES

This work intends to develop a platform for the continuous, automatic, and non-invasive diagnosis of delirium whilst simultaneously giving insights into the potential underlying cognitive processes that drive delirium. This

objective requires the understanding of the clinical pathogenesis of delirium, time series analysis for classification, as well as the abstract incorporation of clinical and cognitive processes in a classification model. With these in mind, this thesis is aimed at addressing the following objectives:

1. Develop a time-series [EHR](#)-based classification system for delirium. For this objective, we aim to collect data from a homogenous [AICU](#) population and use state-of-the-art methods, such as Long-Short Term Memorys ([LSTMs](#)), to model delirium.
2. Identify avenues for novel direct biological signal acquisition. We specifically aim to create a platform that is accurate, precise, non-invasive, and clinically safe for eye tracking in [AICU](#).
3. Use the biological signals acquired from the platform for the classification of delirium. For this, we aim to use further state-of-the-art methods, such as Temporal Convolutional Networks ([TCNs](#)) to classify the acquired biological signals.
4. Explain how these biological signals arise from cognitive processes that drive delirium. We aim to develop novel biologically-inspired architectures that can explain the classifiers under established neuro-cognitive architectures.

We group these objectives into two main categories, namely time-series analysis, and biological signal processing.

1.1.1 *Time-Series Analysis*

Whilst much prior work has gone into the clinical prediction, manual diagnosis, and prognostic consequences of delirium, little work has been done on how delirium as a syndrome has changed. This limits the ability to infer the future burden of disease on the [AICU](#) patients and limits the ability to

plan to cater for the increased demand. We aim to explore, descriptively, how delirium has changed in the AICU at a population level and then use routinely collected EHR data for the diagnosis of delirium, at a patient level.

Also using EHR data, we aim to use routinely collected time-series demographic, biochemical, and observational data on patients in AICU to classify delirium. Delirium, being a syndrome and not a disease, is heterogeneous in its cause. Thus, we aim to use limit the analysis to a single homogeneous disease - namely COVID-19.

1.1.2 *Biological Signal Processing*

Creating time-series classifiers based on routinely collected EHR data provides little in the way of understanding the biological and cognitive underpinnings that drive delirium's manifestations. The routinely collected EHR data provides correlations of disease processes that also cause delirium, and whilst we envisage that these will provide appropriate classification accuracy, they limit the understanding of the syndrome itself. Hence, we aim to focus on collecting novel biological signals from patients with delirium. We focus on visual attention, specifically eye movements as an *overt* manifestation of *covert* cognitive attention where we aim to create a platform for the continuous, non-invasive acquisition of eye movements. We then aim to deploy this platform on patients in AICU, collect their data and analyse it to understand the underlying processes that drive *in-attention* as a surrogate marker of cognition and memory.

1.2 CONTRIBUTIONS

The contributions of this thesis are grouped into five main categories: delirium classification through longitudinal EHR data, development and validation of

an eye tracking platform suitable for use in [AICU](#), deployment of a feasibility trial for the use of the eye-tracking platform in [AICU](#), the classification of delirium from eye movement data, and the development of a novel architecture to explain the visual attention processing in delirium.

1.2.1 *Delirium Classification from Electronic Healthcare Records Data*

We developed an [LSTM](#)-based classification model for the classification of delirium from routinely collected [EHR](#) data in a homogeneous, COVID-19, [AICU](#) population. We demonstrated the model's performance on a validation dataset where we achieved state-of-the-art performance compared to existing models in this population. We also demonstrated that the accuracy of the classifier increases as more data is available indicating that time is an important component for the accurate classification of delirium.

1.2.2 *Development and Validation of Eye Tracking Platform*

We further developed a non-invasive eye tracking platform for the acquisition of eye movements from patients in [AICU](#) in a clinically safe manner. Empiric requirements of accuracy and precision were measured in a clinical setting where we then demonstrated that the performance of the eye-tracking platform exceeded these requirements.

The platform developed required the presence of a robust and fast blinking classifier. Thus, we developed a novel curriculum learning training regime for neural networks which we used on the task of blink detection. We demonstrated that this training scheme reduced the complexity of the required neural networks and resulted in state-of-the-art performance on challenging blinking datasets. The reduction in complexity also achieved higher performance leading to a sampling rate that exceeded Nyquist criteria.

To ascertain where in the scene the patient is looking, we developed a novel gaze-scene intersection method which is fast, accurate, and robust to high noise and sparse inputs.

1.2.3 *Feasibility Study*

We developed a multi-centre study and gained appropriate regulatory approvals from Health Research Authority ([HRA](#)) and Research & Ethics Committee ([REC](#)) to conduct a feasibility study. The study deployed the eye tracking platform across two [AICUs](#) in Chelsea & Westminster Hospital ([CWH](#)) and West Middelsex Hospital ([WMH](#)). We gathered 210 recordings from 35 patients between November 2020 – February 2022.

1.2.4 *Delirium Classification from Eye Movements*

Using the data gathered from the feasibility study using the eye-tracking platform, we developed two time-series classification models. The first model, an eye-movement-only model, demonstrated good classification performance which is improved upon if scene information is also considered. The models were trained on data from [CWH](#) and validated from [WMH](#).

1.2.5 *Competing Forward Models for Visual Attention Understanding in Delirium*

To explain the classification models in a biologically inspired fashion, we developed an architecture suitable for the task of predicting the internal cognitive state from eye-fixation data using multiple competing forward models.

The models are built hierarchically and compete for attribution in a Winner Takes All ([WTA](#)) fashion. The architecture was validated under simulation,

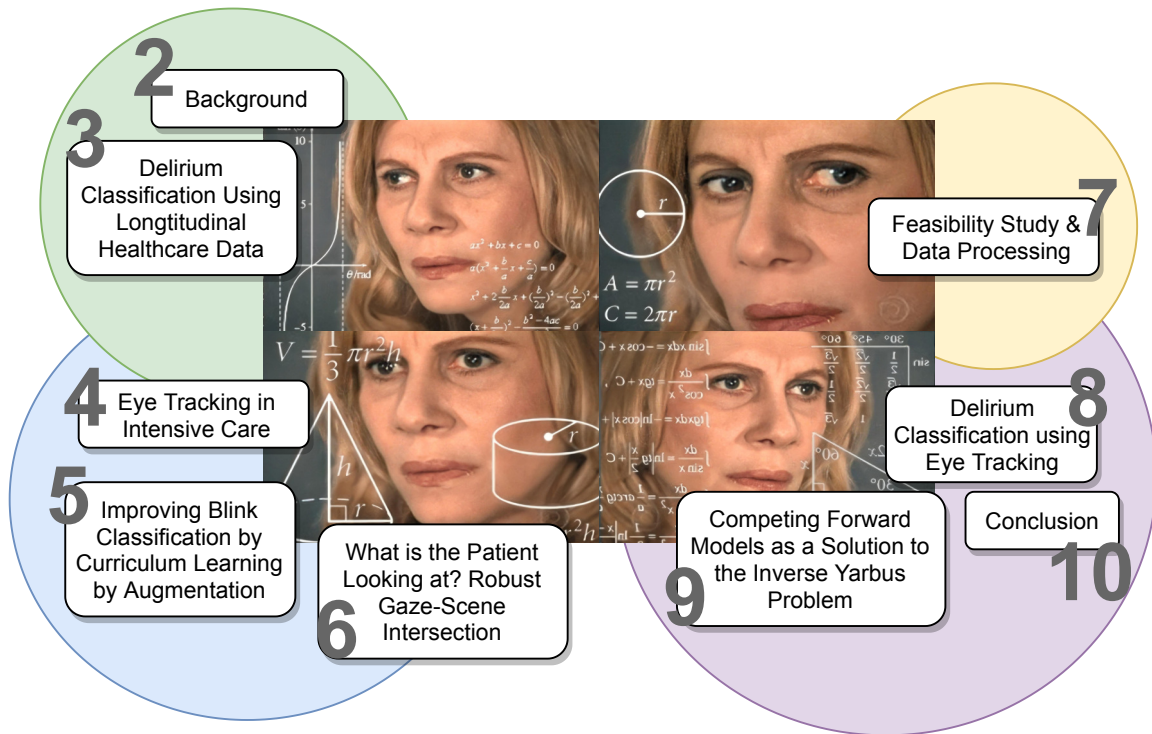


Figure 1.1: Outline of the thesis set to the ‘Confused Lady’ meme demonstrating how eye-movements as a *overt* signal of *covert* internal cognitive mechanisms has permeated popular culture.

then a pre-existing visual search dataset, and then finally used on the data from the feasibility study.

The results of this analysis from the data gathered from our feasibility study demonstrated that episodes of delirium are punctuated by global visual in-attention in a statistically significant manner.

1.3 THESIS OUTLINE

Figure 1.1 illustrates the outline of this thesis pictorially. In detail:

- [Chapter 2](#) reviews the relevant research to this thesis. Specifically, we focus on time-series classification using biological signals and previous work on eye movements for inferring internal cognitive processes.

- [Chapter 3](#) outlines the data, and model, for time-series [EHR](#) classifier of delirium.
- [Chapter 4](#) presents a camera-based platform for the tracking of eye movements in patients in [AICU](#).
- [Chapter 5](#) demonstrates our contribution to improved blink classification facilitating improved accuracy and inference speed.
- [Chapter 6](#) describes our contribution for a novel gaze-scene intersection algorithm to estimate the Point of Regard ([PoR](#)) in patients using the eye-tracking platform on [AICU](#).
- [Chapter 7](#) describes the study that we have conducted, its flow, the patient characteristics, and how the data from each patient's recording session was stored and processed for analysis.
- [Chapter 8](#) describes the Temporal Convolutional Network ([TCN](#)) model specification for two models – an eye movement specific model, and a scene model. We present the architecture, training regime, and results of the model on the dataset collected.
- [Chapter 9](#) details our proposed architecture for the continuous classification of internal cognitive processes where we validate the architecture under simulation, an external visual search dataset, and then go on to use the architecture from data acquired from patients in [AICU](#).
- Finally, [Chapter 10](#) summarises the thesis and present the conclusions of the methodology and results of the work presented. It also highlights the limitations of the work and possible future directions.

1.4 APPENDICES OUTLINE

The thesis has the following appendices:

- [Appendix A](#) presents an algorithm for the comparison of scan-paths between two participants.
- [Appendix B](#) lists the trial documents.

BACKGROUND

Delirium, a cognitive state of confusion, is defined by the Diagnostic and Statistical Manual of Mental Disorders (DSM) (European Delirium Association and American Delirium Society, 2014). In its 5th edition, it lays out specific diagnostic criteria that can be used to diagnose delirium, namely:

1. Disturbance in attention (i.e., reduced ability to direct, focus, sustain, and shift attention) and awareness (reduced orientation to the environment).
2. The disturbance develops over a short period of time (usually hours to a few days) and represents an acute change from baseline attention and awareness, and tends to fluctuate in severity during the course of a day.
3. An additional disturbance in cognition (e.g., memory deficit, disorientation, language, visuo-spatial ability, or perception).
4. The disturbances above are not better explained by a pre-existing, established or evolving neuro-cognitive disorder and do not occur in the context of a severely reduced level of arousal such as coma.
5. There is evidence from the history, physical examination or laboratory findings that the disturbance is a direct physiological consequence of another medical condition, substance intoxication or withdrawal (i.e. due to a drug of abuse or to a medication), or exposure to a toxin, or is due to multiple aetiologies.

The above criteria thus represent an operationalised definition of an acute temporally fluctuating disorder of attention and cognition that are not imme-

diately explained by the patient's pre-existing cognitive disorders (European Delirium Association and American Delirium Society, 2014).

Two phenotypes of delirium have been shown to exist, a *hyperactive* and a *hypoactive* state. Patients can also alternate between the two states with varying frequency in what has been termed a *mixed* phenotype (Peterson et al., 2006).

The *hyperactive* phenotype is heralded by agitation, disorientation in day-night cycles, disturbances in sleep, and is usually a dramatic manifestation. It often attracts the attention of clinical staff and is what most people imagine when one mentions delirium. However, this has been shown to be the minority of patients who suffer with delirium accounting for approximately 5% of all cases of delirium.

By far the most common phenotype of delirium, accounting for the remaining 95%, is *hypoactive* delirium. This is a quiet, introspective state of cognitive and inattention where patients are often have reduced level of consciousness, attract little attention from busy clinicians and nursing staff, and are often mistaken for patients who are 'resting' or 'sleeping'. But on formal testing, they are often found to be covertly delirious. This is the hidden iceberg of delirium.

2.1 OPERATIONAL DIAGNOSIS OF DELIRIUM

The definition of delirium, as outlined above by DSM, is robust, sensitive, and specific. However, it requires trained personnel to perform to achieve its high performance metrics. Operationally, and given the large number of patients in Adult Intensive Care Unit (AICU) and the limited number of AICU psychologists that are present, an easy to use, operational diagnostic tools for delirium were developed. The most popular scoring system is Confusion Assessment Method in ICU (CAM-ICU) (Ely et al., 2004).

This is a simple scoring system that attempts to jointly assess consciousness, through the assessment of the Richmond Agitation Sedation Scale (RASS) score (a clinical score ranging from -5 to $+4$ representing the range of consciousness from coma to agitation), and attention by asking the patient to remember a task and execute it on demand. The task is usually to squeeze the operator's hand when the letter 'A' is said by the operator - the operator then spells out a phrase like 'SAVE A HAART' (HAART intentionally spelt with two A's to ascertain attention) or 'CASABLANCA' (where the BL alters the rhythm). The scoring system then dichotomises the number of hand-squeezing errors where ≥ 2 signifies inattention (Ely et al., 2004). This scoring system has received wide-spread adoption owing to its simplicity and high inter-operator reliability.

Whilst CAM-ICU has facilitate the diagnosis of delirium on AICU, limitations of the use of a scoring system still exist - namely that they are snapshots in time of the cognitive and attentive state of the patient and as delirium is time-varying, can miss episodes of delirium. The scoring system also places burden of work on clinicians and nursing staff and is thus not performed as judiciously as other clinical tasks.

2.2 PREDICTION OF DELIRIUM

Work has been performed on the prediction of delirium. Similar to mortality prediction models, such as the Acute Physiology and Chronic Health Evaluation II (APACHE-II) models, mathematical models have been developed that use data from patient admission variables and predict the development of delirium. Early PREdiction of DELIRium in ICu patients (E-PRE-DELIRIC) is one such model that uses demographic, observational, and laboratory data to to predict the probability of the patient developing delirium during their AICU stay (Wassenaar et al., 2018).

Whilst the model has been validated, it does suffer with relatively poor calibration and has thus received little clinical use beyond informing clinicians of patients with low-risk of developing delirium.

2.3 OUTCOMES OF PATIENTS WHO DEVELOP DELIRIUM

The development of delirium in [AICU](#) carries significant consequences. Patients who develop delirium have globally worse outcomes compared to matched cohorts who do not develop delirium, as will be demonstrated below. The main outcomes that have been studied include mortality, morbidity and complications of the hospital stay, length of stay, cost of stay, and cognitive performance following discharge.

In a matched case-control series of patients who have suffered a neck of femur fracture – patients who developed delirium suffered vastly worse in-patient and post-discharge mortality extending up to 1 year following discharge (McCusker et al., [2002](#)). This trend also applies to a variety of patient cohorts from specialities ranging from Cancer to Cardiothoracics (Bakker et al., [2012](#); Seiler et al., [2021](#))

In-hospital length of stay is also significant longer for patients who develop delirium (Ely et al., [2001b](#); Han et al., [2011](#)). This is as a result of a large amount of clinical uncertainty in future outcomes that result from delirium and thus the patients are deemed unsuitable for discharge. It also relates to the complication that can occur as a result of delirium - increased rates of infection, reduced nutritional intake, increased muscle mass loss (sarcopenia) and ultimately a physiological inability to cope with small pathological stressors - like infection or metabolic derangement.

Economically, patients who develop delirium costs more to keep in hospital (Leslie and Inouye, [2011](#)). This is independent of their length of stay and represents an increased nursing workload, poly-pharmacy, increased cost

of investigations, and increased multi-disciplinary input from a variety of allied health professionals including physiotherapists, occupational therapists, pharmacists and doctors.

Cognitive outcomes following discharge are also worse. In a study, where patients who were discharged home underwent cognitive testing, patients who developed delirium had scores similar to patients with mild Alzheimer's disease 6 months following Intensive Care Unit (ICU) discharge (Jackson et al., 2004). This represents a large burden of disease on the patient, their social network, and on society.

The burden of delirium on AICU is also increasing. To explore how delirium has changed over time, we used an internal data-set at Chelsea & Westminster Hospital (CWH) to explore variables associated with delirium. The variables studied in this dataset relate to factors that are known to cause delirium, namely Age and Severity of disease through the APACHE-II model. We also explored the E-PRE-DELIRIC score.

Figure 2.1 demonstrates data from from 2011, the year of the inception of the dataset, until the analysis date of March 2018. The Age of the patients being admitted to AICU is increasing slowly which mimics national trends of changing behaviour towards the admission of the elderly into AICU.

This trend, however, was not present in the severity of disease; Figure 2.2 demonstrated the APACHE-II score over the same date window. There is no visual discerning trend of increased severity from the periods of 2011 – 2018.

However, Figure 2.3 demonstrates how the expected rates of delirium using E-PRE-DELIRIC have changed over time – there is a clear visual trend of an increase in the predicted development of delirium. This has potentially serious implications as it seems that E-PRE-DELIRIC is also correlated with outcomes as per Figure 2.4.

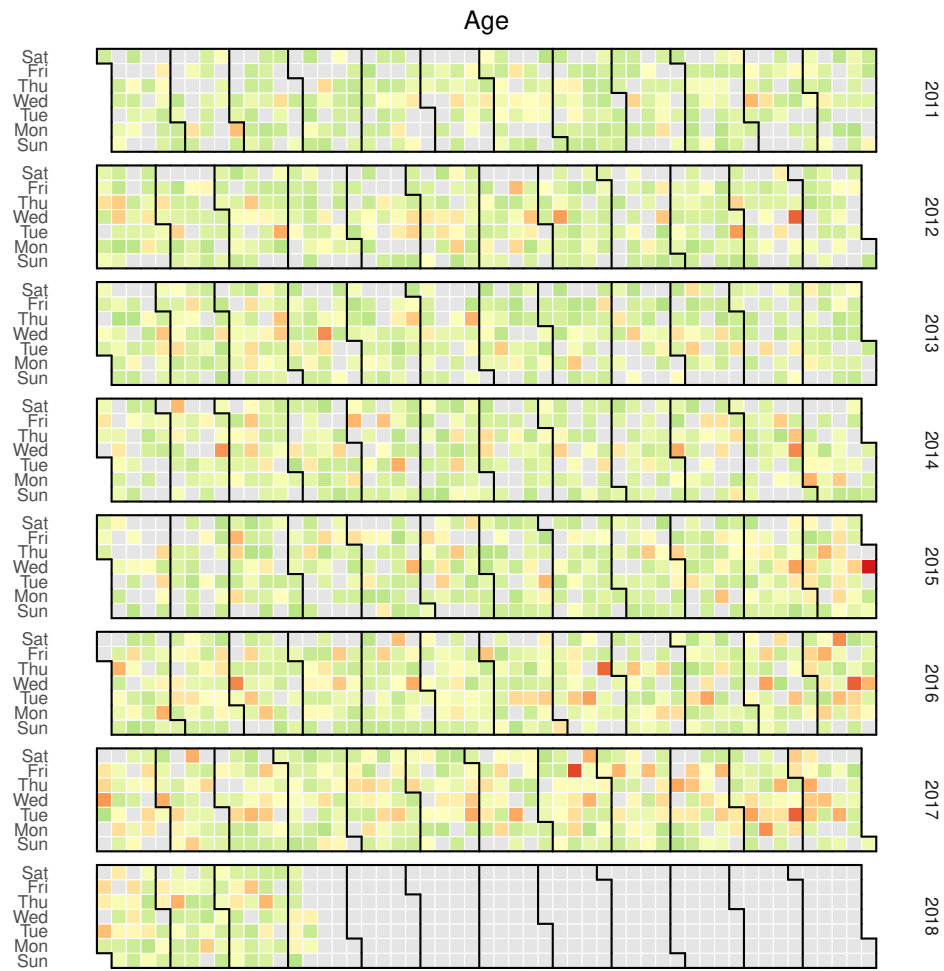


Figure 2.1: Calendar plot of admission age of patients admitted to AICU over time starting from 2011 until 2018 demonstrating stable age of admission in the dataset. Grey indicates lack of new admissions and thus no data, Green indicates low average age and Red indicates high average age.

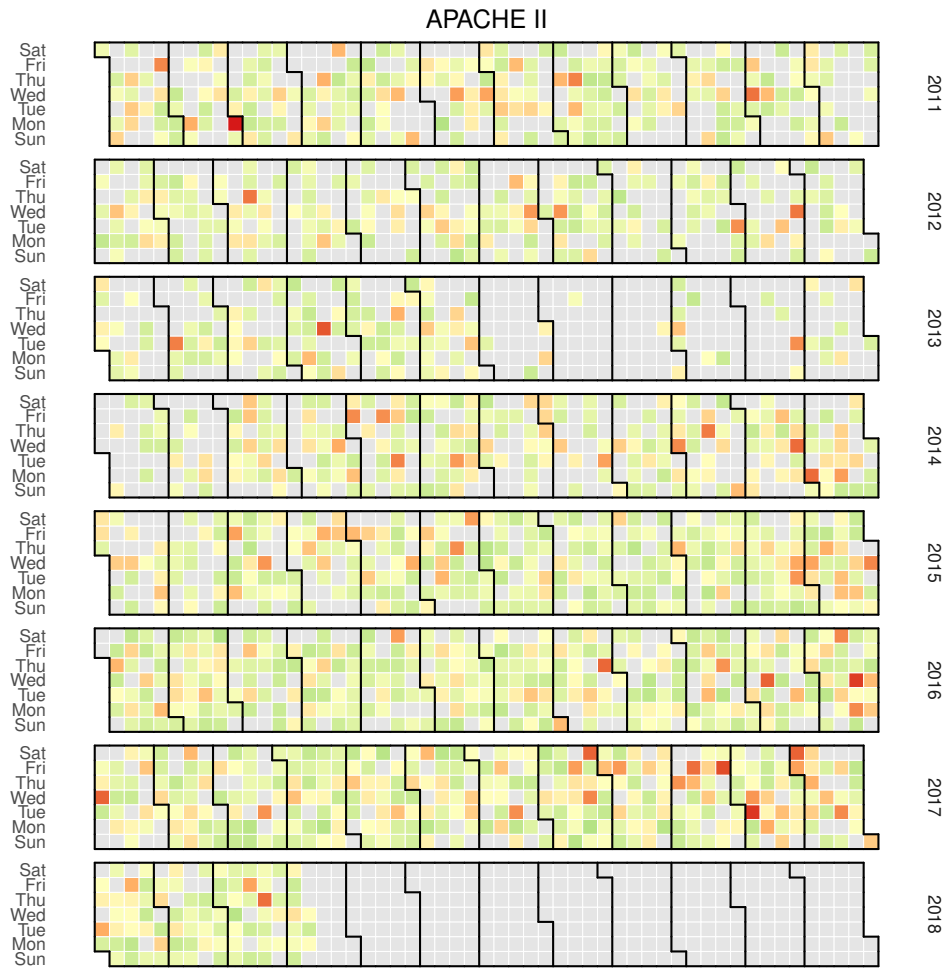


Figure 2.2: Calendar plot of the severity of disease as modelled by [APACHE-II](#) over the dataset's duration of 2011 to 2018 demonstrating stable severity of disease in the dataset. Grey indicates lack of new admissions and thus no data, Green indicates low average acuity and Red is high average acuity.

2.4 EYE-MOVEMENTS IN DISEASE

The role of eye-movements in neuro-psychiatric diseases has a long tradition. Many developmental and degenerative diseases have pathognomonic changes on eye-movement analysis. In this section, we review the role of eye tracking in specific diseases and the techniques used to ascertain those eye-movements.

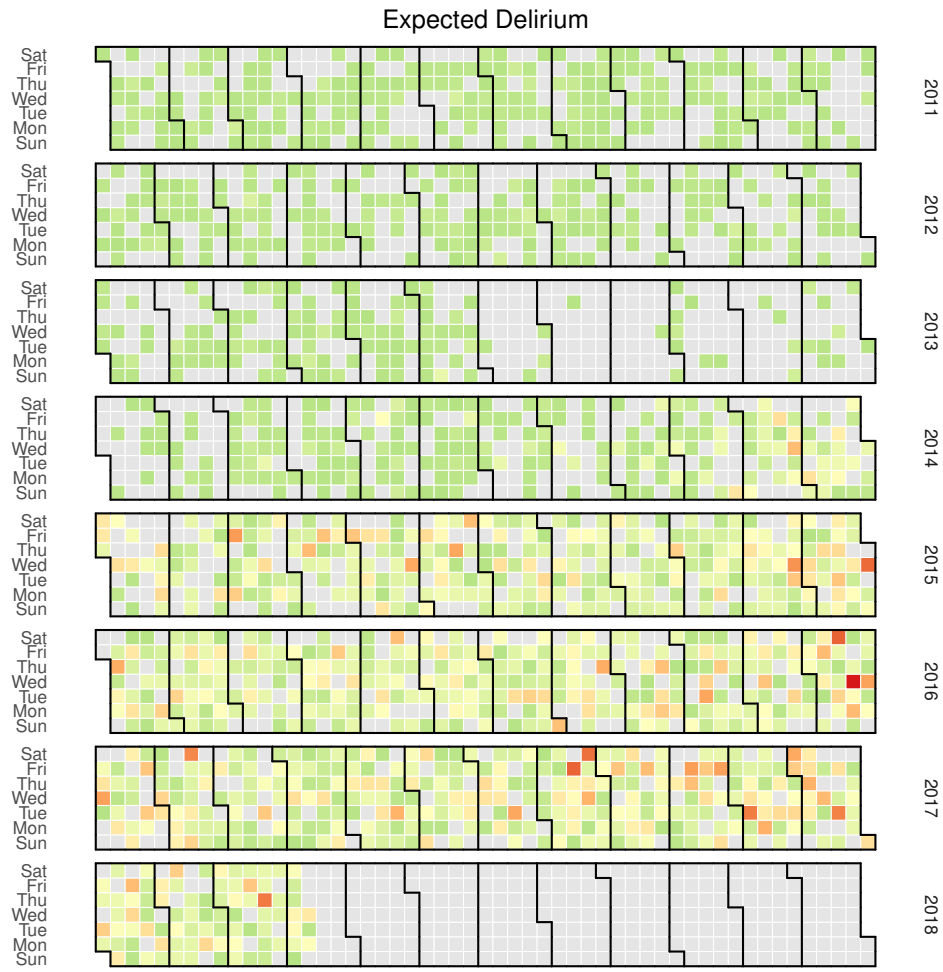


Figure 2.3: Calendar plot of the [E-PRE-DELIRIC](#) score over the dataset's duration of 2011 to 2018 demonstrating that risk of delirium is increasing over the dataset. Grey indicates lack of new admissions and thus no data, Green indicates low risk of delirium and Red is high risk of delirium.

2.4.1 *Eye-Movements in Dementia*

Alzheimer's disease, a form of dementia characterised by loss of cortical neurons resulting in loss of cognition and memory, is a devastating disease. Its diagnosis requires observational, radiological, and sometimes histological investigations, with expertise to interpret the data placing it in a clinical context. Thus, its diagnosis is of paramount importance and carries a significant impact for the patient, their family, and for society (Lehman et al., 2004).

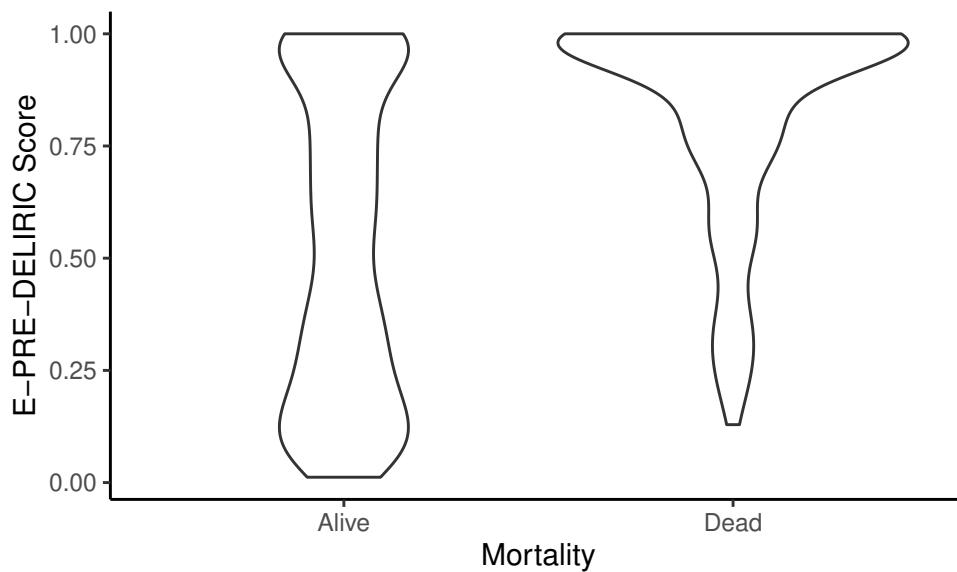


Figure 2.4: Violin plots of patient's outcomes and [E-PRE-DELIRIC](#) score demonstrating that patients who have died have a much higher [E-PRE-DELIRIC](#) score compared to those who were discharged alive.

Eye-tracking was used early in the disease's history. It was demonstrated that patients with Alzheimer's disease didn't fully explore scenes that contained incongruous or irregular figures compared to ones that contained congruous or regular figures - e. g., the region of a slide that contained a horse that had his hind legs painted over was not looked at at all (Daffner et al., 1992). This *in-attention* persists even after a 12-month follow-up (Crawford, 2015). A similar finding was demonstrated in patients exploring the visual features of human faces. In this study, 'diminished curiosity' was reported to be significantly different between the patients with Alzheimer's disease compared to the baseline population (Chan et al., 2018).

Similar findings have been reported in patients with Parkinson's disease and Parkinson's dementia. In this population, impairment of rapid eye movements (saccades) and the fixations interspersed between them were pathological compared to patients without those diseases (Archibald et al., 2013).

Thus, one can reasonably conclude that neurological degenerative cognitive disorders have a measurable effect on eye-movements that are consistent and repeatable.

2.4.2 *Eye-Movements in Psychiatric Diseases*

Much of the success of eye-tracking in psychiatric diseases stems from the finding that patients with schizophrenia perform worse on the anti-saccade test compared to controls (Munoz and Everling, 2004). Other eye-movements, specifically smooth-pursuit and saccade generation, are also exhibited between patients with schizophrenia, and affective disorders, compared to normal controls (Abel, Levin, and Holzman, 1992).

In a more natural natural setting, free-viewing performance of human faces also demonstrated reduced visual search of faces compared to controls but also compared to other patients with affective disorders - such as depression (Loughland, Williams, and Gordon, 2002). This in-attention to facial features is striking especially given that schizophrenia is not a degenerative disease, but is classically thought of as a neuronal network connectivity disorder.

2.5 TIME-SERIES CLASSIFICATION

Much of the work in this thesis relies on the classification of time-series data, often in situation where the entire time-series of a patient (as in the case of the use of time-series data from Electronic Health Record (EHR) for the classification of delirium, see Chapter 3) or a recording (as is the case for using eye-movements for the classification in delirium, see Chapters 8 and 9) has a single label – a supervised classification scheme. To avoid feature engineering and variable selection as there isn't an established method for feature engineering or variable selection in this field, we opt to use

models that can use raw data as their input to output the required class probabilities. Two such models are Long-Short Term Memory ([LSTM](#)) and Temporal Convolutional Networks ([TCNs](#)). Here, we give a brief overview of the underlying architectures that facilitate their time-series performance.

2.5.1 Long Short-Term Memory Modules

An [LSTM](#) is a type of Recurrent Neural Network ([RNN](#)) that can process arbitrary length sequences of data to output a fixed-length vector that represents that sequence. The [LSTM](#) is composed of an input gate, an output gate, and a forget gate, where they work in unison to regulate the flow of information in and out of the [LSTM](#). The fixed-length vector can then be used to classify the time-series data into the required classes. The main advantage of an [LSTM](#) over a classic [RNN](#) is the ability of the [LSTM](#) to deal with longer sequences due to tackling the vanishing gradient problem during training through back-propagation. Formally:

$$f_t = \sigma_g(W_f x_t + U_f h_{t-1} + b_f) \quad (2.1)$$

$$i_t = \sigma_g(W_i x_t + U_i h_{t-1} + b_i) \quad (2.2)$$

$$o_t = \sigma_g(W_o x_t + U_o h_{t-1} + b_o) \quad (2.3)$$

$$\tilde{c}_t = \sigma_c(W_c x_t + U_c h_{t-1} + b_c) \quad (2.4)$$

$$c_t = f_t \circ c_{t-1} + i_t \circ \tilde{c}_t \quad (2.5)$$

$$h_t = o_t \circ \sigma_h(c_t) \quad (2.6)$$

where $\vec{x}_t \in \mathbb{R}^d$ is the input vector to the [LSTM](#) at time-step t , $f_t \in (0, 1)^h$ is the forget gate's activation vector, $i_t \in (0, 1)^h$ is the input/update gate's activation vector, $o_t \in (0, 1)^h$ is the output gate's activation vector, $h_t \in (-1, 1)^h$ is the hidden/output state vector, $\tilde{c}_t \in (-1, 1)^h$ is the cell input activation vector, $c_t \in \mathbb{R}^h$ is the cell state vector. $W \in \mathbb{R}^{h \times d}$, $U \in \mathbb{R}^{h \times u}$, and $b \in \mathbb{R}^h$ are the

weight and bias matrices learnt during training. Subscripts d and h refer to the size of the input features and hidden units respectively. Initial values of $c_0 = h_0 = 0$, and $\circ$ is the element-wise product. σ_g is the sigmoid function, and σ_c is the hyperbolic tangent function (Gers, Schmidhuber, and Cummins, 2000; Hochreiter et al., 1997).

Practically, inputs $x_0, x_1, \dots, x_t$ are fed into the network and the last hidden/output vector h_t thus encodes all previous inputs. The network is trained by gradient decent through back-propagation to alter matrices W , U , and b to minimise the loss function with respect to the inputs. As the main theme of this thesis is the binary classification of delirium, the loss function employed is usually binary cross entropy unless otherwise stated. One of the main disadvantages of [LSTMs](#) is the lack of parallelism – the calculation outlined above are required to be performed through the entire sequence as information is propagated through the hidden vector h_t .

[LSTMs](#) are used in [Chapter 3](#) of this thesis to predict delirium from time-series Electronic Patient Records ([EPR](#)) data.

2.5.2 Temporal Convolutional Neural Networks

Similarly to [LSTMs](#), [TCNs](#) aim to convert a sequence of inputs into an output of fixed-length. However, the architecture that performs this is vastly different utilising convolutional as the main mechanism and relying on multiple layers with specific strides to obtain a causal structure. To ensure that each neuron received information from an appropriately sized receptive field, dilatations are used. Thus, this architecture enforces a strong inductive bias where each neuron can only perceived temporally local information. Formally, for input

sequence $x_0, x_1, \dots, x_t \in \mathbb{R}^n$, and a filter $f : \{0, \dots, k-1\} \rightarrow \mathbb{R}$, the dilated convolution operator F on element s of the sequence is:

$$F(s) = (x *_d f)(s) = \sum_{i=0}^{k-1} f(i) \cdot x_{s-d \cdot i} \quad (2.7)$$

where d is the convolutional layer's dilation factor, k is the filter size, and $s - d \cdot i$ creates a causal structure that avoids data leakage from the future. To increase the receptive field of the TCN, d is increased exponentially with the depth of the network (Bai, Kolter, and Koltun, 2018).

To achieve large receptive fields, neural networks with a large number of layers is required. Thus, residual blocks are used that learn modifications to the identify mapping are used as this has been shown to facilitate the creation of massively deep architectures that ultimately improve performance (Bai, Kolter, and Koltun, 2018; He et al., 2015).

An empiric evaluation of sequence modelling demonstrated the classification prowess of TCNs in its ability to attend to inputs with large temporal distances (Bai, Kolter, and Koltun, 2018). Thus, we use TCNs in Chapter 8 to predict delirium from time-series data gathered by our proposed eye-tracking platform.

2.6 PERFORMANCE OF CLASSIFIERS

In order to analyse the performance of the models that will be presented later on - specifically using LSTMs for the classification of delirium from EHR data (Chapter 3), and the use of TCNs for the classification of delirium from eye-movement data (Chapter 8), graphs and summary statistics of the model's performance are plotted and calculated.

Specifically, model performance is demonstrated by using Receiver Operator Characteristic (ROC) and Precision-Recall Curve (PRC). Area Under Receiver Operator Curve (AUROC) is the summary statistic of the ROC curve

where it demonstrates the accuracy of the model; a number closest to 1.0 indicates perfect classification and 0.5 indicates performance similar to chance. The [PRC](#) curve demonstrates the trade-off between Recall and Precision with the mean Average Precision ([mAP](#)) summarising the curve: the closer the [mAP](#) is to 1.0 the better the model where 1.0 indicates perfect accuracy compared to a baseline model of 0.0.

The trained models are also evaluated using using normalised calibration plots. These are isotonic curves that compare the estimated binned probabilities compared to the fraction of observed outcome where the diagonal line represents perfect concordance between predictions and observations. This curve's summary statistic is the Brier score: a unit-less score between 0.0 and 1.0 where the best-calibrated model attains a score of 0.0 which is calculated as the sum of residual errors between the prediction and the label (Brier, 1950). Formally:

$$\text{Brier Score} = \frac{1}{N} \sum_{i=0}^N (x_i - o_i)^2 \quad (2.8)$$

where o_i is whether the patient i developed delirium or not as a 1 or 0 respectively, x_i is the calculated probability, and N is the number of patients.

2.7 HANDLING MISSING DATA

Often, especially in data from [EHR](#), data can be missing for a wide-range of reasons. An example of this is the different temporal frequencies of pathology samples such as blood tests; some, like gas concentration measurements, are performed almost hourly, whilst other blood tests are performed at a lower frequency.

This is important for several reasons. The cause of missingness speaks to the underlying processes that caused the missingness in the first place – could missing data be attributable to an entirely different clinical management

that biases results? Secondly, [LSTMs](#) and [TCNs](#) cannot handle missing data explicitly and thus, in-order to use them, one must deal with missing data.

One class of techniques aims to *fill* in the missing data. This is often not performed to try and *predict* what those values might be, but instead, offers a way of using algorithms whilst propagating uncertainty by using multiply imputed datasets. One popular method of imputation is called Multiple Imputation using Chained Equations ([MICE](#)) (White, Royston, and Wood, [2011](#)). This is an iterative procedure that builds N complete datasets by building multiple models that predict the missing variables. By using different starting random seeds, different predictions are made of the possible values that those missing values can hold. Then, one can use each imputed dataset to train a model resulting in N models with N performance characteristics that can then be pooled for an estimate of the uncertainty that missing values can hold.

A specific type of imputation strategy, the iterative imputer, performs the imputation scheme in a round-robin fashion until convergence where the values to do change within a certain pre-specified limit. The implementation of the exact algorithm within the iterative imputer can be modified depending on requirements, such as data type.

DELIRIUM CLASSIFICATION USING ROUTINELY COLLECTED HEALTHCARE DATA

This chapter aims to address the question of whether time-series data can classify delirium. We focus on routinely collected data rather than a specific bio-markers or blood-tests in order to generalise our findings. This is of operational and pragmatic importance – by using routinely collected data, we reduce the burden of diagnosis on patients and clinicians. We also focus on modelling the entire sequence of time-series data, rather than what has been previously attempted in the delirium classification literature of using a single time-point (usually admission variables). This is advantageous in several ways - primarily in that there exists more data and therefore using this data should lead to more accurate classification, but also classical statistical techniques (e. g., using linear regression) potentially require the modelling of complex non-linear time-varying relationships between the predictor and the outcome which requires domain knowledge, for each predictor. Rather than perform this modelling approach which is fraught with heuristics and uncertainty, we instead opt to use Long-Short Term Memorys (LSTMs) as time-series estimators as they can model these complex non-linear interactions implicitly thus avoiding feature engineering. We begin this chapter by introducing delirium within COVID-19, what has currently been done on delirium prediction using Electronic Health Record (EHR) data, followed by our proposed modelling approach.

The COVID-19 pandemic heralded a vast number of acutely unwell patients, many of which suffered a severe form of COVID-19 which is recognised to have a spectrum of neurological complications including delirium in up

to 65% of patients (Mao et al., 2020; Rogers et al., 2020). The pathogenesis of delirium in Adult Intensive Care Unit (AICU) is traditionally thought to be due to a combination of disease severity, hypoxia, high inflammatory response, and thrombosis – all elements present in patients with COVID-19 (Cerejeira et al., 2010; Hirsch et al., 2015; Palmbergen et al., 2012; Tay et al., 2020). Thus, early delirium prediction models have formed a significant role as part of a process to improve patient outcomes, benchmark practices and direct resources, especially if further waves of COVID-19 are to occur.

Prior prediction models, like Early PREdiction of DELIRium in ICu patients (E-PRE-DELIRIC), have been demonstrated to perform poorly in patients with COVID-19 (Al-Hindawi et al., 2020; Wassenaar et al., 2015). We hypothesise that this is because the patient cohort underpinning the model is vastly different from the more homogeneous COVID-19 cohort. Specifically, E-PRE-DELIRIC lacks support from patients who have had a long length of stay, severe respiratory failure, and are young – all elements present in COVID-19.

To address this, we aim to develop and validate an improved model with the specific aim of predicting delirium in patients with COVID-19. As the length of stay of patients with COVID-19 on AICU is relatively long, we hypothesise that in-hospital variables affect the development of delirium and thus aim to utilise time-series data from the entire patient's stay for prediction. To maximise utility and facilitate automatic calculation in Electronic Patient Records (EPR), we focus on demographic, observational, and pathological data rather than clinical notes.

3.1 METHODOLOGY

The study was conducted during the COVID-19 infection outbreak as part of dynamic evaluation and optimisation of clinical pathways at Chelsea and Westminster NHS Foundation Trust. As the report involves routinely

collected, non-identifiable clinical audit data, ethical approval was waived as per the UK policy framework for health and social care (*UK Policy Framework for Health and Social Care Research 2020*).

3.1.1 Data Collection and Transformation

Anonymised data from SARS-CoV-2 positive patients admitted to AICU at the Chelsea and Westminster Hospital was collected from the 1st of February 2020 to the 1st of February 2021. Data from the first wave (1st of February 2020 to the 1st of October 2020) were used for model creation, training, and validation whereas data from the second wave (1st of October to the 1st of February 2021) were used for external testing. All patients included in the study underwent assessment by the Confusion Assessment Method in ICU (CAM-ICU) (Ely et al., 2001a). Patients were defined as having delirium if their CAM-ICU was positive and were defined to be non-delirious if the CAM-ICU remained negative throughout their stay in AICU. Patients in whom delirium assessment, due to precluding factors in CAM-ICU, were excluded from the study.

Patient flow through the study is presented in Figure 3.1 and Table 3.1 lists the patient characteristics stratified by delirium status and model training and testing phases.

Variables were selected based on expert domain knowledge to maintain biological plausibility. Specific attention was given to variables that were known to exist in both the delirium and COVID-19 literature; specifically variables that related to physiological response, hypoxia, inflammation, and thrombosis (Table 3.2) (Klok et al., 2020). Continuous variables were truncated in time to the first positive CAM-ICU. Categorical variables that are known to correlate with the development of delirium, namely history of cognitive impairment, history of alcohol abuse, and history of corticosteroid use were also included.

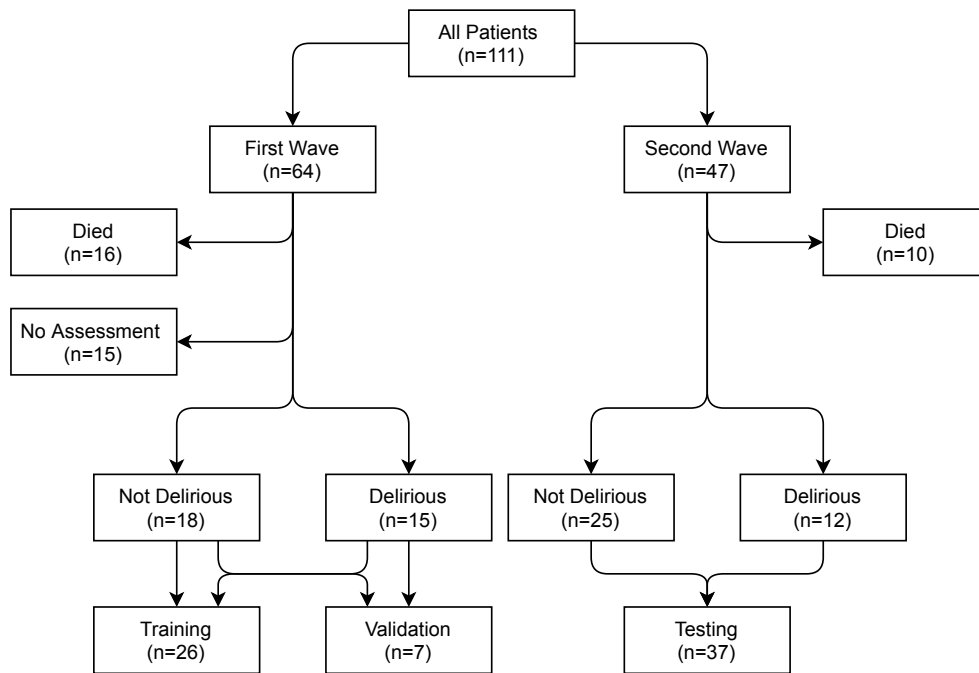


Figure 3.1: CONSORT diagram of patient flow in the study looking at the use of [EHR](#) data for the classification of delirium. Patients in the first wave were used for model building and validation whereas patients in the second wave were used for testing.

Raw time-series data from the entire patient's stay was used without filtering or outlier removal. As the dataset contained missing values, an imputation scheme was employed to propagate uncertainty. This strategy creates N datasets and imputes the missing values using an Iterative Bayesian Ridge estimation until convergence – as explained in [Section 2.7](#). Training data was then split into a training (80%) and validation (20%) cohort. Continuous variables were centred to a mean of zero and a standard deviation of 1 prior to fitting.

Table 3.1: Patient Characteristics split across training and testing cohorts for the classification of delirium using routinely collected healthcare data. Testing cohort (second wave) had reduced severity of disease as demonstrated by a lower mean APACHE-II score and lower Rockwood Frailty score demonstrating the changing admission criteria landscape between the first wave and second.

	Delirious		Not Delirious		Overall	
	Testing (N=12)	Training (N=15)	Testing (N=25)	Training (N=18)	Testing (N=27)	Training (N=33)
Age						
Mean (SD)	56.5 (20.5)	56.9 (12.9)	62.2 (16.8)	52.4 (14.0)	61.8 (16.7)	54.5 (13.5)
Gender						
Female	5.00 (41.7%)	7.00 (46.7%)	8.00 (32.0%)	5.00 (27.8%)	13.0 (35.1%)	12.0 (36.4%)
Male	7.00 (58.3%)	8.00 (53.3%)	17.0 (68.0%)	12.0 (66.7%)	24.0 (64.9%)	20.0 (60.6%)
APACHE II						
Mean (SD)	16.9 (1.44)	22.1 (17.3)	24.2 (17.4)	15.5 (9.18)	23.7 (16.9)	18.6 (13.8)
Rockwood Clinical Frailty Score						
Mean (SD)	3.00 (1.0)	3.20 (0.775)	3.72 (1.06)	2.94 (0.429)	3.32 (1.04)	3.06 (0.619)

3.1.2 *Analysis and Model*

Data Analysis was performed in R version 4.0.1 and Python 3.7. Tensorflow version 2.2 was used for neural network computation and scikit-learn 0.23 was used for performance metric calculation. R was used for data cleaning while Python was used for missing data imputation, hyperparameter optimisation, model selection, training, evaluation, and testing.

Building upon a clinically validated neural network architecture, the proposed model is composed of a multilayer shallow fully connected neural network starting with a [LSTM](#) module (Abdulaal et al., 2020). The last hidden output from the [LSTM](#) module is used to encode the data series into a single latent variable of fixed size where this vector is then used for classification through 6 fully connected dense layers resulting in a single prediction per time series. Rectified Linear Unit ([ReLU](#)) were used as non-linearity functions with the exception being the last layer's activation being a Sigmoid to normalise the output between 0 and 1 as the probability of the patient's variables belonging to a non-delirious or delirious label, respectively. Binary Cross-Entropy was used for the loss function. AdamW optimiser was used with hyper-parameters $\beta_1 = 0.9$ and $\beta_2 = 0.99$ (Loshchilov and Hutter, 2018). As patients had differing lengths of stays, a batch size of 1 was used to avoid padding. Training occurred over 30 epochs. Extensive regularisation was employed to reduce over-fitting and improve generalisation of the model; specifically, dropout layers, ℓ_2 weight regularisation, and early stopping.

A random search optimisation strategy was utilised to find the optimal set of hyper-parameters to maximise accuracy. A nested cross-validation scheme was used where the training dataset was split into 10 folds where 9 folds were used to train a model with specific hyper-parameters and the last fold was used for validation repeated 10 times and the results averaged (Browne, 2000). The hyper-parameter optimisation random search was bounded within

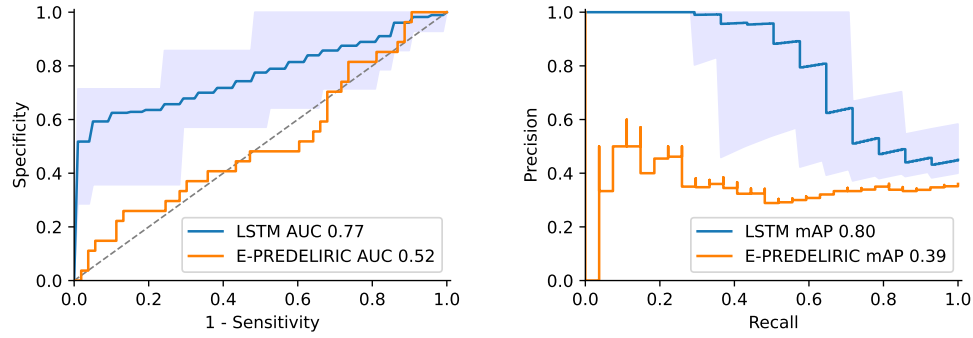
the accepted range of each hyper-parameter. The tuned hyper-parameters were: learning rate, ℓ_2 regularisation, and dropout rate (Bergstra and Bengio, 2012; Gal and Ghahramani, 2016). The best performing model's parameters were stored and used for validation with the best-validated model used for testing.

The performance of the model are presented using Receiver Operator Characteristic (ROC) and Precision-Recall Curve (PRC) curves with their summary statistics of Area Under Receiver Operator Curve (AUROC) and mean Average Precision (mAP). Calibration is presented using the Brier score. Further details of these were presented in Section 2.6.

To understand the effect of time on the accuracy of the models, the external testing datasets (following imputation of missing features) were truncated in time at fractions of the entire sequence of variables and the performance of the model was assessed using AUROC and mAP criteria. The shortest sequence length was 1 at the earliest time – akin to estimating delirium from admission variables.

3.2 RESULTS

For model creation and validation, a total of 66 patients were screened between the 1st of March and the 1st of October 2020. Patients who died prior to a delirium assessment was possible were excluded from the dataset ($n = 16$, 24.2%). Following a notes search, 15 (22.7%) patients were excluded as their delirium status was undetermined. This resulted in a dataset of 35 patients: 23 (65.7%) patients with delirium and 12 (34.2%) patients without delirium (Figure 3.1). These were split into a training ($n = 28$, 80%) dataset and a validation ($n = 7$, 20%). The patients in the second wave were used for external testing ($n = 47$) but 10 patients were excluded as they died before an assessment of delirium was possible (Richmond Agitation Sedation



(a) Receiver Operator Characteristic (ROC) performance curve (b) Precision-Recall Curve (PRC) performance curve

Figure 3.2: Receiver Operator Characteristic (ROC) curves and Precision-Recall Curve (PRC) for classification of delirium from time-series data. The proposed model demonstrates an AUROC of 0.77 with tight confidence. The mAP in the PRC is 0.80 demonstrates an adequate trade-off between sensitivity and positive predictive value. In the ROC curve, the lower triangle demonstrates an area where classification is worse than chance (flip of a coin) and the upper triangle demonstrates better than chance classification.

Scale (RASS) < -3 throughout the duration of their AICU stay) thus leaving 37 patients for testing. Table 3.1 demonstrates the demographics of patients in both the training and testing cohorts.

The random hyper-parameter search resulted in a Dropout probability of 0.2, ℓ_2 layer weight penalty strength of 1×10^{-3} and a learning rate of 1×10^{-4} . This resulted in the best performing model during validation that was then tested on the second wave of patients with COVID-19. The final network architecture is presented in Figure 3.3. Table 3.2 lists the variables that were extracted per patient across their stay on AICU and then used within the model. The table also lists population means and standard deviations used to centre inputs.

The ROC curve for the proposed model using the testing dataset is demonstrated in Figure 3.2. The ROC curves are calculated for each imputed dataset ($n = 20$) achieving an AUROC of 0.77. The Recall-Precision curve also demonstrates good accuracy with a mAP of 0.80 indicating good trade-off between recall and precision. The presented model is adequately calibrated with a Brier score of 0.32.

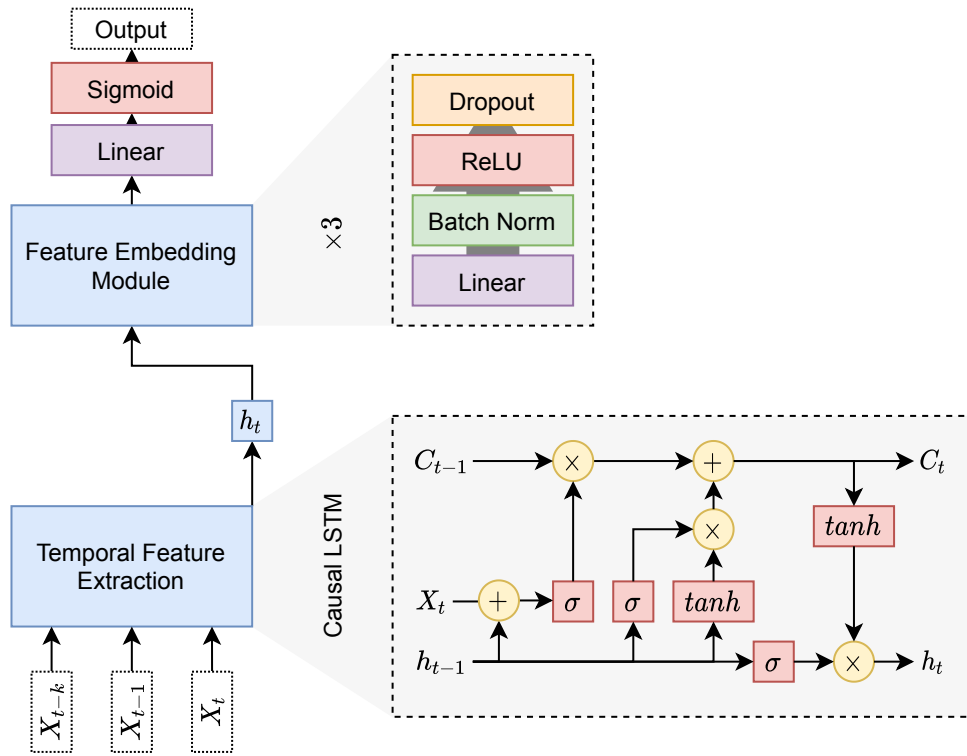
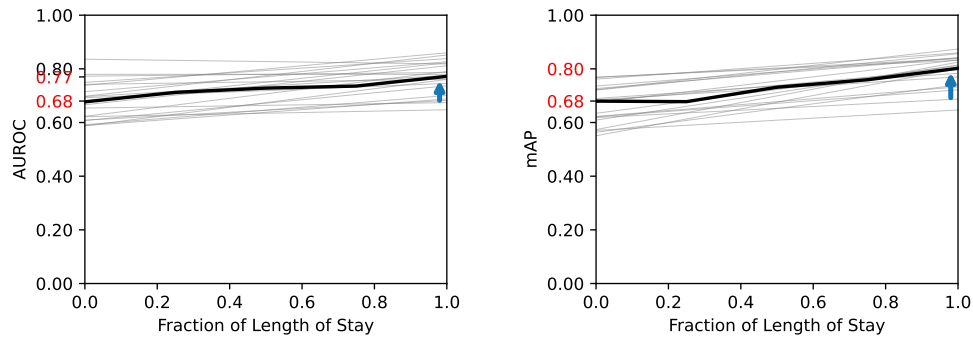


Figure 3.3: The architecture of the neural network used for the analysis of EHR data to classify delirium presented pictorially. The inputs (X at time t) consist of the 17 variables outlined in Table 3.2. Each layer's output then forms the input into the next layer. The first module, a causal LSTM extracts temporal features into a fixed vector. The Feature Embedding Module then embeds these features in higher dimensional space followed by a final linear transformation layer to output the probabilities that this time series corresponds to a Delirious or Non-Delirious patient.

Figure 3.4 demonstrates the increase in accuracy over time when increasing fractions of the data available is presented to the model. The AUROC increases by 0.1 (Wilcoxon Rank Sum Exact Test, $W = 318$, $P < 0.00001$) and the mAP increases by 0.12 from initial admission (Wilcoxon Rank Sum Exact Test, $W = 90$, $P < 0.00001$) to when the entire patient's length of stay is available for the model to predict from.



(a) Area Under Receiver Operator Curve (AUROC) as a function of sequence length (b) mean Average Precision (mAP) as a function of sequence length curve

Figure 3.4: The effect of time, considered as a fraction of the total data available, on the accuracy of the model at classifying delirium from EHR data. As more data is available, the model's accuracy monotonically increases from an AUROC of 0.67 to 0.77 and a mAP of 0.68 to 0.80. Gray lines are individual models with the Black line indicating the average accuracy gain. Blue arrows demonstrate the accuracy uplift.

3.3 DISCUSSION

The development, internal validation, and testing of a neural network model was presented which accurately predicts the development of delirium in patients with COVID-19 using time-series data. To maximise learning from the data and avoid placement of empiric domain specialist knowledge through feature engineering, an LSTM neural network was used on raw time-series data. The neural network, during the training, will decide on which features are important at which time points through altering the weights and biases of the artificial neurons to minimise the loss function (binary cross-entropy) through back-propagation. While this methodology does not shed light on biological machinery or causality, it does provide state-of-the-art performance to an AUROC of 0.77 and a mAP of 0.80 compared to E-PRE-DELIRIC's 0.48 (Figure 3.2).

As with many clinical data sets, the clinical reason and the timing of tests including the recording of the results is varied and thus will contain missing data as the temporal frequency of tests and measurements do not

match. For example, haemoglobin oxygen saturation (SpO_2) is measured continuously but is recorded on [EPR](#) hourly whilst D-Dimer is only measured and recorded daily. Thus, missing data was imputed as neural networks cannot handle missing data explicitly. This imputation procedure was carried out multiple times to propagate uncertainty in the value of the missing variable. To understand the effect of imputation on model performance, 20 datasets were imputed which resulted in 20 independent models. The individual performance of each model demonstrates that the imputation procedure produces negligible effect on the overall model performance.

To evaluate the hypothesis that variables within the patient's stay in [AICU](#) influence the development of delirium, we investigated the accuracy of the models over time. With increasing data, the model increases its accuracy in its predictions with a 0.1 increase in [AUROC](#) and a 0.12 increase in [mAP](#) in a statistically significant manner. This findings suggests two possible hypotheses; either that delirium prediction benefits from learning about patient trajectories during their [AICU](#) stay, or that the classification of delirium is in-determined until more data is available - both hypotheses provide early proof that adding in more data from the patient's stay leads to more accurate classification. While the accuracy of the model using only the first set of observations is adequate ([AUROC](#) 0.67 and [mAP](#) 0.68), the model's accuracy improves as more data becomes available throughout the patient's stay in [AICU](#).

The use of the second wave of patients as a testing cohort adds credence to the validity of this model. These prediction models can be clinically useful in multiple ways – by using the proposed model as the denominator, the model acts as a way of standardising practice as an index which can then be used for comparison between different interventions. This model specialises in predicting delirium in COVID-19 patients which is prudent should further waves of COVID-19 occur. The second, and most appealing, use of this model

is to provide insights into the likely trajectory of patients as they move from the acute phase of illness and into rehabilitation. This rehabilitation period's success largely relies on patient involvement which can only reliably occur if the development of delirium is considered, and clinical actions are taken to reduce the impact of delirium.

There are a number of limitations to the study, the primary limitation is the limited number of patients that the neural network model was able to learn from. This limitation was mitigated by extensive regularisation through ℓ_2 layer weight penalisation, drop out, and early stopping. This ensured that the model does not simply memorise the data, hence its performance on the testing dataset which it has never seen. Recent theoretical work suggests that the number of parameters of the neural network is a poor indicator of network complexity and that over-parametrisation with extensive regularisation leads to better unbiased estimators with low variance (Belkin et al., 2019; Zhang et al., 2021). The second limitation is that the data originated from a single-centre and validity in other centres is yet to be determined.

3.4 CONCLUSION

The development of an artificial neural network-based model was presented with superior performance using shared variables between delirium and COVID-19 literature. Patients with COVID-19 exhibit a long duration of stay in AICU thus making the prediction of delirium from admission variables (as per E-PRE-DELIRIC) biologically implausible. This model thus uses values from the patient's entire stay to achieve state-of-the-art performance compared to pre-existing techniques.

Future work aims to couple the probability of delirium from this model into clinical actions to demonstrate any potential clinical impact on predic-

Table 3.2: Parameters used in the time-series model for the classification of delirium from routinely collected healthcare data. The mean and standard deviations of each variable are used to centre continuous variables during training. At testing, the means and standard deviations are used to centre the values prior to inference.

Parameter	Population Standardising Parameters	
	Mean	Standard Deviation
Age (years)	57.8	10.5
FiO ₂ (%)	43.7	14.8
PaO ₂ (kPa)	10.3	3.4
PaCO ₂ (kPa)	6.3	2.0
SpO ₂ (%)	95.3	3.7
Lactate (mmol/L)	1.2	0.5
D-Dimer (mmol/L)	4148.0	1039.5
Ferritin ($\mu\text{g/L}$)	1273.9	201.0
Sodium (mmol/L)	142.6	2.7
Potassium (mmol/L)	4.5	0.5
C-Reactive Protein ($\mu\text{g/L}$)	149.4	29.2
Urea (mmol/L)	10.3	3.5
C-Reactive Protein (mg/L)	149.4	52.1
White Cell Count ($\times 10^9$)	13.2	3.4
Neutrophil Count ($\times 10^9$)	12.9	2.5
Lymphocyte Count ($\times 10^9$)	1.2	0.3
PaO ₂ /FiO ₂ Ratio	26.3	12.6
History of Cognitive Impairment		NA
History of Steroid Use		NA
History of Alcohol Abuse		NA

tion. Further work would also aim at reducing the number of features to a minimum set of informative variables.

AN EYE-TRACKING PLATFORM FOR PATIENTS IN INTENSIVE CARE

As outlined in [Chapter 2](#), eye-tracking technology has had success in inferring *covert* internal mental states (Amadori et al., 2020; Wang, Amadori, and Demiris, 2018). Clinically, eye-tracking is established in the numerous fields where its success has led to the suggestion that it can be used for the diagnosis of delirium through *inattention* (Exton and Leonard, 2009; Trillenberg, Lencer, and Heide, 2004). The application of eye-tracking in Adult Intensive Care Unit ([AICU](#)) has been held-back by the technological limitations of currently available systems. To our knowledge, no suitable system has been developed or implemented that facilitates the acquisition of eye movements from delirious patients in a clinical setting.

This chapter details the empiric requirements of an eye-tracking platform on [AICU](#), the challenges that are faced by clinical and environmental factors. We then describe our proposed platform, its development and validation, and preliminary results on clinical safety and posterior predictive checks.

4.1 EMPIRIC SYSTEM REQUIREMENTS

[AICU](#) presents a challenge for eye tracking technology. The mixture of patient, disease, environment, staff and medical equipment regulations all coalesce together placing the following clinical empirical requirements:

NON-INVASIVE Due to the nature of patients and their disease in [AICU](#), a non-invasive system is required owing to the need for clinical hands-on care.

Having a device that instruments the patient will interfere with clinical care and is thus potentially unsafe. This also suggests that a device that facilitates free-viewing is required – a system that does not restrict the patient for signal acquisition.

CALIBRATION-FREE In delirium, patients are unable to follow commands making calibration through command-following difficult. The system must be able to track the patient's gaze without an explicit calibration step that maps eye movements to scene gaze positions.

ACCURATE The accuracy of the system must be adequate for use in AICU. A formal definition of accuracy is presented in [Equation \(4.5\)](#).

PRECISE Precision, defined as the reproducibility of measurement, is important as low precision leads to uncertainty on the gaze estimate. Low uncertainty facilitates subsequent analysis through the reliable classification of eye movements. A formal definition of precision is shown in [Equation \(4.6\)](#).

PERFORMANT Eye movement research suggests that fixations typically last between 0.1 – 0.5 seconds giving a minimum required frequency of 20Hz by Nyquist criteria (Feng, 2006). This system would ideally run on commercially available hardware.

4.2 DESIGN AND PERFORMANCE ANALYSIS

To meet the challenges set out by the empirical requirements, we have designed a system that decouples the regression of the patient's eye and head pose from where they are looking in the scene through the use of a dual camera system as per [Figure 4.1](#). One camera located in-front of the patient at

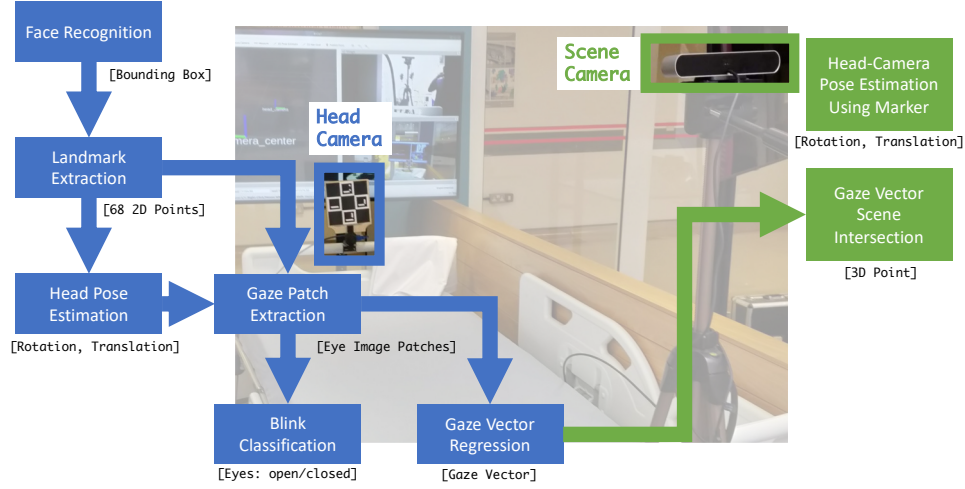


Figure 4.1: An overview of the proposed system for eye tracking in AICU. Two cameras are used: a patient facing camera (head-camera, blue) and a patient-perspective facing depth camera (scene-camera, green) connected to a laptop (outside the viewpoint) running Robot Operating System (ROS). The flow of information through the system is demonstrated with the head-camera and scene-camera running concurrently.

the foot end of the bed facing the patient (termed the head-camera) and another behind the patient facing the same direction as the patient (termed the scene camera). The head-camera is a Red-Green-Blue (RGB) camera whereas the scene-camera is required to be a Red-Green-Blue-Depth (RGBD) camera. The decoupling of the eye gaze regression and gaze inference through the use of two cameras satisfies the first requirement of *Non-invasiveness* as per Section 4.1.

SCENE-CAMERA The scene-camera performs head-camera pose estimation and gaze vector-scene intersection. The head-camera’s pose is measured using a ChArUco board which is fixed above the head-camera – a once only calibration procedure yields a static pose between the head-camera and its board (An et al., 2018). The resolution of the scene-camera is dependent on its ability to resolve this marker reliably - we utilise a 1920×1080 pixel resolution at 30Hz.

Once the gaze-vector has been estimated from the head-camera, it is the scene-camera that resolves that vector as a point in the scene. This is done using a ray-octree intersection that is created from the point-cloud received from scene camera (Glassner, 1984).

HEAD-CAMERA The patient facing *head-camera* carries out multiple steps that ultimately result in the head-pose and gaze-vector of the patient relative to the head-camera. The head-camera's pose as estimated from the scene-camera creates a transformation tree that facilitates the location of the patient's gaze relative to the scene-camera which acquires the point of view of the patient.

The first step is face detection. Following experimental evaluation between various face detectors, the Single Shot Scale-invariant Face Detector (S^3FD) face detector was used for its robust detection at extreme head poses (Zhang et al., 2017). The output is a bounding box around potential faces and confidence of each detection. The camera specification is linked to its ability to reliably produce a high resolution image of the face - we utilise a 1024×768 pixel resolution at 50Hz. To facilitate an appropriately focused image for further downstream computation, an open-loop focusing technique was employed. Focus was calculated as the variance of the laplacian of the gray-scaled image; the intuition is that an appropriately focused image will contain a large amount of edges and thus high variance, formally:

$$Laplace(f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \quad (4.1)$$

$$focus(f) = variance(Laplace(f)) \quad (4.2)$$

where f is the gray-scale image.

In our implementation, and to ease the burden of setup, the focus of the patient's face is communicated to the operator by a set of beeps which vary in their frequency between beeps as a surrogate marker of how focused the

image was: the greater the lack of focus, the shorter the time-span between beeps, which increase in their time-span as the focus is altered towards an image with high laplacian variance. A function that satisfies this is:

$$frequency(f) = \begin{cases} \max(1, \xi - focus(f))^{0.5}, & \text{if } focus(f) \geq \xi \\ 0, & \text{otherwise} \end{cases} \quad (4.3)$$

where ξ is a pre-defined threshold. Thus the operator would alter the focus manually until the beeping as ceased below ξ ; the value of which was found empirically to be 50.

The entire pipeline commences following image acquisition from the *head-camera* followed by face-detection. For each detected face, facial-landmarks are extracted in order to estimate the patient's head-pose (Ohayon and Rivlin, 2006). Facial landmark extraction can be performed by several well-established methods, the most popular of which are Multi-Task Cascaded Convolutional Networks for Joint Face Detection and Alignment (MTCNN), Fine-Grained Head Pose Estimation Without Keypoints (HopeNet), and Face Alignment in Full Pose Range: A 3D Total Solution (3DDFA). These were evaluated for head-pose estimation using a composite score comprised of accuracy, computational performance at inference time, and utilisation of GPU Memory (Ruiz, Chong, and Rehg, 2018; Zhang et al., 2016; Zhu et al., 2019). As 3DDFA had the lowest composite score, it was chosen as the landmark extractor. By knowing the camera calibration matrix and having a predefined model, a direct least squares method is utilised following a Random Sample Consensus (RANSAC) scheme to obtain the pose of the patient's head relative to the head-camera (Fischler and Bolles, 1981; Hesch and Roumeliotis, 2011). This step is then followed by an iterative algorithm to solve the perspective-n-point problem which refines the first, RANSAC direct least-squares, approach which results in negligible degradation of performance by a large up-lift in accuracy and stability of the estimated head-pose. The static 3D model's size

is parametrised by the user’s inter-pupillary distance with 0.06 metres used as the default.

The landmarks extracted from the previous stage also result in the extraction of two image patches of the patient’s eyes which are then used for gaze regression. Real-Time Eye Gaze Estimation in Natural Environments ([RT-GENE](#)), a neural network based model that regresses gaze vector from eye patches, was chosen for its published accuracy in natural viewing environments. For blink-detection, Real-Time Blink Estimation in Natural Environments ([RT-BENE](#)) was used following improvement through a data augmentation scheme and the use of a ResNet-18 as the backbone for increased speed and accuracy (see [Chapter 5](#)) (Cortacero, Fischer, and Demiris, 2019; Fischer, Chang, and Demiris, 2018). [RT-GENE](#) was modified in this work to optimise inter-process communication and reduce latency between image capture and gaze estimation by removing publishers of intermediate calculations. The neural networks in [RT-GENE](#) were also retrained using a ‘Smooth-L1’ loss, formally:

$$\ell(x, y) = \begin{cases} 0.5(x - y)^2 / \beta, & \text{if } |x - y| < \beta \\ |x - y| - 0.5\beta, & \text{otherwise} \end{cases} \quad (4.4)$$

where x is the estimate of the ground-truth y , and $|*|$ is the absolute function of its parameter $*$. This loss function creates a criterion where the mean absolute error is used if the error between the output of the neural network and the ground truth is above a predefined threshold (β), or a mean squared term if the error is below β . This loss function is thus less sensitive to outliers than the original mean squared error loss term used in [RT-GENE](#) and has been shown to stabilise training by the prevention of exploding gradients (Girshick, 2015).

The resolution of the head-camera (1024×768), with a sufficiently zoomed-in lense that results in the patient’s face occupied most of the image resolution,

results in eye-patches that are down-sampled for the consumption through the pipeline at a clinically safe distance of approximately 2 metres. The use of [RT-GENE](#) and [RT-BENE](#) in this pipeline satisfies the requirement for the system being *Calibration-Free* as per [Section 4.1](#).

The required empiric accuracy and precision of an eye tracker in [AICU](#) were estimated based on trigonometric measurements of information in the scene. Measurements of the scene from the point of view of 5 patient from 5 different environments were undertaken across our [AICU](#). Empiric accuracy was defined as the average angular resolution required to differentiate between objects in the scene. This led to the measurement of angles between objects from the perspective of a patient.

To measure an empiric requirement for precision, the angular size of objects from the perspective of a patient were measured. In this context, the size of the object in the scene correlates to repeatability – the smaller the size of object, the higher the precision required for the identification of the object the patient is looking at.

4.3 MEASUREMENTS OF ACCURACY AND PRECISION

To measure the achieved accuracy and precision of the system, six healthy, vision-corrected, volunteers were recruited to undergo testing of the system ([Table 4.1](#) lists their characteristics). An [AICU](#) replicated environment was created under laboratory conditions and the bed was inclined at 30 degrees. The participant's head pose and gaze vector were regressed from the head camera placed in the bottom left hand side of the screen. Each participant was asked to look at the centre of an Aruco marker that would appear at a random location of the screen that was 0.1 metres in size for a screen placed 2 metres away from the participant's head measuring 0.55 metres in height and 0.89 metres in width. For each marker shown, 40 gaze vector

Table 4.1: Demographics data of participants in the validation experiment and the preliminary patients in Continuous Non-Invasive Eye Tracking for the Diagnosis of Delirium (CONfuSED).

Variable	Validation (n=6)	CONfuSED (n=5)
Age	25 - 32	18 - 73
Gender (male = 1)	5	3
Vision Corrected (Glasses = 1)	2	2

measurements were taken and then marker moved to a new random location on the screen that is yet to be measured. Experimental end point was defined as full coverage of the screen by target markers.

For each marker location, the first 10 and last 10 (temporal) measurements were discarded concentrating on the middle 20 measurements. For each marker location, the accuracy and precision were calculated as:

$$\text{Accuracy} = \frac{1}{N} \sum_{i=1}^N \cos^{-1}(|gaze_i| \circ |target|) \quad (4.5)$$

$$\text{Precision} = \sqrt{\frac{1}{N} \sum_{i=1}^{N-1} \cos^{-1}(|gaze_{i+1}| \circ |gaze_i|)^2} \quad (4.6)$$

where $\circ$ is the vector dot product and $|*|$ returns a normalised vector from its potentially non-normalised vector argument $*$. The gaze vector is measured to start from the head pose's transform and extends to a unit length. Target vector is measured to also start from the head pose transform and extends to the target.

For visualisation, median accuracy for each marker is then rasterised onto an image with the resolution of the screen which is then smoothed with a Gaussian kernel.

As part of a clinical feasibility pilot trial that aims to look at the use of this system on the detection of delirium in AICU, the system was deployed on 5 patients (see Table 4.1). We report aggregated descriptive statistical plots

to demonstrate the performance of the system in patients suffering from episodes of delirium compared to episodes without delirium.

In assessing the acceptability of the system clinically, a questionnaire was deployed to the staff that were nursing the patients recruited into CONfuSED as shown in Table 4.2 (MacMurchy et al., 2017).

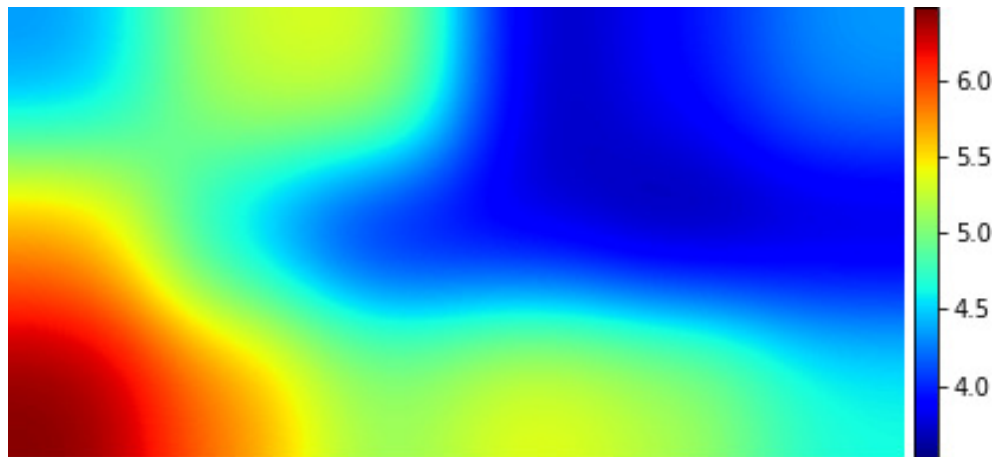
4.4 RESULTS

The platform for gaze regression presented above hinges on the dichotomisation of the head-camera from the scene-camera, and by utilising two cameras placed at a clinically safe distance, the system meets the clinical requirement of being non-invasive. The use of neural networks to regress the gaze-vector (RT-GENE) and classify blinking (RT-BENE) ensure that patient calibration is not required for the accurate and continuous measurement of gaze throughout the patient's AICU stay.

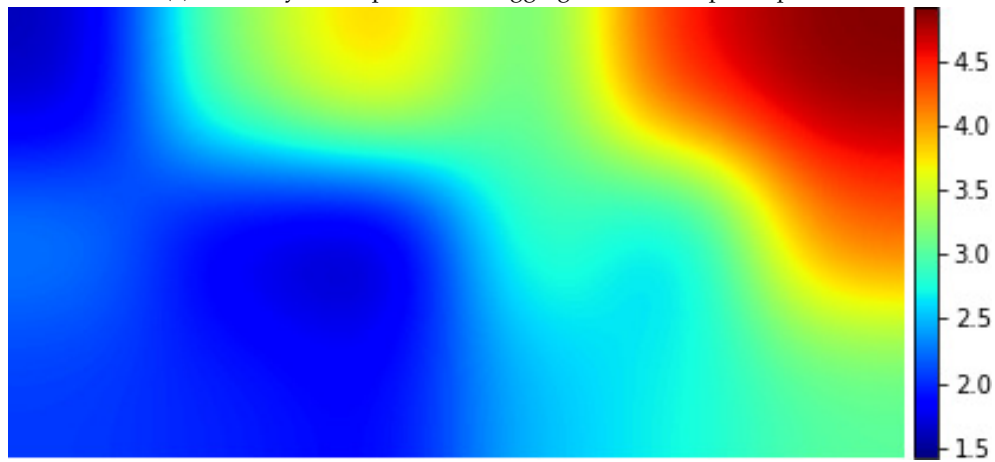
Figure 4.3 demonstrates the distribution of empiric requirements of accuracy and precision as measured from several AICU bed-spaces. The required accuracy and precision resulted from measurements of objects from the perspective of the patient in an AICU bed-space. Empiric accuracy was a result of measurement of the angular distance between objects in the scene from the perspective of the patient whilst precision was the angular size of those

Table 4.2: The questionnaire used assess the acceptability of the use of cameras on day to day activities of the nursing staff on AICU.

Question	Answer
Did this study affect your ability to care for your patients?	Yes/No
Did this study affect your interactions with patients/families?	Yes/No
Did this study affect your interactions with other nurses?	Yes/No
Did this study affect your interactions with doctors?	Yes/No



(a) Accuracy result per marker aggregated over all participants



(b) Precision result per marker aggregated over all participants

Figure 4.2: Aggregated results for all participants demonstrating the accuracy and precision per marker where X and Y axes represent the ArUco marker location on screen in Pixels. The sensitivity of the camera's position relative to the viewer is evident where the highest accuracy and precision is when the user's head pose is directed at the camera.

objects. This reveals that an accuracy of 15.4° and a precision of 6.4° is thus required for a clinically useful system.

Figure 4.3 also demonstrates the accuracy and precision results from the replicated laboratory AICU across all participants pooled across all marker locations. Median accuracy is 4.4° with a median precision of 2.7° as defined by Equations (4.5) and (4.6) respectively. The results of the laboratory validation demonstrate that the system's accuracy and precision exceed what is required in an AICU and thus the platform is suitable for use in AICU.

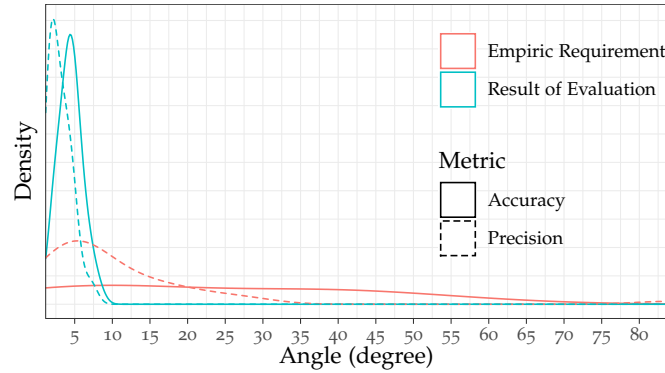


Figure 4.3: Results of the empiric requirements of accuracy and precision compared to achieved accuracy and precision of our proposed eye tracker in AICU. Density plots demonstrating the required accuracy and precision from empirical measurements and the system’s metrics as evaluated in the replicated environment in the laboratory.

We also explored how the accuracy of the system changes as the placement of the camera relative to the participant’s head changes. In the laboratory experiment, grouping by markers, Figure 4.2 displays aggregated results across all participants where each marker’s median accuracy and precision are rasterised as an image the size of the marker and then smoothed with a Gaussian kernel. As the utility of this eye tracking system is in AICU, a median error of 4.4° is lower than the 15.4° . The figure demonstrates the sensitivity of the camera’s placement relative to the user’s head as the error of the system is not uniform; median accuracy is lowest at the bottom left which, from the perspective of the head-camera, represents the most challenging eye image to regress as the eye is mostly closed. Median precision is worse in the top-right where the eye is the most open but the head is rotated giving only 1 eye to regress a gaze vector thus giving the least precision.

Optimisation of the pipeline through internal validation experiments has lead to acceptable performance on commercially available hardware. Total rate of 28Hz is of sufficient performance to yield to eye movement analysis.

To ensure that the platform maintains external validity, distributions of measured patient collected data were checked against that of the distribution of the training data. The system was successfully deployed on 5 patients where

some of them developed episodes of delirium [Figure 4.4](#) demonstrates the measurements that the platform facilitates on those patients in [AICU](#). Blinks, eye and head angles are displayed as aggregated density plots. Extreme head angles are as a result of occlusions of the patient's face - examples of such occlusions are the patient's hands occluding their face, nurses interrupting line of sight between the camera and the patient and nasogastric tubes. These periods of occlusions could be mediated by increasing the threshold for face detection above the current value of 0.6 thus increasing the specificity of detections ignoring periods of occlusion. The figure also demonstrates the aggregate statistics from the datasets used to create [RT-GENE](#) and [RT-BENE](#) along side. The distributions of blinks and gaze data is similar across the patients and the dataset indicating likely accurate acquisition.

The questionnaire outlined in [Table 4.2](#) was conducted on 11 nursing staff, all of whom reported "No" to every question indicating acceptability of the system in clinical practice.

Reflecting on the performance of the platform, it has clearly met the system requirements. Whilst more accuracy and precision is always welcome, this would likely come at the cost of to requirements - such as non-invasiveness if the cameras are to be placed closer to the patient, or the use of deeper neural networks resulting in higher computational requirements for realtime inference.

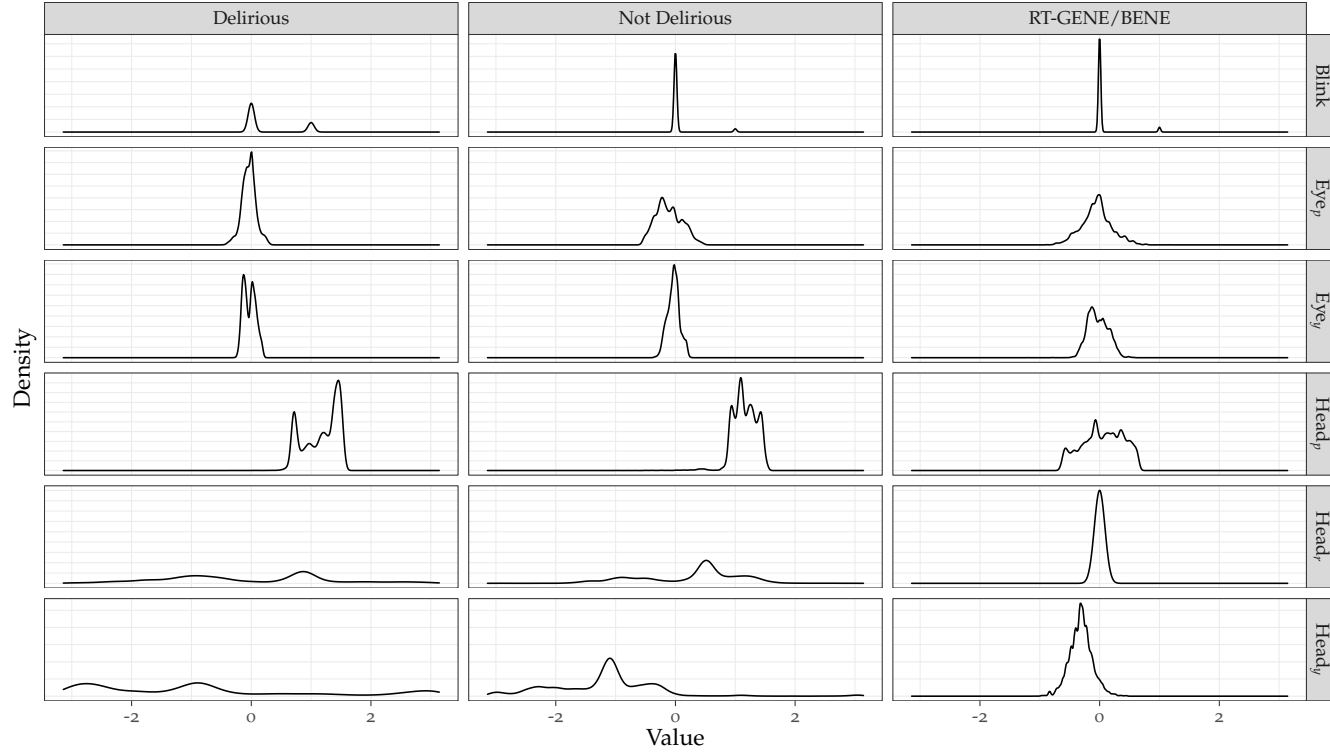


Figure 4.4: Posterior density checks for our proposed eye tracking platform in AICU demonstrating similar ranges of data when compared to the training datasets. Density plots of aggregated data from 5 patients stratified by delirium status. Eye and Head refer to Eye and Head pose angles from the system while subscripts y , p and r refer to yaw, pitch and roll respectively. Blink is classified as per RT-BENE. Blink and Gaze statistics are of similar distributions between RT-GENE/BENE and that of non-delirious patients. Of note, no Head_r data is present in RT-GENE-RT-BENE datasets as faces processed to remove roll prior to inference.

IMPROVING BLINKING CLASSIFICATION BY CURRICULUM LEARNING BY AUGMENTATION

This chapter aims to introduce specific improvements to the blink detection module of the eye-tracking platform presented in [Chapter 4](#). We outline the utility of blinking as a biological signal, previous work on blink classification and our proposed training regime to improve blinking classification accuracy.

Blink detection is a useful biological signal. It is used in multiple applications including for the assessment of drowsiness whilst driving, as a human-computer interface, and for the estimation of cognitive work-load (Królak and Strumillo, 2012; Rahman, Sirshar, and Khan, 2015; Wilson, 2002). Blink detectors are also used to gate deep convolutional neural network gaze regressors as these will regress the gaze regardless if the eye is open or closed. A state-of-the-art blink detector achieves a mean Average Precision (mAP) of 0.757 on a challenging and realistic data set (Cortacero, Fischer, and Demiris, 2019).

Data augmentations have been demonstrated to increase the accuracy of neural networks in various domains (Cubuk et al., 2020; Moreno-Barea, Jerez, and Franco, 2020; Romero et al., 2018). Tailored data augmentations are thought to inject specialist prior domain knowledge leaving augmentation-invariant features as the signal to model, e.g. by augmenting the colour of a picture of dog in a classification task, the creator of the network is indirectly imbuing knowledge that the colour of the dog is not important for the classification task as dogs can come in a variety of colours. These data augmentation schemes thus expand the generalisability of the model into

inputs that the original dataset did not contain (Dao et al., 2019; van der Wilk et al., 2018).

Similarly, *curriculum learning* has been demonstrated to increase the performance of classification (Bengio et al., 2009). The original idea is inspired by the way humans have been taught – a teacher presents a student with a curriculum which increases in difficulty as the student learns the simple basics and expands their knowledge. The difficulty of the data presented increase at a particular pace throughout the training period – akin to a teacher’s curriculum (Bengio et al., 2009; Hachohen and Weinshall, 2019). The result of such a process is an increase in accuracy of the model trained (Bengio et al., 2009). However, creating custom curricula requires the manual labelling of data points which is an arduous task. Thus, automated techniques for labelling the difficulty of each data point have been proposed. These can be categorised into Curriculum Learning by Transfer (CLbT) and Curriculum Learning by Bootstrapping (CLbB). The rate of increase in difficulty, termed the *pacing* function, has also been under investigation; linear, trigonometric and exponentially increasing difficulty functions have been employed in the literature (Sarafianos et al., 2018; Vijjini, Anuranjana, and Mamidi, 2021).

In this chapter, we aim to combine *curriculum learning* and *data augmentation* in a scheme we are terming *Curriculum Learning by Augmentation* (CLbA). This training scheme parametrises augmentations by a *strength* parameter that can then be utilised to increase the difficulty of the data points presented to the neural network using curriculum learning. This method thus removes the requirement for manual labelling of each data point’s difficulty or the use of a pre-trained neural network to label difficulty as is required by CLbT and CLbB (Hachohen and Weinshall, 2019). We demonstrate this training scheme on a classification task of blink detection using a frequency and difficulty imbalanced dataset where our classifier outperforms pre-existing

curriculum learning techniques setting a new state of the art in blink detection performance.

5.1 RELATED WORKS

Curriculum learning takes inspiration from human learning of data $\mathbb{X} = \{(x_i, y_i)_{i=0}^N\}$ where $x_i \in \mathbb{R}^d$ corresponds to a data point with label y_i is presented to a learning algorithm at a certain *difficulty* and *pace* (Hacohen and Weinshall, 2019). A pre-trained *guide* expands $\mathbb{X}$ into $\{(x_i, y_i, d_i)_{i=0}^N\}$ where d_i is the difficulty assigned to x_i .

Previous curriculum learning techniques require a pre-trained neural network. Curriculum Learning by Transfer (CLbT) relies on a pre-trained teacher neural network as a guide to the difficulty of the data point; the teacher, a non-curriculum trained neural network, scores each data point's difficulty proportionally to the magnitude of that data point's confidence across all classes. Curriculum Learning by Bootstrapping (CLbB) is a similar concept - the same neural network is trained twice, the first time without a curriculum where the first network is used to score the difficulty of each data point prior to the training of the second network. The difficulty score is given as the magnitude of the loss function of each data point. In either cases, a previously trained neural network is required, such neural networks will be referred to hereon as *guides*.

Focusing on CLbB, the *difficulty* of each sample x can then be used to mask a classification loss function by:

$$\text{CLbB Loss} = \frac{1}{|\mathcal{X}|} \sum_{\forall (x,y) \in \mathcal{X}} \text{loss}(x,y) \times \begin{cases} 1, & \text{if } \|G(x) - y\| \leq \gamma \\ 0, & \text{otherwise} \end{cases} \quad \mathcal{X} \subset \mathbb{X} \quad (5.1)$$

where $\mathcal{X}$ is a random subset of dataset $\mathbb{X}$, $|*|$ is the cardinality function of its argument $*$, $loss$ is any classification loss function (e.g. binary cross entropy), $G(*)$ is the output of the *guide*, and γ is the current difficulty paced during training. Thus, predictions that are *easy* are learnt first followed by increasingly difficult samples.

5.2 METHOD

5.2.1 Blink Datasets

We focus our efforts on the Real-Time Blink Estimation in Natural Environments ([RT-BENE](#)) dataset (Cortacero, Fischer, and Demiris, 2019). It consists of 243,714 eye images of which 10,444 were closed and 218,548 were open. The rest ($n=14,722$) were uncertain and hence were excluded from the dataset resulting in an imbalanced dataset. In training a classifier, Cortacero *et. al* used a weighted binary cross-entropy loss function that penalised incorrect classification of blinking eyes heavily proportionally to class balances. The dataset was used to train a neural network that achieved state-of-the-art [mAP](#) of 0.997 on the Eyeblick8 dataset. However, on its own dataset, which is more challenging, it achieved a [mAP](#) of 0.757 while utilising DenseNet-121 as the Convolutional Neural Network ([CNN](#)) backbone with subsequent reduction in computational performance due to the use of a large model.

We hypothesise that the lower accuracy achieved on [RT-BENE](#)'s dataset is due to inadequate coverage of the data distribution - Eyeblick8's dataset is of participants sitting in front of the camera in a well-lit environment but [RT-BENE](#)'s dataset originates from more subjects, across a wider variety of lighting conditions while the subjects are unconstrained and are thus free to look around which dramatically increases the difficulty of the problem.

5.2.2 Curriculum by Augmentation

To avoid the requirement of a guide, we instead parametrise the *strength* of augmentations by parameter δ as a surrogate to difficulty. This leads to the simplification of the loss function to:

$$\text{CLbA Loss} = \frac{1}{|\mathcal{X}|} \sum_{(x,y) \in \mathcal{X}} \text{loss}(x^+, y) \quad x^+ \sim A(x, \delta) \quad \mathcal{X} \tilde{\subset} \mathbb{X} \quad (5.2)$$

where $\mathcal{X}$ is a random subset of dataset $\mathbb{X}$, $|\cdot|$ is the cardinality function of its argument $\cdot$, and loss is any classification loss function (e.g., binary cross entropy). Function $A(x, \delta) = (a_1(x, \delta), a_2(x, \delta), \dots, a_n(x, \delta))$ is a stochastic process that applies a sequence of transformations a_n on input x with strength δ to generate a view x^+ . Function $A(x, \delta)$ thus transforms the image x whilst preserving its major semantic characteristics. The proposed loss function thus removes the requirement for a *guide* and a mask.

We create a set of data augmentations that are aimed at solving specific problems in the existing dataset (Karras et al., 2020). In [RT-BENE](#), the eye patches used for blink classification are extracted using a landmark extractor which has its own accuracy profile and thus the eye patches extracted are inherently temporally and spatially noisy, and subject to camera and lens characteristics such as resolution, signal-to-noise, distortion and focus. The data augmentation schemes proposed are thus aimed at improving the robustness of the classifier to inaccuracies of the eye patches. The data augmentations used are:

1. Resize and Crop
2. Perspective Transform
3. Colour augmentation of Brightness, Contrast, Hue, and Saturation

4. Gaussian Blur

These augmentations are aimed at tackling the weaknesses in [RT-BENE's](#) dataset - those of a limited set of lighting conditions and skin tones. To protect against classification failure at inference time, we also augment the dataset with transformations that alter the quality of the eye patches due to camera or lens characteristics.

The strength of the set of augmentations A is controlled by $\delta \in [0, 1]$ where 0 indicates no data augmentation and 1 indicates maximal augmentation. The parameters of each transformation are tuned such that the maximal values (i. e., $\delta = 1$) are representative of a possible eye patch. Each transform is thus able to randomly transform an eye patch up to the strength given by δ . As there are many possible augmentations from the transformations pipeline, even small values of δ make it unlikely that the classifier is presented with the same image again.

In detail, the strength of augmentations related to resize and cropping was controlled as deviations from $1 - \delta$ where low values of δ ensure that little resizing and cropping is possible but higher values deviate from 1 resulting in larger cropping and resizing. Similarly, augmentations of perspective was implemented to increase with $\frac{\delta}{2}$. This maintains a level of plausibility to the perspective shifts that the model could expect.

Brightness and Contrast increased linearly with δ . Here, 0 indicates that no change to the original image whereas 1 indicates maximal change. Hue and Saturation were set to $\frac{\delta}{2}$. Gaussian blur (radius of 1) was controlled with δ being a probability that the blur would be applied (or skipped with a probability of $1 - \delta$).

For each image in the dataset, the resulting image is thus a combination of possible augmentation expanding the available dataset significantly. [Figure 5.1](#) demonstrates how a subjects' eye patches changes with increasing δ and the resulting data augmentations available.

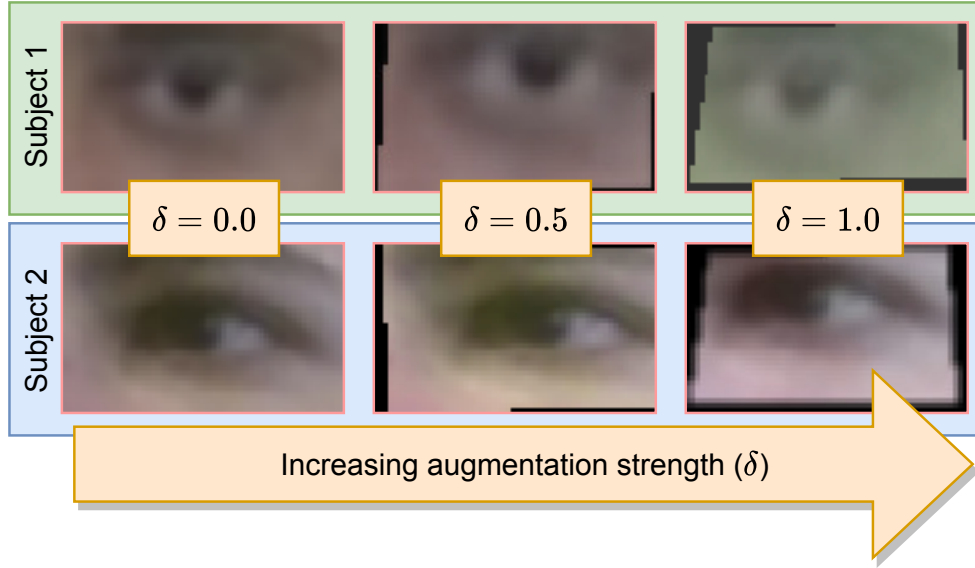


Figure 5.1: The role of the δ as the strength operator for our proposed CLbA training regime. Original image patch ($\delta = 0$) and a sample of the possible data augmentations with increasing strength (and thus difficulty) of augmentation ($\delta = 0.5$ and $\delta = 1.0$ respectively).

We compare our proposed algorithm with CLbB with a linear *pacing* function. To maintain equipoise between the proposed algorithm and CLbB, we utilise the same augmentation scheme for CLbB but with maximal strength (i. e., $\delta = 1$).

5.2.3 Pacing

To utilise curriculum learning under the proposed CLbA architecture, one sets a pacing function that would increment δ from 0 to 1 throughout training. This strategy couples the strength of the data augmentation with the progression of training. Early epochs of training thus learn simple non-transformed eye patches which slowly increase in difficulty. The rate of increase of δ can be

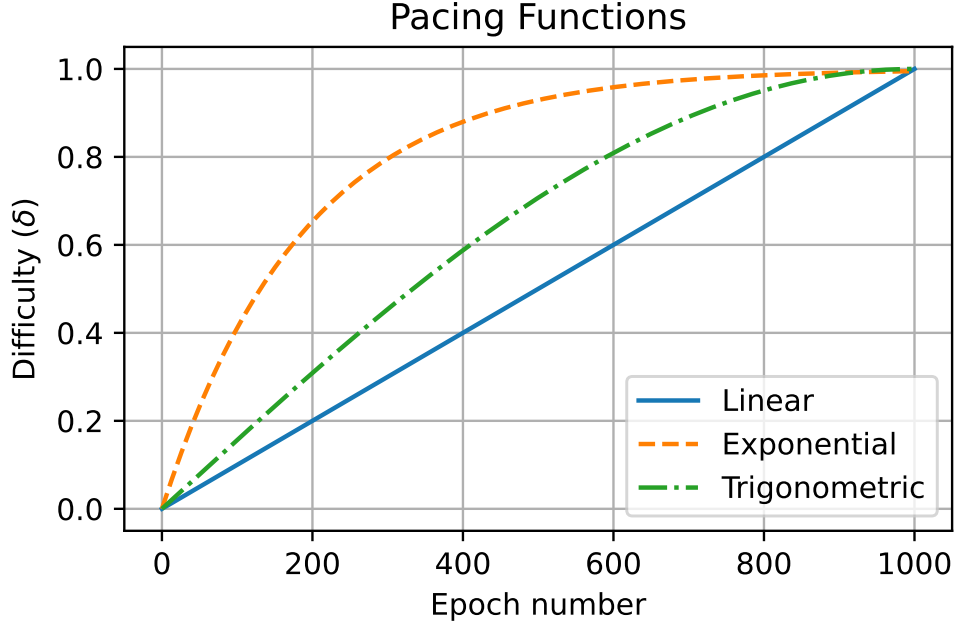


Figure 5.2: Three possible pacing functions that increase the difficulty (δ) of the data-points as training progresses in a curriculum fashion as part of our proposed CLbA training regime.

incremented at particular rates, for CLbA, we experiment with the following pacing functions:

$$\text{Linear} \rightarrow \delta = \frac{\ell}{L} \quad (5.3)$$

$$\text{Exponential} \rightarrow \delta = 1 - (1 + \lambda)^{-\ell} \quad (5.4)$$

$$\text{Trigonometric} \rightarrow \delta = \sin\left(\frac{\ell}{L} \times \frac{\pi}{2}\right) \quad (5.5)$$

where ℓ is the current epoch, L is the maximum number of epochs for training, and λ is the shape of the exponential function (empirically set to 5×10^{-3}) – Figure 5.2 demonstrates those pacing functions visually.

To maintain fair comparison with the CLbB baseline experiments, we increment γ in Equation (5.1) linearly between 0 to 1 throughout training.

5.2.4 *Neural Network Architecture and Training*

Augmented images of the eyes are fed through an ImageNet pre-trained backbone (ResNet-18 in our experiments), the outputs of which are then flattened and concatenated into a fully connected layer followed by batch normalisation and drop-out ($p = 0.2$) followed by a final linear layer that outputs the class probabilities.

The neural network was trained using PyTorch (Paszke et al., 2019). Adam was used as the optimiser with $\beta_1 = 0.9$ and $\beta_2 = 0.99$ (Kingma and Ba, 2017). Eye patch images were resized to 36×60 in pixel size and a batch size of 1024 was used. Learning rate was set to 1×10^{-3} . Weighted Binary Cross Entropy was used for the loss function where data points belonging to the blink class were weighed $\times 24$ compared to the non-blinking class.

Three-fold cross-validation was used for model building with the model achieving the lowest validation loss retained for testing. Four subjects were withheld for testing and were not used for training, validation or model creation. We report performance on those four subjects.

We train four neural networks with different initialisation using a different random seed. We then use the ensemble of these networks and average the outputs for maximal performance.

5.2.5 *Assessment of Model Performance*

To assess the performance of the trained network, Receiver Operator Characteristic (ROC), Precision-Recall Curve (PRC), and their respective summary scalars of Area Under Receiver Operator Curve (AUROC) and mAP are used, whilst model calibration is demonstrated using normalised calibration plots and the Brier score as per Section 2.6.

At inference time, the networks' final layer is passed through a sigmoid function that outputs the likely probability that the eyes are closed. However, as blinking is a binary outcome, the threshold which maximises accuracy is of significance. We report the threshold that maximises the F1-score, which is the harmonic mean of the precision and recall. For the latter, we sweep across thresholds calculating the F1-score at each threshold and pick the threshold with the maximum F1-score.

5.3 RESULTS

Figures 5.3 and 5.4 demonstrates that CLbA produced models that are accurate beating existing state-of-the-art blink detectors and existing curriculum learning techniques; the mAP of the ResNet-18 ensemble model is 0.971 which is significantly higher than the previous state-of-the-art DenseNet-121 score of 0.757 whilst using simpler architecture.

Meanwhile, pre-existing curriculum learning techniques, specifically CLbB, produce worse results than simple augmentation. This is due to class difficulty imbalance; the blinking class is more difficult to predict with a larger error between the bootstrapped network and the label. It is thus presented late in the training process at which point, the neural network has already reached a degenerate solution and fails to predict the blinking class.

Figure 5.5 demonstrates the calibration of the models as a normalised calibration graph. The Brier-score was calculated for each model which demonstrated that the ResNet18 Ensemble models, trained using our proposed CLbA approach, are not only the most accurate but also exhibit the best calibration with a score of 0.01 which is also mirrored in the threshold for the highest F1-score at 0.81. The curriculum data augmentation also improves the model calibration significantly providing tangible evidence that this training methodology improves out of sample performance.

Table 5.1 demonstrates the computational performance of the architectures at inference time on an NVIDIA Titan V. GPU and CPU clocks were synchronised and 1000 runs were performed and their mean presented. An inference rate of 102Hz on commercially available hardware is of sufficient computational speed to satisfy Nyquist criteria presuming a blink duration of 100-150milliseconds, thus aiding real-time acquisition (Bristow et al., 2005; Vaidyanathan, 2001).

The previous state-of-the-art blink detector required the use of DenseNet-121, a powerful but computationally heavy network (Cortacero, Fischer, and Demiris, 2019). Through the use of the CLbA, simpler and more performant architectures, such as ResNet-18, can be used with higher accuracy.

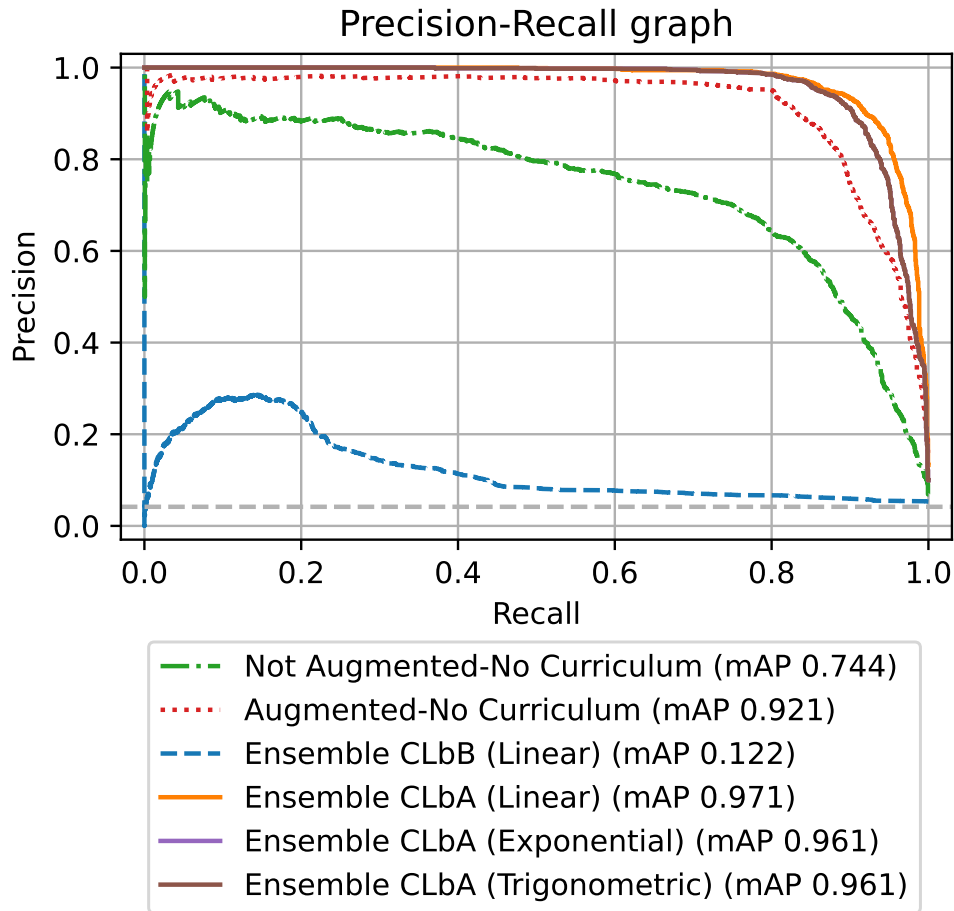


Figure 5.3: The performance of the blinking classifier trained using our proposed [CLbA](#) training scheme. An ensemble of ResNet-18 models achieves state-of-the-art mAP of 0.971 compared to [CLbB](#) of 0.122. Ensemble [CLbA](#) using a trigonometric pacing function and one that uses an exponential pacing function are completely overlapping indicating identical performance.

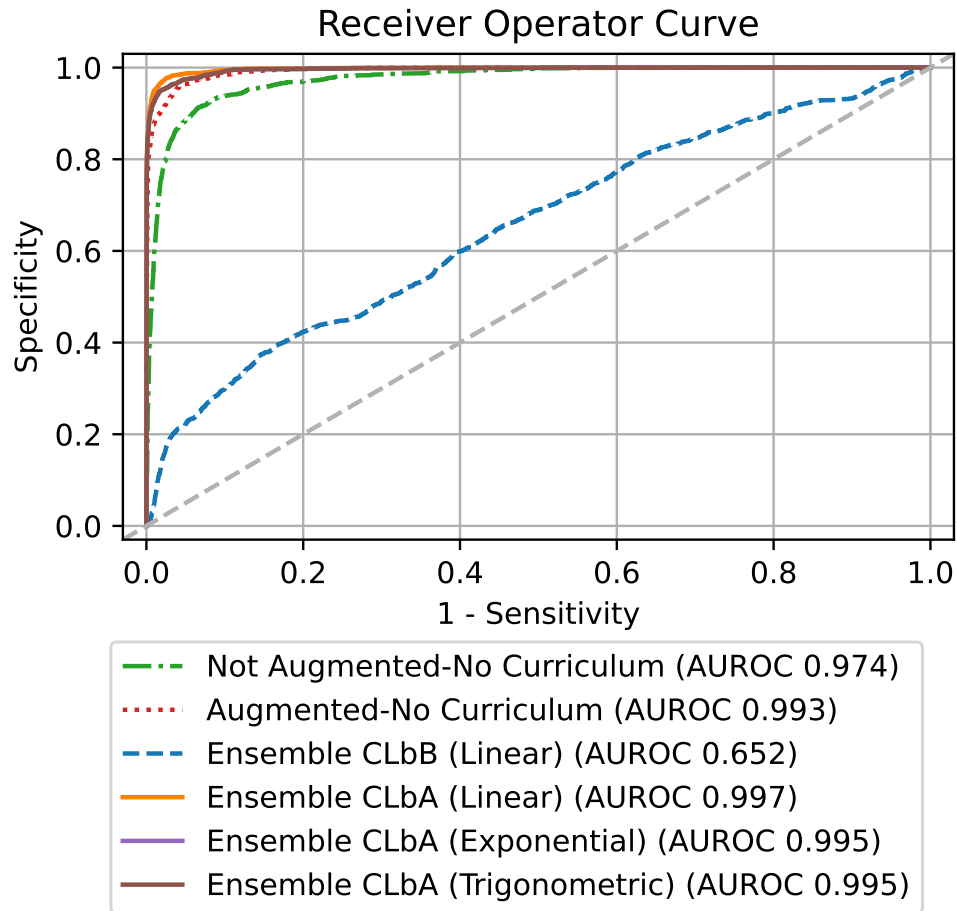


Figure 5.4: The performance of the blinking classifier trained using our [CLbA](#) training regime. The [AUROC](#) of the [CLbA](#) approach is near maximal whilst the [CLbB](#) scheme performs poorly. Ensemble [CLbA](#) using a trigonometric pacing function and one that uses an exponential pacing function are completely overlapping indicating identical performance.

Table 5.1: Performance of our proposed deep neural networks for blink detection and the previous state-of-the-art. ResNet-18 provides the fastest inference time. Previous work (namely Cortacero, Fischer, and Demiris, 2019) does not meet Nyquist criteria and is thus not suitable for real-time inference. Accuracy of ResNet-50 and VGG-16 are not presented in Figures 5.3 to 5.5 to avoid clutter.

Model Name	Depth	Parameters	Inference Rate	Meets Nyquist Criteria?
Densely Connected Convolutional Networks (Cortacero, Fischer, and Demiris, 2019)	121	7.2M	17.6Hz	N
Residual Networks (Ours)	18	11.4M	102.3Hz	Y
Residual Networks (Ours)	50	23.9M	43.4Hz	Y
Visual Geometry Group (Ours)	16	134.7M	38.1Hz	Y

5.3.1 Ablation Studies

We investigate the accuracy contribution of each component of the pipeline including data, augmentation, curriculum learning, and pacing functions to elucidate the effect of each component.

Figures 5.3 and 5.4 demonstrate the results of the ablation studies using ResNet-18. Removing any component of the pipeline achieves an accuracy of 0.744, inline with previous work. Using the augmentation scheme presented, CLbA results in a mAP increase of +0.177 to 0.921. The performance of various pacing functions within CLbA make marginal differences - a linearly paced curriculum performs the best whilst having the best calibration. Trigonometric and exponential functions are similar in performance where they both beat simple augmentation in performance and calibration.

Adding in curriculum learning, the mAP increases by a further +0.036, inline with previous studies on the use of curriculum. Using a 3 model ensemble of ResNet-18 results in a further increase to the state-of-the-art of 0.971.

There were failures of prediction in the testing dataset ($n = 388$, 1.01%). The majority of such failures ($n = 247$, 0.9%) were due to ambiguous labelling where the eye is slightly open with visible sclera but without any visible cornea. The rest of the failures ($n = 141$, 0.4%) were of genuine failures to predict the eyes-closed class - no failure cases exist in predicting the eyes-open class.

5.4 CONCLUSION

We presented a curriculum-based data augmentation scheme that we are terming CLbA. The training scheme implicitly handles frequency and difficulty imbalanced data by increasing augmentation strength as training

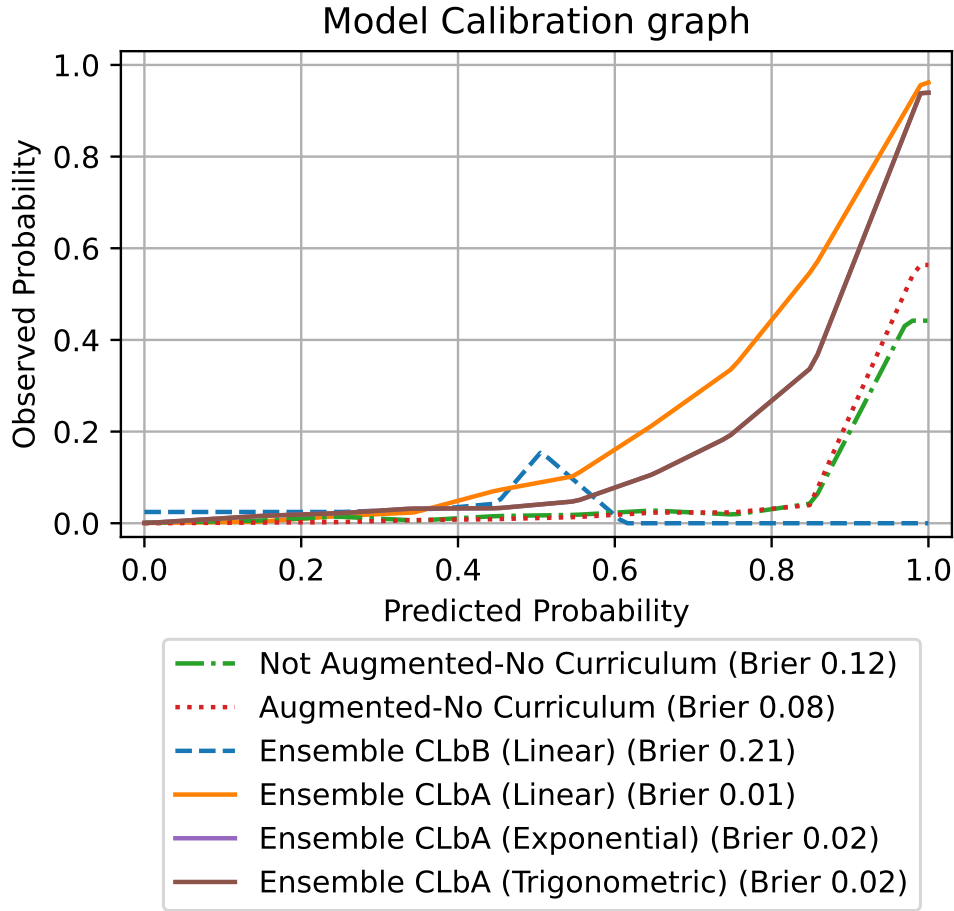


Figure 5.5: Normalised model calibration curve for blinking classification models trained using CLbA against baseline models. Models trained with CLbA produce the best calibrated models whilst CLbB produces poor calibration. Ensemble CLbA using a trigonometric pacing function and one that uses an exponential pacing function are completely overlapping indicating identical performance.

progresses. This results in a increased accuracy and calibration compared to no-curriculum baselines and compared to pre-existing curriculum learning training schemes. We applied the training scheme to improve upon the state-of-the-art accuracy on blink detection using a realistic and challenging dataset.

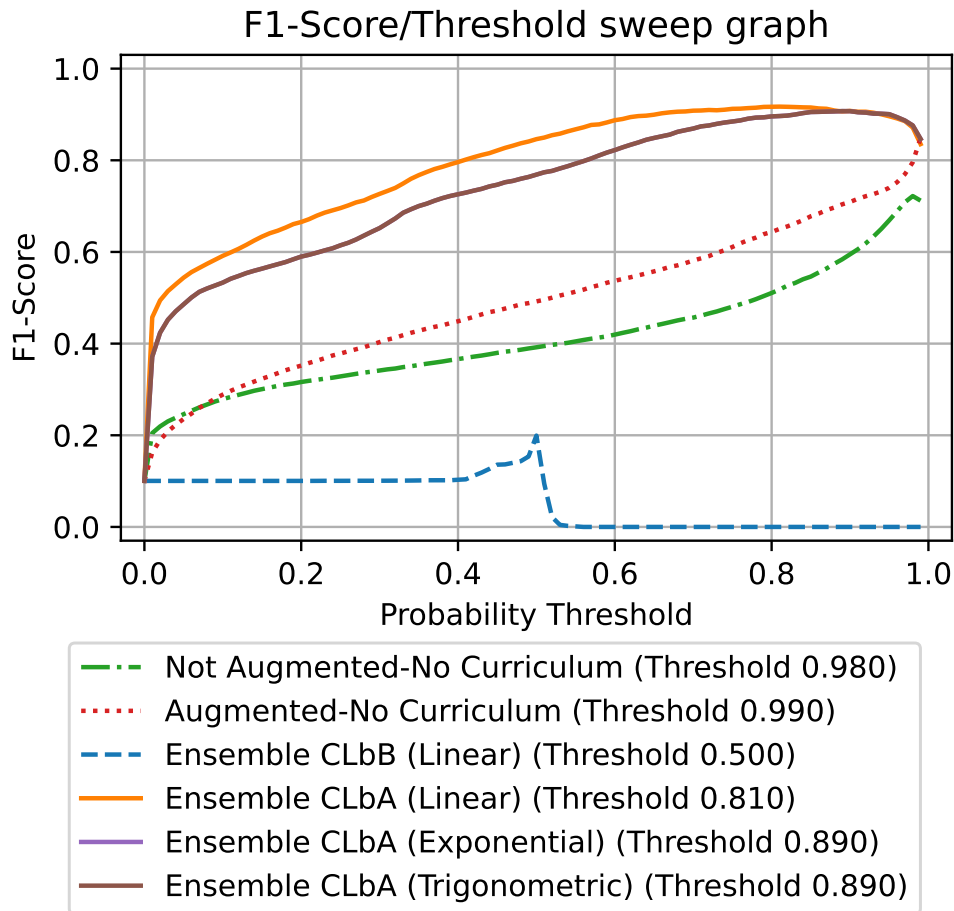


Figure 5.6: Probability thresholds versus F1-score of each blinking classification model. Models trained using our proposed CLbA approach produce appropriately high thresholds maximising the F1-Score whilst baseline models produce lower F1-Scores across all thresholds.

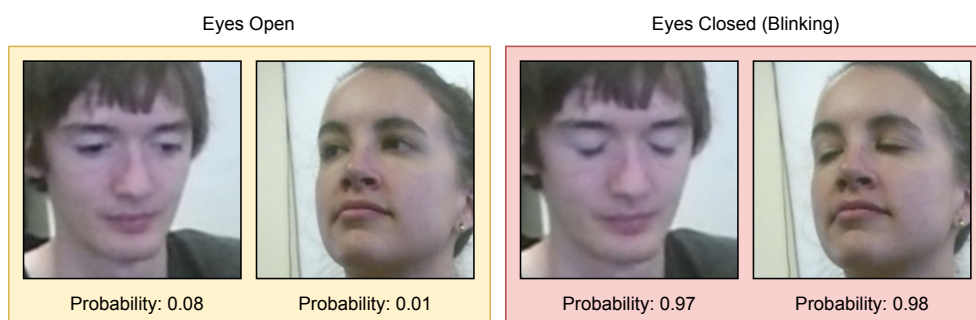


Figure 5.7: Example outputs of the model with respective probabilities in two subjects demonstrating an eyes open and blinking scenarios are classified with high discriminatory power.

WHAT IS THE PATIENT LOOKING AT? ROBUST GAZE-SCENE INTERSECTION

This chapter focuses on the the development of a novel gaze-scene intersection method to estimate the Point of Regard (PoR) – this is the location in the scene that the eye-tracked participant is looking at. We outline the importance of PoR as a concept, the current state-of-the-art in this problem and how our setup mandates a novel solution. We propose an image-based architecture, validate it on a realistic dataset and compare against previous baseline methods.

Following the regression of the gaze vector from the patient’s eye patches, finding the specific part of the scene that they are looking at - often called the PoR or the gaze cue - is of importance. This is of interest in the continuous detection of the location of gaze targets in an eye tracked user in an environment where the scene view is offset/parallaxed from the user’s perspective. Such environments have been deployed in hospitals but similar environments exist in other healthcare environments, vehicles, marketing, and human-robot interactions (Al-Hindawi, Vizcaychipi, and Demiris, 2021; Amadori et al., 2020; Cortacero, Fischer, and Demiris, 2019; Fischer and Demiris, 2016; Ohm et al., 2014).

The location of the user’s gaze in the scene provides cues that facilitate further analysis of internal cognitive states and attention behaviour. Cumulative gaze locations can then be used to generate a temporal heat map to indicate attention in objects in the scene or further analysed into scan paths (Brockmann and Geisel, 2000; Špakov and Miniotas, 2007).

Fixed offset points of view, such as ones in eye-tracking glasses, resolve PoR estimation through calibration. This procedure creates a mapping between a

user’s eye direction and the location in the scene. This mapping, usually a polynomial function with learnable parameters, provides sufficient accuracy due to the low and fixed offset between the user’s viewpoint and the scene camera’s viewpoint (Morimoto, Amir, and Flickner, 2002; Nyström et al., 2013). This calibration procedure is participant and session-specific and can be prone to error, drift and failure (Morimoto, Amir, and Flickner, 2002). However, in situations where the user’s head pose and the scene camera’s pose are not static, and thus exhibit dynamic parallax, such mappings become highly non-linear and become intractable to find as existing mappings do not exist or, worse, produce incorrect results (Ramanauskas, 2006). Thus, these solutions do not work if used in circumstances where the head-pose of the participant has a dynamic offset from the scene camera, such as free-viewing conditions without instrumentation.

If a scene depth map and a gaze vector are given *a-prior*, a brute-force solution is to project $\mathbb{R}^{u,v,d} \rightarrow \mathbb{R}^{x,y,z}$ using the camera’s intrinsics matrix to a point cloud which can then be used to find the point closest to the gaze ray (Wei et al., 2018). However, the step to convert the depth image into a point cloud and then initialising an appropriate dichotomisation scheme is computationally expensive. Thus, image-space techniques have been developed that are aimed at not using point clouds but using depth images directly.

A similar problem, has been tackled in the literature where the **PoR** was estimated from the 3rd person perspective (Fang et al., 2021). A synthetic depth map was generated from a single Red-Green-Blue (**RGB**) image which is then reformulated to facilitate an intersection between the depth map and a generated Field of View (**FOV**) map. This intersection is then used as an input to two neural networks that regress the location of gaze vector in the scene. The algorithm has two key hyper-parameters, the width (σ) of the **FOV** generator and the depth subsection parameter δ which were empirically set to 6 (representing 60° **FOV**) and 0.3 respectively. The architecture’s reliance on

synthetic depth data, and the requirement of the presence of the participant to be present in the frame, makes its immediate use unsuitable for use in an environment where the scene camera is located in a quasi-egocentric manner.

Our contributions are:

1. A geometrically inspired approach to scene dichotomisation that facilitates the calculation of the intersection between gaze vector.
2. Hyper-parameters that are intrinsically explainable and linked to the performance and accuracy of the gaze tracker.
3. A fixed performance envelope of $O(J)$ complexity using subsampling of J points in the cone of gaze.

We demonstrate that this approach is accurate under sparse depth images and is robust to large parallax between the head pose and scene camera position from data obtained in a deployed gaze tracking, clinical, environment. [Figure 6.1](#) demonstrates the results of the proposed architecture on an example scene.

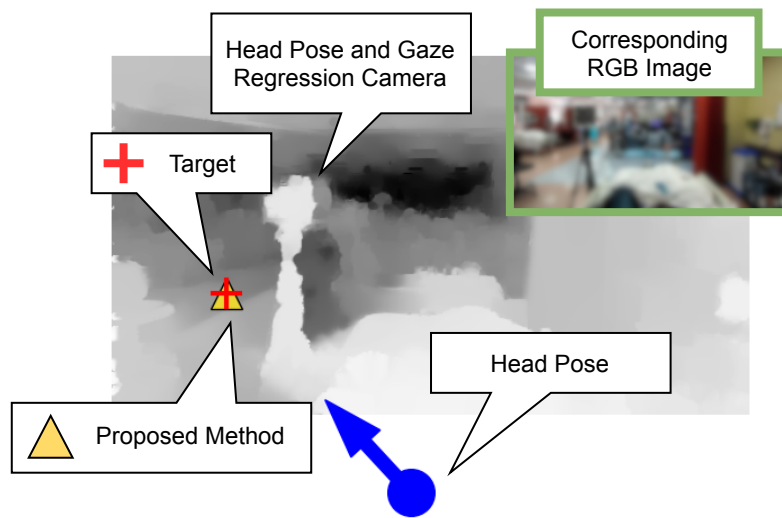


Figure 6.1: A pictorial result of our proposed **PoR** estimation method. An in-painted and rectified depth image with head position and gaze direction indicated as the blue circle and arrow respectively. The red cross is the target label with the detected gaze-scene intersection point labelled as a yellow triangle (RGB image blurred to preserve privacy) with near 100% accuracy.

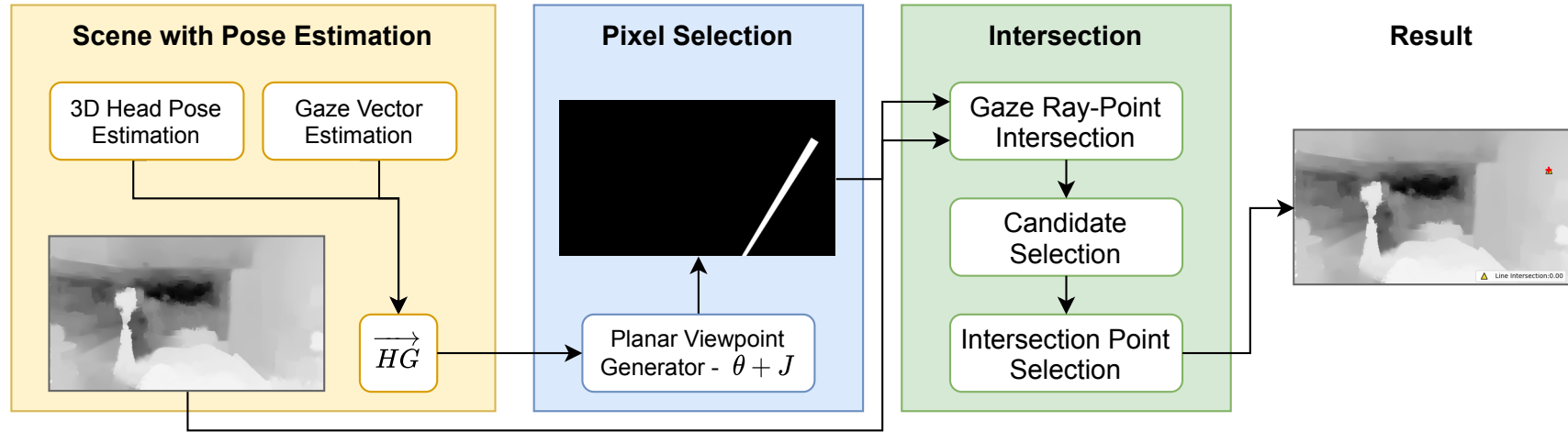


Figure 6.2: The proposed architecture for **PoR** regression. Scene information including depth image (RGB image blurred) is provided where the Head Pose H and Gaze vector G are estimated. A planar cone is then rasterised starting at H with direction G with a central angle of θ and subsamples J . This rasterised image intersects the depth image. Gaze vector $\vec{HG}$ then intersects the set where candidate pixels are then classified whether they are inside the cone of gaze. The successful candidates are sorted and the intersection point selected.

6.1 METHOD

As per [Figure 6.2](#), the first stage of the architecture requires the RGB and depth images, the head location and gaze direction of the participant in scene coordinates to be known. The second stage involves rendering a planar cone map emanating from the viewpoint of the participant which is parametrised by the accuracy of the gaze tracker. This map then subsets the depth map resulting in a small number of pixels tested to intersect with the gaze vector in the third stage. This third, and final stage, intersects each point with the gaze vector, classifies potential candidates as being inside or outside the gaze cone, and selects the most promising candidate to output.

Let $I_D \in \mathbb{R}^{h \times w}$ represent a optionally sparse and optionally rectified depth image with width I_{D_w} , height I_{D_h} with a single channel representing depth. This depth image, as acquired by depth cameras, can contain ‘holes’ that are as a result of background occlusions by foreground objects. In-painting of the depth image, where the holes are then ‘filled’, is often performed for further down-stream processing. We use one such algorithm for in-painting of depth images when mentioned (Bertalmio, Bertozzi, and Sapiro, 2001). This map is to be acquired from a quasi-egocentric viewpoint with known camera intrinsics matrix K . The inverse of this matrix (K^{-1}) is used to convert world points to pixels in camera-space assuming a Pinhole camera model (Dawson-Howe and Vernon, 1994):

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \quad (6.1)$$

where (f_x, f_y) are focal lengths and (c_x, c_y) are the image coordinates of the camera’s centre measured during calibration.

In a rectified image, the focal lengths in K can be the width and height of the image and image centre is half the width and height respectively.

Let $\overrightarrow{HG}$ be a vector that defines to start of the gaze at the head-pose H in coordinates relative to the scene-camera and with direction G . This is given *a-priori* from an eye tracker that also outputs the position of the head relative to the scene camera.

6.1.1 Planar Viewpoint Generator

The Planar Viewpoint Generator (PVG), parametrised by a central angle θ representing the gaze vector of the user from the perspective from the scene-camera starting at $\overrightarrow{HG}$, rasterises a subset of the scene that is viewable by the participant into set S_L . Geometrically, θ represents the accuracy of the eye tracker. Then, S_L is used to create a limited set of 3D world points such that:

$$S_L = \{i_1, i_2, \dots, i_m\} \subseteq I_D \quad (6.2)$$

$$I_d = I_D \cap S_L \quad (6.3)$$

where i_m is the index of I_D and m is the number of elements of the rasterised image and can be intersected with I_D such that the resulting image I_d is a subset of I_D .

As the number of elements in I_d ultimately dictate the performance of the architecture, rather than using the entire set, we instead choose to subsample J points. A 2D Gaussian distribution centred on H is drawn and multiplied by the result of the PVG and then sampled from. We empirically set the scale of the Gaussian distribution to be half the maximal depth, i. e. $\frac{\max(I_D)}{2.0}$. This ensure that points closer to H are sampled at a higher density.

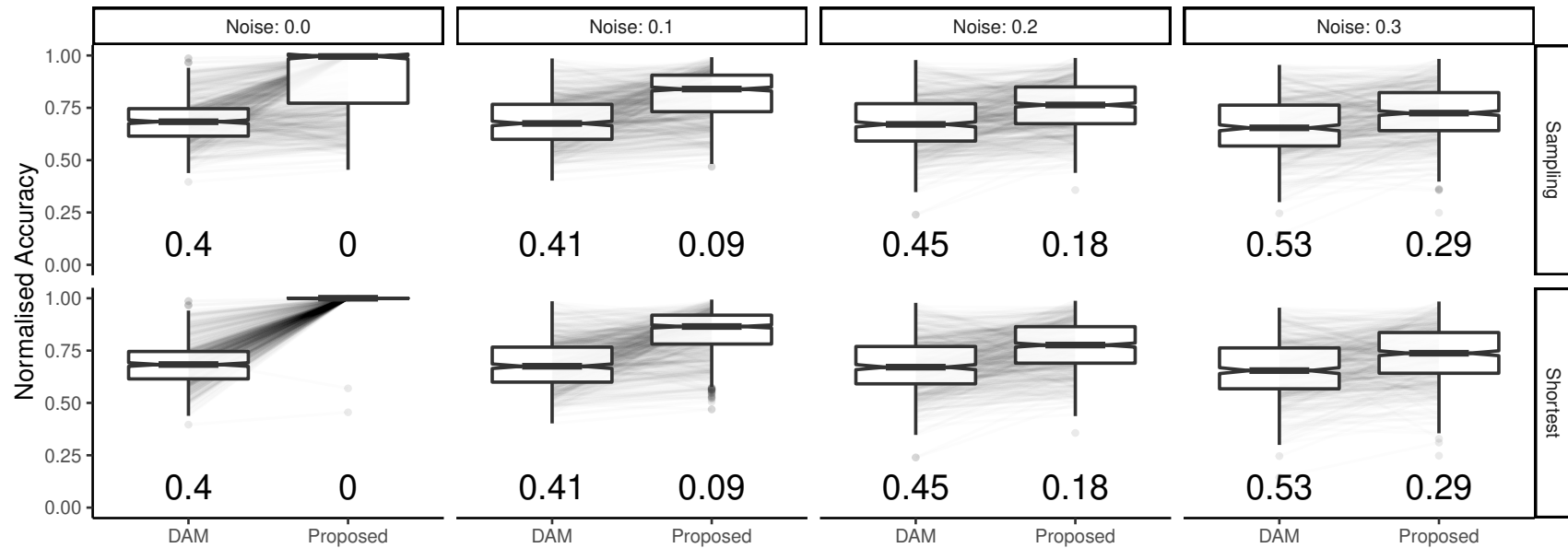


Figure 6.3: Our **PoR** estimation algorithm’s performance compared to Dual Attention Module with increasing noise and thus decreasing precision of the eye-tracker demonstrating that our proposed method surpasses the previous method at all noise levels. Fractions below each plot represent the failures in each technique as a fraction of the overall number of runs. Notches in the box plot that do not overlap indicate statistical significance. Plots are split according to reduction strategy where the Sampling strategy N was set to 5.

6.1.2 Intersection

Using a Pinhole camera model with intrinsics matrix K (Dawson-Howe and Vernon, 1994), the world coordinates of pixels in I_d can then be calculated:

$$W_L = \{K^{-1} \cdot [u, v, I_{D_{u,v}}]^T\} \quad \forall (u, v) \in I_d \quad (6.4)$$

where u and v are pixel indices and $I_{D_{u,v}}$ is the depth of the pixel at u, v . W_L now contains a set of points in the same coordinate system as the gaze vector $\overrightarrow{HG}$. The distance between the gaze vector $\overrightarrow{HG}$ and each point p in set W_L can be found by first reformulating the vector $\overrightarrow{HG}$ to be a line in the form of $HG = H + tG$ where t is a scalar that gives the locus of the line along direction vector G and is set to $\max(I_d)$. The distance between p and $\overrightarrow{HG}$ can thus be given as the normalised projection onto that line:

$$\text{distance}(p, HG) = \|(p - H) - ((p - H) \cdot G)G\| \quad (6.5)$$

Each potential candidate is then classified as lying inside the cone of gaze into set C_l :

$$C_l = \{\text{distance} : \text{distance} \leq \sin(\theta) \times \|p - H\|\} \quad \forall p \in W_L \quad (6.6)$$

Multiple candidates in C_l could be the potential intersection point but only one is required. Several plausible reduction strategies exist, the simplest is to pick point p with the smallest distance in set C_l . However, this has the potential effect of aliasing whereby a gaze vector *skimming* the actual intersection point to land on a much further location (i.e. a wall) which has a smaller distance. To tackle this potential problem, a further reduction strategy is explored where set C_l is sorted in ascending order and the top N candidates are selected. The point p with the shortest $\|p - H\|_2$ is chosen as

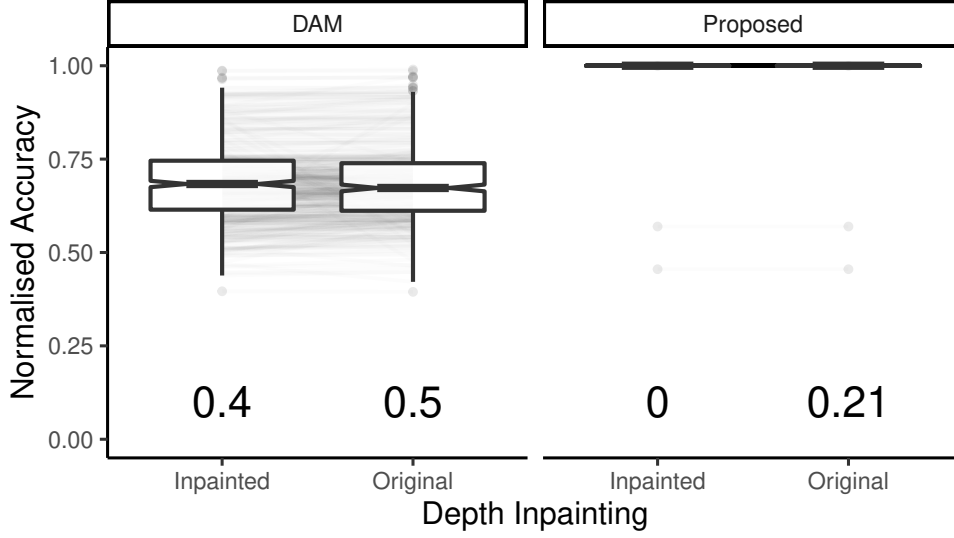


Figure 6.4: The effect of depth in-painting on the performance of the proposed [PoR](#) estimation algorithm. Depth in-painting has negligible effect on accuracy but does reduce the number of failures. Noise is set to 0 in this experiment.

the intersection point potentially alleviating the aliasing issue. The second strategy collapses into the first strategy when $N = 1$.

6.2 DATASET

The above algorithm is validated using depth images acquired from a Stereo-Labs ZED2 camera in an Adult Intensive Care Unit ([AICU](#)) setting at Chelsea and Westminster Hospital NHS Foundation Trust (HRA and REC Approval 20/LO/0162) ([Al-Hindawi, Vizcaychipi, and Demiris, 2021](#)). Dual Attention Guided Gaze Target Detection in the Wild ([DAM](#))’s estimated depth map was replaced with the actual depth map in the dataset. The resolution of the depth image is 1920×1080 . The dataset consists of 23 recordings of depth from a quasi-egocentric view each lasting 10 minutes. The depth image of first frame is taken from each recording in this experiment. Due to lacking ground truth data of the [PoR](#) of each participant, for each recording, we chose to simulate head pose H and gaze vector G .

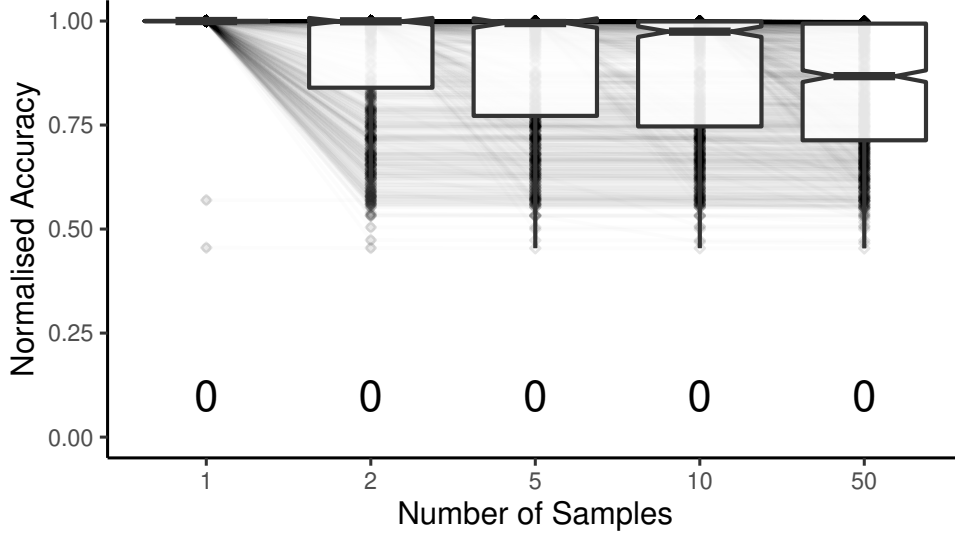


Figure 6.5: Ablation study for optimal number of closest samples for PoR estimation. The number of potential candidates to sample from set C_l relating to accuracy. The naive strategy of picking the voxel closest to the ray has the highest accuracy under experimental conditions whereas the sample strategy ($N > 1$) decreases accuracy.

Gaze target location in pixels was sampled randomly as $G_t \sim \mathcal{U}_{u,v}(0, L_{D_{width}} : L_{D_{height}})$ whilst the head-pose, located in world coordinates is sampled into pixel space by $H_{u,v} = K^{-1} \cdot [\sim \mathcal{N}(0, 0.2), \sim \mathcal{N}(1, 0.2), 1]^T$. These distributions are sufficiently wide enough to induce high random offsets between the gaze direction of the participant and the scene camera. Experimentally, increasing Gaussian noise ($\sim \mathcal{N}(0, 1) \times \text{diag}(L_D)$) is added to G_t to ascertain the reliability of the proposed algorithm against noise. Image locations where either algorithm failed were recorded. A random seed of 0 was used for reproducibility.

For each recording, 1,000 G_t gaze locations are sampled and processed to result in the proposed gaze-scene intersection point p . The Euclidean distance in pixels between G_t and p normalised by the maximal possible distance was taken as:

$$\text{accuracy} = 1 - \frac{\|G_t - p\|_2}{\sqrt{I_{D_w}^2 + I_{D_h}^2}} \quad (6.7)$$

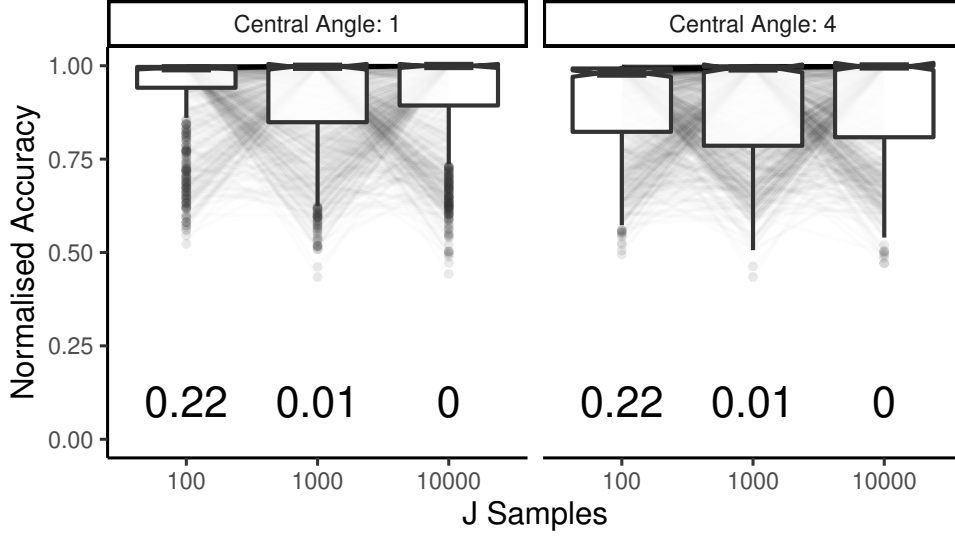


Figure 6.6: Ablation study for optimal number of samples of the planar-viewpoint generator for PoR estimation. The number of samples J tested against $\vec{HG}$ against normalised accuracy. As J increases, the number of successful queries increases.

where $\|*\|_2$ is the ℓ_2 norm operator of its argument vector $*$, I_{D_w} and I_{D_h} are the width and height of the image I_D . This accuracy term can then be used to evaluate our proposed algorithm.

Comparison interval on box plots were calculated for statistical significance by:

$$\text{Comparison Interval} = \text{median} \pm 1.57 \times \frac{\text{IQR}}{\sqrt{n}} \quad (6.8)$$

where IQR is the inter-quartile range defined by the 25th and 75th percentiles and n is the number of data points. Non-overlapping notches thus suggest non-parametric statistical significance (Chambers, 2017).

6.3 RESULTS

Figure 6.3 demonstrate the accuracy of the proposed algorithm at increasing noise levels. At low levels of noise, the proposed algorithm provides near-perfect accuracy in a statistically significant manner which naturally

converges to the accuracy achieved by (Fang et al., 2021) whilst also increasing the number of answered queries compared to DAM. The statistical significance is demonstrated by the non-intersection of Figure 6.3 median’s notch.

6.4 ABLATION STUDIES

In testing the accuracy of our proposed algorithm against depth images as acquired directly from the sensor and thus without hole in-painting, Figure 6.4 demonstrates that in-painting does not increase the overall accuracy but does reduce the number of failures.

The number of candidates considered following classification (N) was experimentally validated. Figure 6.5 demonstrates that as the number of samples (N) increases, the normalised accuracy decreases. It is worth noting that the median increase of accuracy results from points increasing from areas of low accuracy to high accuracy, rather than a general decrease of accuracy. This suggests the simplest strategy of the closest point to vector $\overrightarrow{HG}$ is the one that provides the highest accuracy and thus aliasing is not a likely scenario in this dataset.

The number of samples J dictates the number of intersections between $\overrightarrow{HG}$ and $p \forall p \in W_l$ and the size of set C_l which can be tuned for performance and accuracy. Figure 6.6 demonstrates the normalised accuracy achieved with increasing the number of J samples across two cone angles. A J value of 1000 is an optimal trade-off between performance and accuracy with only 1% of simulated runs not producing a result

6.5 CONCLUSION

We presented an algorithm for the estimation of PoR that is accurate and robust under a variety of difficult situations in a quasi-egocentric manner.

Extensive ablation studies demonstrate the importance of each sub-module while the architecture's hyper-parameters being derived from system requirements.

This chapter proposed a solution to a specific problem of [PoR](#) estimation using our unique eye-tracking platform. The architecture proposed in this chapter culminates all the requirements of an eye-tracking platform. In the next chapter, we outline a feasibility study that uses the entire platform for data acquisition from patients in [AICU](#).

FEASIBILITY STUDY DETAILS AND DATA PROCESSING

Following the development and validation of the eye tracking platform outlined in [Chapter 4](#) and the improvements made in [Chapters 5](#) and [6](#), appropriate regulatory approvals were sought for the deployment of the platform in Adult Intensive Care Unit ([AICU](#)). The aim was to conduct a feasibility trial looking at the role of eye movements for the diagnosis of delirium in [AICU](#). This chapter outlines the details of the study conducted, including how consent was handled given that delirium is a capacity losing state, inclusion and exclusion criteria, and patient characteristics of the patients recruited. This chapter also outlines how the data was handled prior to subsequent analysis. [Appendix B](#) provides all of the trial documents including the study's protocol and information sheets.

7.1 STUDY DETAILS

As the platform requires two cameras and a laptop, we sought to balance the cohort of patient's in the dataset to comprise of patients that are representative of the general [AICU](#) population but also avoid including a large number of patients who do not develop delirium. Thus, we performed study enrichment by including patients at moderately high risk of developing delirium through the calculation of each patient's Early PREdiction of DELIRium in ICu patients (E-PRE-DELIRIC) score – [Figure 7.1](#) demonstrates the study design and flow of each patient.

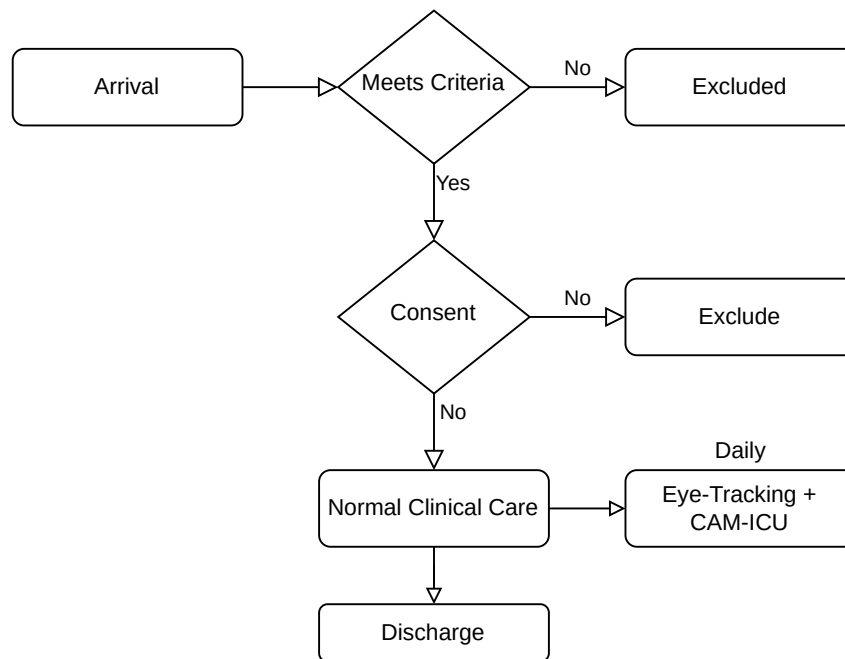


Figure 7.1: Flow of patients through our pilot feasibility study titled Continuous Non-Invasive Eye Tracking for the Diagnosis of Delirium ([CONFuSED](#)). Upon arrival, patients are checked for eligibility and if they meet the criteria, appropriate consent/advice is sought from the patient if capacitous or from their next of kin or a nominated consultee if not capacitous. Upon inclusion into the study, at least daily measurement of delirium through Confusion Assessment Method in ICU ([CAM-ICU](#)) is performed alongside measurement of the patient's gaze as per [Figure 7.2](#).

Following recruitment, patients underwent a daily assessments of delirium through [CAM-ICU](#) with concurrent measurements of eye movements for 10 minutes daily until discharge from [AICU](#) (Ely et al., 2001a).

7.1.1 Ethics Approval and Consent to Participate

The study protocol was approved by Health Research Authority ([HRA](#)) and Research & Ethics Committee ([REC](#)) (20/LO/0162) and was conducted in accordance with the Helsinki declaration. It was registered and adopted by the National Institute of Health Research ([NIHR](#)) facilitating the use of research nurses and registered on ClinicalTrials.gov (Al-Hindawi, Vizcaychipi, and Demiris, 2021).

Upon admission to [AICU](#), consent to enter this study was obtained in one of two ways; as the trial studied delirium, a capacity losing state, an assessment of capacity was performed and if present, fully informed consent was gained from the patient directly. Should the participant's mental state, competence and capacity mean that they were unable to provide consent, the patient's relatives/friends were approached for advice using the same procedure. In situations where a relative/friend was not available, a nominated consultee was sought which was one of the Intensivists who are not involved in the care of the patient at the time of advice. Should the patient recover their capacity, fully informed consent was undertaken to ensure that their participation in the study is in-line with their wishes.

Documents pertinent to the study, including the Protocol and relevant information sheets are presented in [Appendix B](#).

7.1.2 *Inclusion Criteria*

We aim to include patients that are representative of the overall patient population in [AICU](#), thus, with the above criteria for enrichment, our inclusion criteria is:

- Age ≥ 18 years
- [E-PRE-DELIRIC](#) score $\geq 20\%$
- Expected Length of Stay ≥ 2 days

7.1.3 *Exclusion Criteria*

Owing to the difficult diagnostic nature of delirium-on-dementia, we chose to exclude patients with a pre-existing diagnosis of dementia. We also aim to exclude patients in which the eye tracking platform would fail either due to

equipment reasons (failure to track) or due to patients reason where the gaze location does not represent their *covert* attention. Thus, our exclusion criteria is:

- Lack of Consent
- Pre-existing diagnosis of Dementia
- Significant Visual Impairment
- Non-concordant eyes
- Inability for facial recognition and eye tracking to be performed reliably

Figure 7.1 illustrates the flow of patients throughout the study.

7.1.4 Sample Size

Given that we are proposing a feasibility study, and owing to the exploratory nature of the study without prior effect size, a preliminary patient size was set to 30 as per (Lancaster, Dodd, and Williamson, 2004).

To avoid im-balanced data, we enriched our cohort by pre-selection avoiding patients at low-risk of delirium. To maximise external validity of the classification models, we used data from Chelsea & Westminster Hospital (CWH) for training whilst data from West Middlessex Hospital (WMH) was used for external validation. However, the entire dataset was used for the biologically inspired architecture as parameter optimisation was performed under simulation and external datasets.

Table 7.1: Patient Characteristic enrolled in the feasibility trial titled [CONfuSED](#). Pertinent variables related to outcomes of interests as well as variables that predict the development of delirium are stratified by data collection centre.

	Chelsea and Westminster	West Middlesex	Overall
Sessions	186 (71.0%)	76 (29.0%)	210
Patients	27 (77.1%)	8 (22.9%)	35
CAM-ICU			
Positive	75 (40.3%)	39 (51.3%)	114 (43.5%)
Negative	69 (37.1%)	27 (35.5%)	96 (36.6%)
Unable to measure	42 (22.6%)	10 (13.2%)	52 (19.8%)
Gender			
Female	8 (29.6%)	3 (37.5%)	11 (31.4%)
Male	19 (70.4%)	5 (62.5%)	24 (68.6%)
Age (Years)			
Mean (SD)	57.3 (19.3)	58.0 (14.7)	57.5 (18.1)
Frailty Score			
Mean (SD)	3.48 (1.05)	1.50 (0.535)	3.03 (1.27)
APACHE-II Probability (%)			
Mean (SD)	23.5 (14.1)	29.5 (20.7)	24.9 (15.7)
Urea (mmol/L)			
Mean (SD)	10.9 (7.08)	10.8 (5.96)	10.9 (6.76)
Corticosteroids			
No	26 (96.3%)	8 (100%)	34 (97.1%)
Yes	1 (3.7%)	0 (0%)	1 (2.9%)
Admission Type			
Elective	5 (18.5%)	0 (0%)	5 (14.3%)
Emergency	22 (81.5%)	8 (100%)	30 (85.7%)
ICU Length of Stay (Days)			
Mean (SD)	22.2 (25.9)	27.3 (24.1)	23.4 (25.2)
Unit Outcome			
Alive	23 (85.2%)	7 (87.5%)	30 (85.7%)
Dead	4 (14.8%)	1 (12.5%)	5 (14.3%)

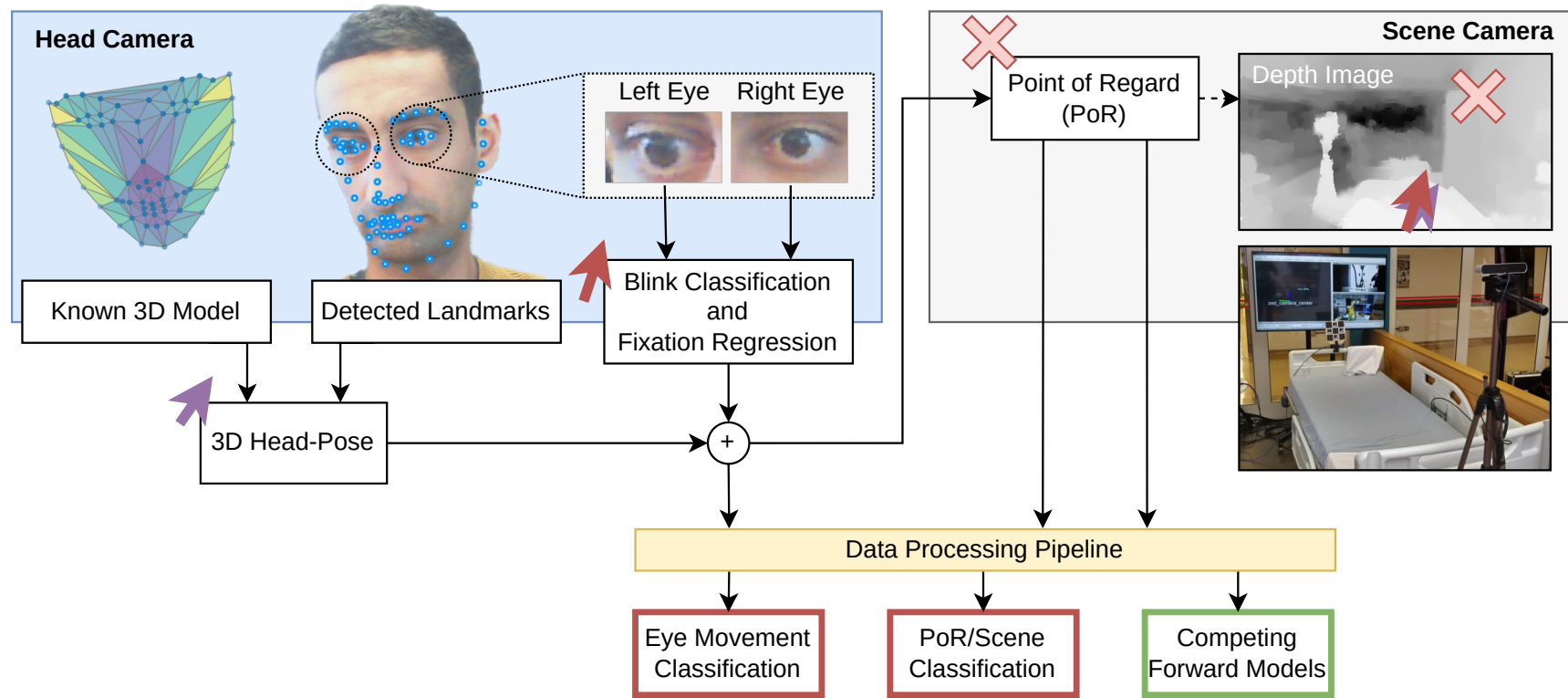


Figure 7.2: An outline of the eye tracking system used in CONfuSED alongside the origin of the data used in the each experiment.

7.2 DATA PROCESSING

The camera set-up was deployed across [CWH](#) and [WMH](#) between the periods of November 2020 to February 2022. The setup, as illustrated in [Figure 7.2](#), captures multiple biological signals in real-time asynchronously. Each biological signal contains a header of the time that that signal was acquired which can then be used to synchronise the data into a coherent dataset. The biological signals acquired were:

1. Facial landmarks. These are 68 points, in camera-space, located on the patient outlining their jaw, nose, eyes, and eye-brows. These are ultimately used for head-pose estimation, eye-angle estimation, and blink estimation. The latter two roles are performed following extraction of the two sub-images of the patient's eyes. These patches undergo de-rotation.
2. Head-pose. The extracted landmarks are then compared against a predefined 3D model of a generic head and the re-projection error between extracted landmarks and the 3D models. The result is the translation of the patient's head relative to the camera and the patient's heads rotation relative to the camera.
3. Eye angles. The eye patches extracted are then used to estimate the vertical and horizontal eye angles relative to the patient's head.
4. Blinking. The same eye patches are used to estimate the probability of whether the patient is blinking or not.

Each of the above biological signals, as well as the Red-Green-Blue-Depth ([RGBD](#)) data from the scene camera and the transformations tree, are each stored in a *message*, a data structure composed of a header which outlines the time the signals was acquired and the data itself. Each data-structure is typed, facilitating the correct reading of the data from storage.

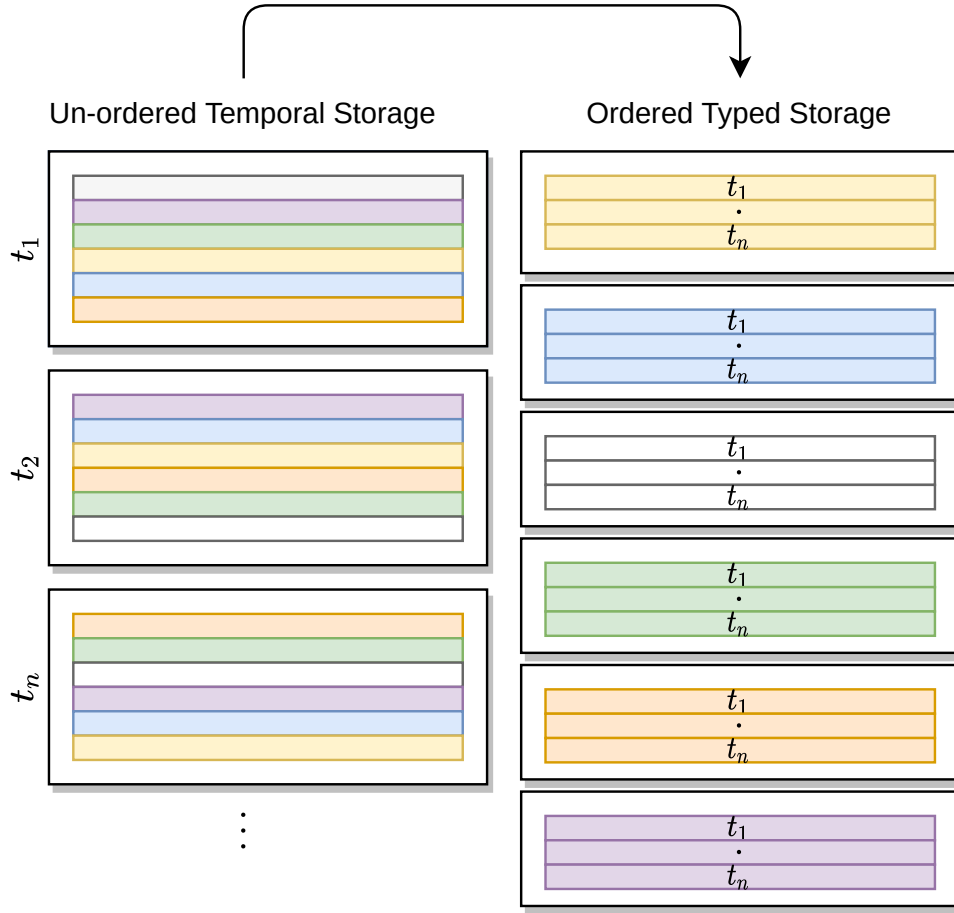


Figure 7.3: Data preparation performed in an offline stage following data acquisition. Raw un-ordered sequential messages are ordered into ordered temporal storage facilitating analysis.

Each of the acquired biological signals, as well as the [RGBD](#) data from the scene camera and the transformations tree, are each stored in a *message*, a data structure composed of a header which outlines the time the signals was acquired and the data itself. Each data-structure is typed, facilitating the correct reading of the data from storage.

However, non-ordered sequential data storage is not appropriate for machine learning. Thus, the an offline processing stage takes the raw data messages and coverts them to a format suitable for digestion, this is depicted in [Figure 7.3](#). The un-ordered sequential data is thus temporally ordered and validated.

7.3 FIXATION CLASSIFICATION

Following the conversion of the measured data into ordered temporal data, a post-processing stage performs data cleaning which removes gaze and scene data where the blink classifier deems that the patient was blinking as per [Chapter 5](#). Following this, gaze data is converted to fixations based on a dispersion filter. This filter calculates the pair-wise cosine distances across a set of gaze vectors and the inverse cosine of the maximum pair-wise distance is taken to the angle of that set of dispersion, formally:

$$g_v = \{g_1, g_2, \dots, g_t\} \quad (7.1)$$

$$dist = 1 - (|g_t| \circ |g_{t+1}|) \quad \forall (g_t, g_{t+1}) \in g_v \quad (7.2)$$

$$dispersion = \cos^{-1}(1 - \max(dist)) \quad (7.3)$$

where $\circ$ is the vector dot product and $|*|$ returns a normalised vector from its potentially non-normalised vector argument $*$. The result is a dispersion, measured in radians, of the set g_v . The set is composed of gaze vectors where the number of elements in the set comprise a duration of 600ms (Irwin, 1992; Kotval and Goldberg, 1998). If the dispersion of the set g_v is below a threshold, that set is deemed to be a fixation, this is depicted pictorially in [Figure 7.4](#). The dispersion threshold is set to the precision of the gaze tracker as per [Chapter 4](#).

Following fixation classification, the Point of Regard (PoR) is measured by the intersection of the fixation vector with the scene as as per [Chapter 6](#). This results in a complete dataset per recording session per participant. We use the High-performance data management and storage suite (HDF5) storage format for its performance, scalability, and provability upon deployment in wide-ranging tasks (Collette, 2013).

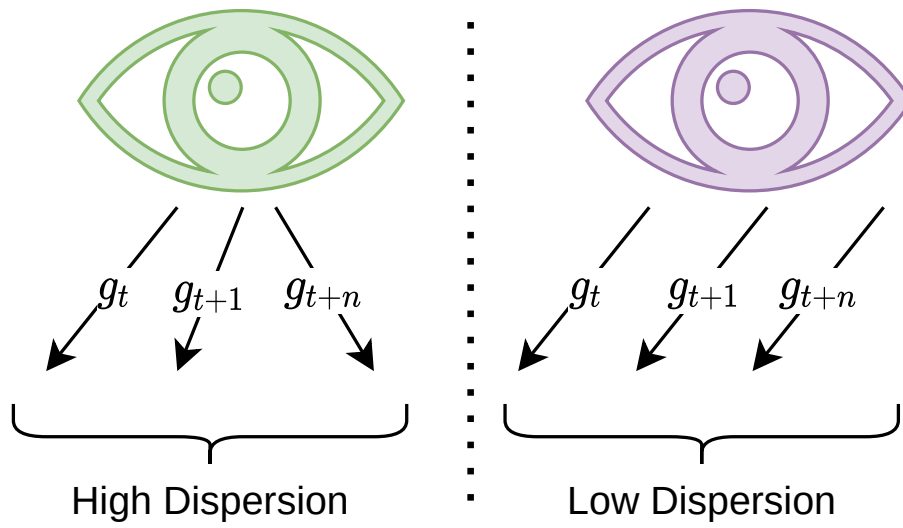


Figure 7.4: Two scenarios presented illustrating fixation classification. Consecutive gaze angles are placed in a set and the maximum angle between successive gaze vectors is taken as the dispersion. Areas of low dispersion are deemed to be fixations.

7.4 CONCLUSION

This chapter outlined the conduct of a feasibility study and the characteristics of the patients enrolled. It also outlined how data from the study was handled and stored.

The following chapters will use the data gathered and prepared in this chapter for further analysis.

DELIRIUM CLASSIFICATION USING EYE TRACKING DATA

This chapter aims to outline the development of a classifier that uses eye tracking data from our eye-tracking platform as previously outlined in [Chapter 4](#) and processed as per [Chapter 7](#).

As our data is large time-series in nature, with variable length, a causal classification technique that can attend to distant time points is required for optimal classification. Whilst both Long-Short Term Memory ([LSTM](#)) and Temporal Convolutional Network ([TCN](#)) can both handle arbitrarily long sequences, multi-layer [TCN](#) have been shown to be superior to [LSTM](#) for long sequences and thus form the basis of our classification models (Bai, Kolter, and Koltun, 2018).

The [TCN](#) acts as a time-series encoder outputting a fixed-length vector that represents the time series in multi-dimensional space. The [TCN](#) 's last hidden output thus forms an encoding of the entire time series. This encoding is then fed into a set of fully connected linear layers that output a logit. This logit is then converted to a probability using the softmax function, indicating whether the networks classify the patient's time series as being delirious or not.

We propose two models: the first model aims to use the temporal data of eye movements, i. e. horizontal and vertical eye angles relative to the patient's head, to predict delirium whilst the second model aims to use temporal Point of Regard ([PoR](#)) data, representing the gaze-scene intersection, as inputs to predict delirium. We present the models specifications, the training regime, and the performance of the models on the data collected from patients in

Adult Intensive Care Unit (AICU) as part of Continuous Non-Invasive Eye Tracking for the Diagnosis of Delirium (CONfuSED) (see Chapter 7).

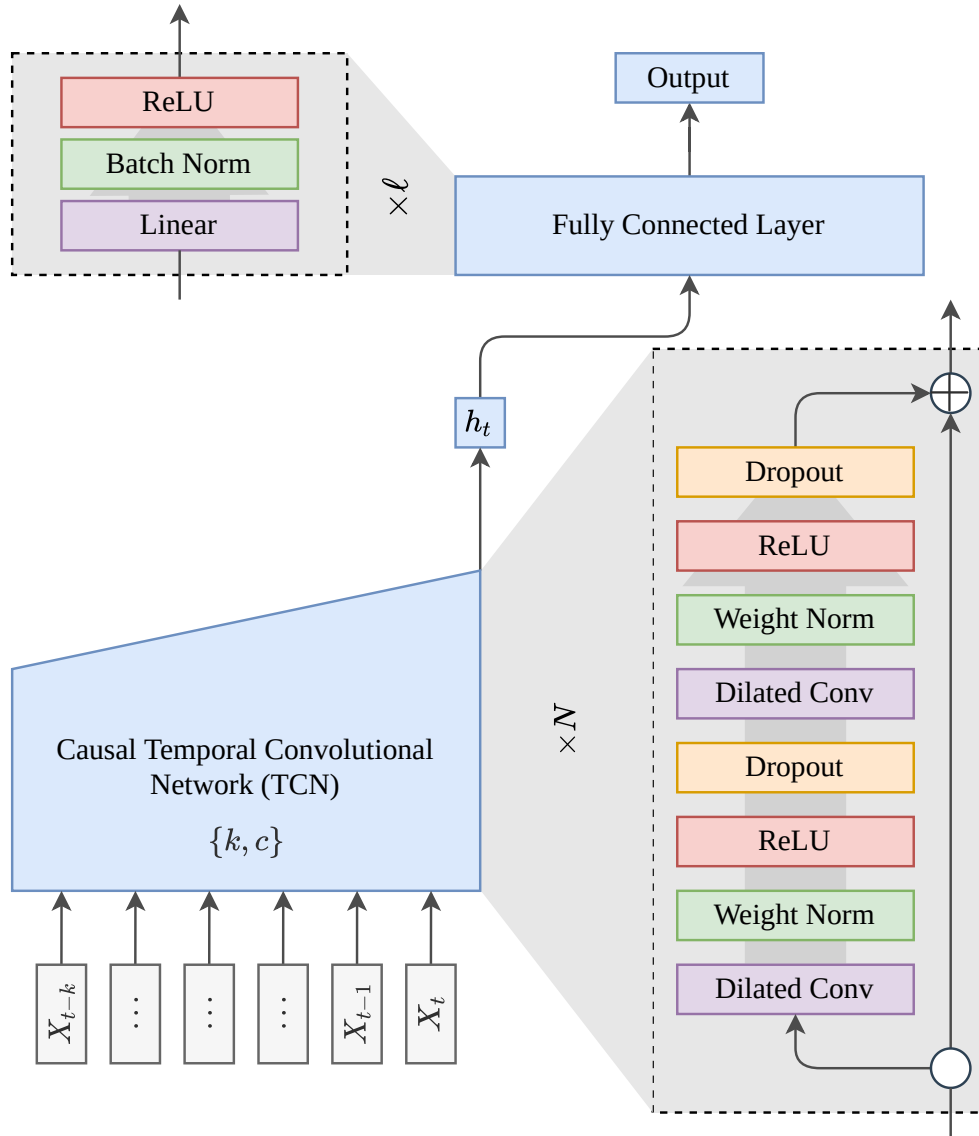


Figure 8.1: The architecture for delirium classification from our eye tracking platform. Time-series inputs are encoded to a fixed-sized vector using a TCN where a set of fully connected layers are used to convert the encoded vector to the probability of delirium. X_t represents the input at time point t . h_t represents the hidden latent fixed vector size that encodes previous time points. N and ℓ are the number of layers of the TCN and the fully connected layers respectively.

8.1 DELIRIUM CLASSIFICATION FROM EYE MOVEMENTS MODEL

The TCN for the task of predicting delirium from temporal eye movement data is composed of 8 temporal layers followed by 6 fully connected linear layers; Figure 8.1 illustrates the architecture. Each temporal block consists of a 1 dimensional convolutional filter, with a channel size of 256 and a kernel size of 7 followed by a non-linear activation function (ReLU) and dropout ($p = 0.05$) for regularisation. The output of the network is thus the probability of delirium given the labelled sequence of temporal eye-movements.

8.2 DELIRIUM CLASSIFICATION FROM POINT OF REGARD MODEL

Similar to the above model, this TCN aims to predict delirium from the temporal PoR data. The part of the scene that the patient is looking at, the PoR, was extracted as per Chapter 6. A crop of the scene, a 224×224 patch, was then encoded into a 1024 length vector using a ResNet-52 neural network pre-trained on ImageNet (He et al., 2015). Thus, the inputs into the TCN (X_t in Figure 8.1) are thus the fixed-length vectors.

The model is composed of an 8-layer TCN followed by 5 fully connected linear layers. Each temporal block consists of a 1 dimensional convolutional filter, with a channel size of 512 and a kernel size of 7 followed by a non-linear activation function (ReLU) and dropout ($p = 0.05$) for regularisation (Bai, Kolter, and Koltun, 2018).

8.3 TRAINING REGIME

AdamW optimiser was used with hyper-parameters $\beta_1 = 0.9$, $\beta_2 = 0.99$ and a learning rate of 1×10^{-3} (Loshchilov and Hutter, 2018). Nested cross-validation was used for hyper-parameter optimisation with early stopping for

further regularisation. Given that our data is balanced, non-weighted binary cross-entropy was used for the loss function.

A random search optimisation strategy was utilised to find the optimal set of hyper-parameters to maximise performance. A nested cross-validation scheme was used where the training dataset was split into 10 folds where 9 folds were used to train a model with specific hyper-parameters and the last fold was used for validation repeated 10 times and the results averaged (Browne, 2000). The hyper-parameter optimisation random search was bounded within the accepted range of each hyper-parameter. The tuned hyper-parameters were: TCN layers, channel size, kernel size, number of fully connected layers, learning rate, ℓ_2 regularisation, and dropout rate (Bergstra and Bengio, 2012; Gal and Ghahramani, 2016). The best performing model's parameters were stored and used for validation with the best-validated model used for testing.

The performance of the models are presented using Receiver Operator Characteristic (ROC) and Precision-Recall Curve (PRC) curves with their summary statistics of Area Under Receiver Operator Curve (AUROC) and mean Average Precision (mAP). Calibration is presented using the Brier score. Further details of these were presented in Section 2.6.

8.4 THRESHOLD SELECTION

To identify the threshold at which to dichotomise the probabilistic output of the models, we plotted a sweep of the threshold between $0 \rightarrow 1$ and calculated the F-1 score, the harmonic mean of the precision and recall at each threshold.

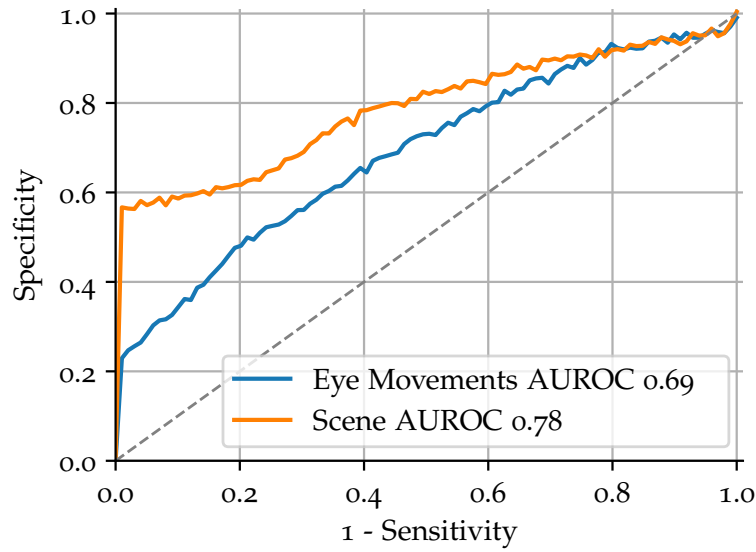


Figure 8.2: Receiver Operating Characteristic Curve as a performance indicator of the TCN for the binary classification task of diagnosing delirium. Two models are presented, an eye-movements only model, and a PoR model. The eye-movements model classifies delirium based on time-series eye movement data, namely the horizontal and vertical eye angles of each fixation with an AUROC of 0.69. Whereas the scene model classifies delirium from time-series PoR data to an AUROC of 0.78. Both classification models use the architecture in Figure 8.1.

8.5 RESULTS

A total of 169,941 fixations across 262 sessions between Chelsea & Westminster Hospital (CWH) and West Middelsex Hospital (WMH) were recorded. Data from CWH was used for training and hyper-parameter optimisation whereas data from WMH was used for external validation only. Results in the following sections are reported using validation data from WMH only.

8.5.1 Eye movement differ between patients with and without delirium

If visual attention is a good surrogate marker of cognitive attention, then eye movement behaviour would be different between patients who are delirious and those who are not. To test this hypothesis, we trained a model to clas-

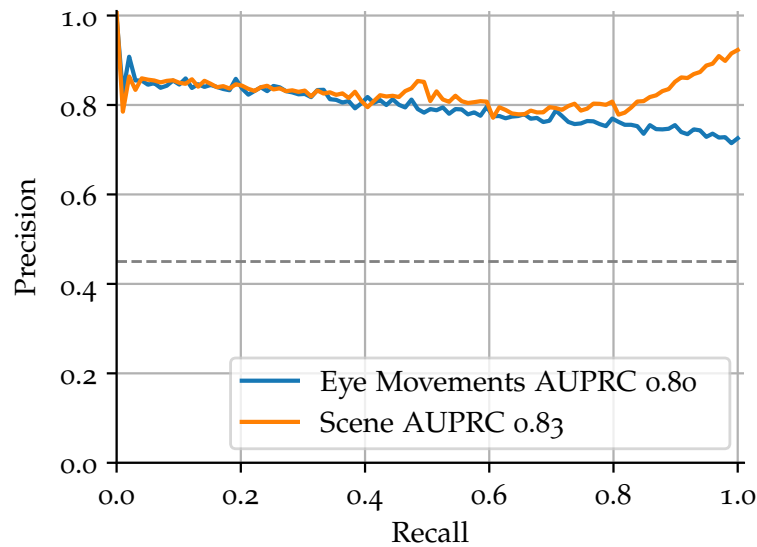


Figure 8.3: Precision-Recall Curve as a performance indicator of the TCN for the binary classification task of diagnosing delirium. Two models are presented, an eye-movements only model and a PoR model. The eye-movements model classifies delirium based on time-series eye movement data achieving an Area Under the Precision Recall Curve (AUPRC) of 0.80. Whereas the scene model, that contains time-series PoR data, achieves an AUPRC of 0.83. Both classification models use the architecture in Figure 8.1.

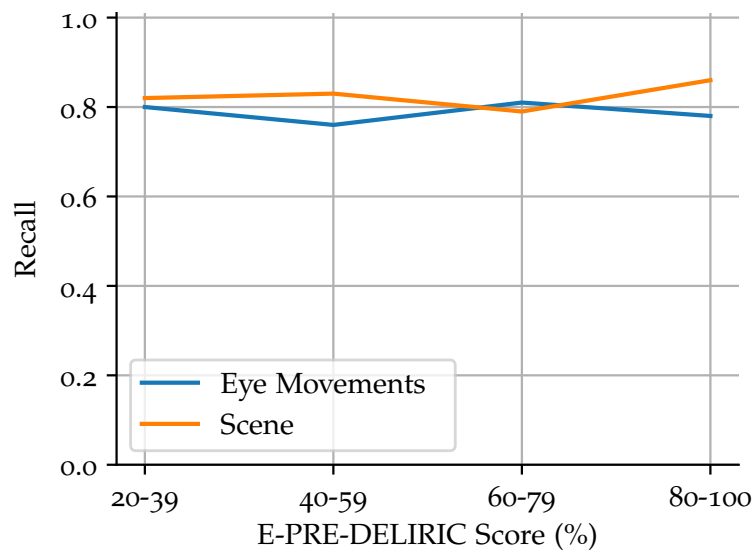


Figure 8.4: The proposed classifier's recall split by patient's Early PREdiction of DELIRium in ICu patients (E-PRE-DELIRIC) score. Both the eye movements model and the scene model demonstrate good re-call throughout the range of potential risk as calculated by E-PRE-DELIRIC. Probability threshold for recall calculation was set to maximise the F1-Score.

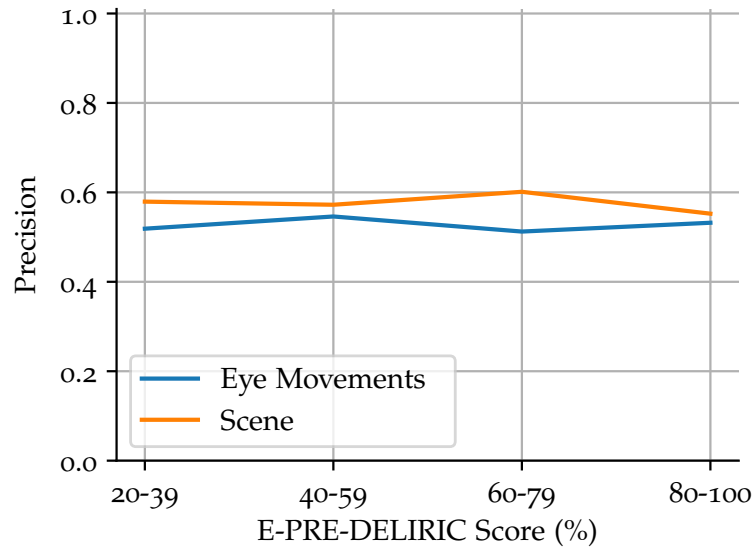


Figure 8.5: The proposed classifier's precision split by patient's *E-PRE-DELIRIC* score. Both the eye movements model and the scene model demonstrate adequate precision throughout the range of potential risk as calculated by *E-PRE-DELIRIC*. Probability threshold for precision calculation was set to maximise the F1-Score.

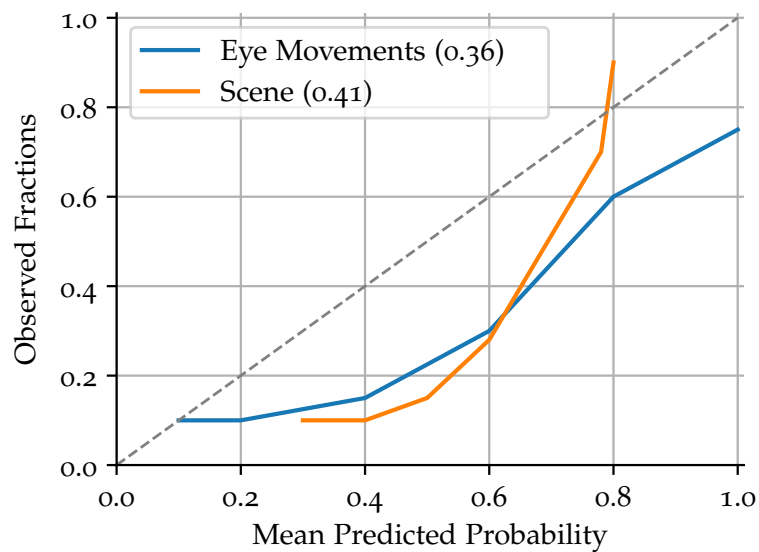


Figure 8.6: Normalised Calibration Plots of the mean predicted risk against the observed fractions for the two models tasked with delirium classification by eye movements or *PoR* time-series data. Numbers in brackets represent the Brier calibration score.

sify delirium purely based on the angles of the patient's eye fixations. The intuition is that, if a model, acting as a universal approximation function,

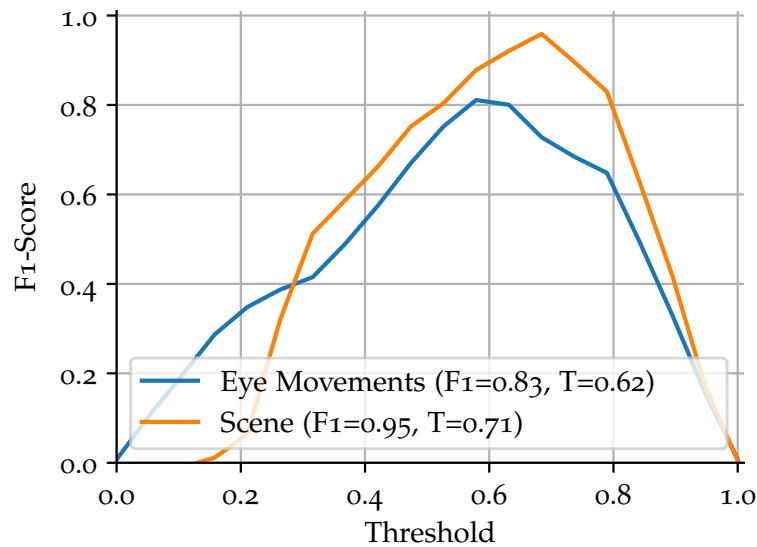


Figure 8.7: Threshold sweep against the F1 score at that threshold where the threshold at the maximum F1 Score is the optimal point to dichotomise the probabilistic output of the classification models. Numbers in brackets represents the F1 score and the thresholds at that F1-Score.

can identify temporal differences of eye movements through its classification validation accuracy, then, a difference exists in the eye movements of delirious and non-delirious patients. The blue plots in [Figures 8.2](#) and [8.3](#) demonstrates that the model achieves adequate discriminatory validation performance with an [AUROC](#) of 0.69 and a [mAP](#) of 0.80 indicating that eye-fixation-angle statistics differ between delirious and non-delirious patients.

8.5.2 *Point of Regard differs between patients with and without delirium*

Eye movement behaviour differences could either be a result of the intrinsic cerebral activity or a result of extrinsic visual information. To answer the question as to what is driving eye movements in delirious and non-delirious patients, we trained another model that uses scene information. The intuition here is that if eye movements are not reactive to the environment, then a model would not be able to classify delirium from scene information coming from a delirious episode or a non-delirious episode.

Thus, a crop of the scene that forms the PoR of the participant's fixation was extracted and a model was then trained on the time series of PoRs. This model demonstrates an increase in the discriminatory performance of the classifier to an AUROC of 0.78 and a mAP of 0.83, a 9% increase in the ROC and a 3% increase in the PRC curve when compared to the eye-fixation-angle classifier as per the orange plots in Figures 8.2 and 8.3. This suggests that the different eye movements of delirious and non-delirious recordings can be accounted for, at least in part, by extrinsic scene information.

Stratifying by predicted delirium risk, our classifier is stable across a variety of predicted delirium risk (Figures 8.4 and 8.5). The threshold to calculate the recall and precision in Figures 8.4 and 8.5 was chosen so as to maximise the F1-Score which was obtained in Figure 8.7. Figure 8.6 presents raw calibration plots without post-hoc calibration.

8.6 CONCLUSION

In this chapter, we have presented two models that classify delirium based on eye-movements and PoR. These time-series models are simple in architecture but provide robust classification performance on an external validation dataset.

Using TCNs, we conducted two experiments. The first study looked at eye movement fixation characteristics between delirium episodes and non-delirium episodes where a time-series classifier was able to dichotomise the data suggesting a significant difference. The second study answered the question as to whether those eye movement differences could be explained by scene information. If eye movements were intrinsic - i.e. originating from internal mechanisms without any external influence, then a classifier trained on scene information would not be able to discriminate between delirious and non-delirious episodes. This study found that scene information, through

the extraction of PoRs, increased classification accuracy suggesting that scene information contributes to the eye movements characteristics of delirium.

The significance of these classifier is of wide importance to the AICU and delirium clinicians. The classifiers, running on commercially available, low-cost hardware, demonstrate that visual attention is diagnostic of delirium. Specifically, the models, following appropriate regulatory approvals, could be readily deployed across AICU automating the diagnosis of delirium where they are envisaged to aid in the timely diagnosis and prompt treatment of delirium and its underlying causes. By utilising two cameras, working in unison around the bed-space, the patient is left un-instrumented ensuring clinical safety while the delirium status of a patient is continuously and non-invasively measured and recorded. Naturally, these classification models would require further evaluation in a prospective setting to aid in the widespread adoption of this platform. Architectural improvements including feature engineering are left for future work which is aimed at maximising classification performance.

Despite these successes, this finding requires further understanding - what scene information differs between session with delirium compares to ones without? This is discussed in the following chapter.

COMPETING FORWARD MODELS AS A SOLUTION TO THE INVERSE YARBUS PROBLEM

Whilst [Chapter 8](#) demonstrates that we can classify delirium based on eye-movements and scene-information, the models themselves provide no ability to explain their reasons for their outputs - specifically, the differences between the eye-fixation-angle characteristics of episodes of delirium compared to episodes without delirium. To achieve this explainability, we draw analogy between delirium and the Inverse Yabus problem. This chapter outlines what that problem is and how we aim to solve it using a biologically inspired approach.

9.1 THE INVERSE-YARBUS PROBLEM

The finding that eye movements are an *overt* manifestation of an internal *covert* cognitive task being undertaken was a crucial finding of Yabus's experiment, and has been termed 'top-down' attention (Yabus, [1967](#)). The inverse Yabus process, where the objective is to infer the internal cognitive task from eye movements has had significant scientific attention with mixed results. Work has called into question the validity of such an objective as a particular gaze-pattern could map to several internal cognitive tasks, none of which can be accurately measured (Greene, Liu, and Wolfe, [2012](#)).

However, success has been had in applying a wide range of computational algorithms in specialist domains and data sets. An ensemble of Support Vector Machines ([SVMs](#)) was used to classify task given gaze patterns and a Fisher Kernel Learning method was used to reinterpret the inverse Yabus

problem (Greene, Liu, and Wolfe, 2012; Kanan et al., 2014; Kumar et al., 2019). Hidden Markov Models (HMMs) were also developed for the task of probabilistic inference under a Bayesian framework (Haji-Abolhassani and Clark, 2014). Top-down causal network analysis was also carried out in an effort to study the influence of specific tasks on distinct gaze statistics with specific effort made to identify the most important eye-movement specific statistic (Boisvert and Bruce, 2016). Lastly, inference of potential targets given a specific search pattern was also attempted where given a-priori knowledge of the object to be search, it was shown that it is possible to infer gaze patterns specific to those objects. (Borji, Lennartz, and Pomplun, 2015).

Conversely to top-down attention, bottom-up gaze influences have also been explored. These are image features that draw gaze fixations. Computational models under free-viewing conditions, and thus without an internal covert task, were developed resulting in a computational model that models gaze patterns given the scene that the participant is looking at (Itti and Koch, 2001; Koch and Ullman, 1987). Early models took inspiration from biological findings of cerebral visual processing and output a saliency map of what the viewer is likely to look at given the presented scene in a probabilistic manner (Itti, Koch, and Niebur, 1998). Many other forms of saliency computational models have been developed, each with their interpretation of saliency (Borji and Itti, 2013).

Combining ‘top-down’ and ‘bottom-up’ attention has had little success. Bayesian frameworks have been employed to integrate the two mechanisms and while theoretical frameworks exist for the joint analysis, many stop short at identifying key ‘top-down’ mechanisms and offer a uniform prior (Zhang et al., 2008).

Focussing on eye-movement statistics, it has been demonstrated that participants exhibit different distributions of fixation amplitudes and directions under different tasks (Le Meur and Coutrot, 2016). Viewing natural scenes

when the participants are freely viewing the scene results in a specific saccade amplitude and direction while viewing a conversation between two people results in significant attention given to the persons' faces. These context-dependent fixation behaviour has been termed '*Saccade Biases*'. Multiple attempts have been made at modelling these biases with different data sets providing evidence for the best fit for Gaussian, Cauchy and Laplace distributions (Le Meur and Coutrot, 2016; Liu et al., 2013; Meur and Coutrot, 2016).

Building upon the theory of mind, and the computational architectures that built upon it, namely Hierarchical Attentive Multiple Models for Execution and Recognition (HAMMER), Hierarchical Modular Selection and Identification for Control (HMOSAIC), and prior work on biologically inspired saliency models, we propose an architecture that aims to infer the internal cognitive task given time series Point of Regard (PoR) data in a probabilistic manner using multiple competing forward models. (Demiris and Khadhour, 2006; Gallagher and Frith, 2003; Haruno, Wolpert, and Kawato, 2001, 2003; Le Meur and Liu, 2015). Forward models have been shown to be empirically present in the visual processing hierarchy of humans under experimental conditions (Mehta and Schaal, 2002). Our proposed architecture formulates multiple competing forward models that are parametrised by the current PoR of the participant and the scene which they are viewing. The models output a set of PoRs and their likelihoods according to in-built parameters that are unique to each forward model. Each forward model is then sampled and the Euclidean distance between the predicted PoR and the actual PoR is calculated. This results in a continuous measure of the *correctness* of that forward model. The distances are normalised and converted into probabilities. The forward model with the highest probability 'wins' the fixation under a Winner Takes All (WTA) scheme. We demonstrate the architecture's utility under simulated conditions where we tune the hyper-parameters of the models, followed by

the exploration of the architecture’s validation accuracy in a visual search task. We finally use the architecture on the data acquired as part of our feasibility trial (see [Chapter 7](#))

9.2 OVERVIEW OF ARCHITECTURE

Multiple competing forward models are built each using a pipeline composed of *saliency*, *saccade biases*, and *inhibition of return* modules where the output of one module is the input into the next. The first module is a saliency model that, given a raw image, would output a probabilistic map of where fixations are likely to result given an intrinsic definition of saliency. Coupled to each saliency module is a simulation of eye gaze movement which is parametrised by the current [PoR](#) and consists of two sub-modules. The first sub-module, termed *saccade biases*, is a probability distribution that jointly encompasses the amplitude and direction of the next fixation given the location of the current fixation, which are often task-specific. The second module, *inhibition of return*, captures the statistical probability that the current fixation inhibits future fixations due to visual processing of the scene at that fixation. These two sub-modules constrains the saliency model into an area of the scene that the next fixation could plausibly reach. The result of each forward model module is thus a list of possible fixations and their respective likelihood.

Given the list of fixations and their respective likelihood per forward model, a Monte-Carlo sampling module outputs the *correctness* of the forward model given the next gaze measurement from the participant PoR_{t+1} . The *correctness* of the forward model is thus the average distance the simulation’s predictions are from the next fixation point PoR_{t+1} . However, due to the errors of measurement of the fixation point, simply outputting the density of the forward model under fixation PoR_{t+1} would result in erroneous measurements. To

alleviate this, the Euclidean distance between fixation PoR_{t+1} and the sample from the set of simulated fixations is used instead of the Euclidean distance.

The result is a set of scores for each forward model with the forward model that is most in-line with fixation PoR_{t+1} having the highest score. These scores are then converted to probabilities using a modified soft-max function.

9.2.1 *Saliency Models*

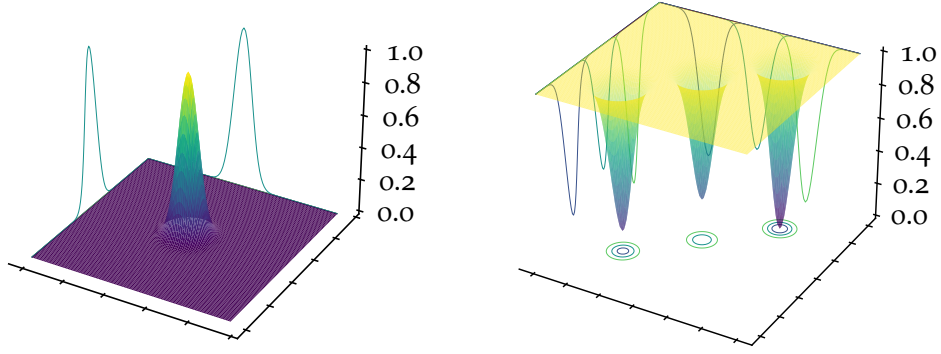
One of the inputs into the forward model, part of the state, is the scene image. This is used to calculate saliency, which is defined as a two-dimensional discrete probability mass function that places density in locations where fixations are expected. The mapping between the image to probability mass function is heterogeneous in definition with multiple saliency models defining the mapping according to different metrics. For this architecture, no restriction on the metrics is placed but a selection of saliency models should be coupled to potential tasks that the participants would be undertaking and highlight potential areas of interest in the scene.

9.2.2 *Saccade Biases*

Rather than utilise pre-existing best fit distributions that aims to cover the different tasks, we choose to not restrict the distribution of saccade biases and thus use the least informative distribution – a Gaussian distribution centred at the fixation with a 2° full-width-half-maximum.

9.2.3 *Inhibition of Return*

Once a participant has landed a fixation on a location, they are not likely to look at that same location again shortly. This phenomenon has been termed



(a) Saccade Bias demonstrating probability of 0 in all image location except at fixation where the peak of 1.0 transmits the saliency of that point on-wards. (b) Inhibition of Return module demonstrating three fixations under decay. Probability of fixation at all locations is 1.0 except where fixations have occurred in the past.

Figure 9.1: Eye-movement specific modules used in the forward models as a solution to the inverse Yarbush problem. Saccade Bias modelled in Figure 9.1a and Inhibition of Return in Figure 9.1b. Saccade Biases are modelled as a distribution with a baseline probability of 0.0 that extends to 1.0 within the standard deviation set by the field of view. Inhibition of Return is modelled as a cumulative distribution where the baseline inhibition is 1.0 indicating that no inhibition takes place following multiplication. At each fixation, a Gaussian is placed that inhibits return to that location where its probability decays back to 0.0. Three fixations are demonstrated in various states of decay.

inhibition of return and can be modelled by a decaying inverse Gaussian function over previous fixations (Le Meur and Liu, 2015), formally:

$$ior(x, t) = \begin{cases} 1, & t = 0, \\ \left[ior(x, t-1) - \Phi(\|y - x_t\|_2) + \sum_{k=1}^{\xi} \frac{\Phi(\|y - x_{t-k}\|_2)}{T} \right] & \text{otherwise, } \forall y \in \Omega \end{cases} \quad (9.1)$$

where $y \in \Omega$ represents all locations in image Ω . x is the gaze point at time t , Φ represents a Gaussian distribution parametrised by a standard deviation set to 2° reflecting the width of the human fovea. $\|\cdot\|_2$ is the Euclidean distance, $[\cdot]$ clips the values between $[0, 1]$. ξ is the number of fixations to consider and is equal to $\min(T, t-1)$ where T is nominally set to 8 (Le Meur and Liu, 2015).

9.2.4 Prediction Error

Given the list of fixations and their probabilities per forward model, a Monte-Carlo sampling module outputs the *correctness* of the forward model given the next gaze measurement from the participant PoR_{t+1} . The *correctness* of the simulation encompasses the average distance the simulation's predictions are from the next fixation point PoR_{t+1} . However, due to the errors of measurement of the fixation point, simply outputting the density of the forward model under fixation PoR_{t+1} would result in erroneous measurements. To alleviate this, the Euclidean distance between fixation PoR_{t+1} and the sample from the set of simulated fixations is used. A single forward model can be represented mathematically as:

$$p_{GM} = p(S) * p(TOSM) * p(IOR) \quad (9.2)$$

where $p(S)$, $p(TOSM)$, $p(IOR)$ are the probability mass functions for saliency, saccade biases and inhibition of return respectively, and p_{GM} represents probability mass function resulting from forward model GM . p_{GM} can then be sampled using a Monte-Carlo sampling scheme and the expectation of that sample is then:

$$d(PoR_{t+1})_{GM} = \mathbb{E}_{n \sim p_{GM}} \left[1 - \frac{\|n - PoR_{t+1}\|_2}{\sqrt{\Omega_w^2 + \Omega_h^2}} \right] \quad (9.3)$$

where $\mathbb{E}$ is the expectation operator that samples n from set p_{GM} for forward model GM . $\| * \|_2$ is the Euclidean distance operator of the argument $*$. To normalise these measurements, the distances are normalised to the maximum possible value (the diagonal of the scene image Ω). As smaller distances are

more inline with measurement, and to covert these into probabilities, 1 minus the quantity is used as the input into a softmax function:

$$\sigma(d_h) = \frac{e^{-\beta d_h}}{\sum_{j=1}^K e^{-\beta d_j}} \quad (9.4)$$

This results in a normalised distribution where the forward model that is most in-line with fixation PoR_{t+1} has the highest probability. Parameter β dictates the sharpness of the softmax function. A grid search strategy was used to find a value for β that maximises separation while minimising uncertainty. This is advantageous in a situation where a gaze point is between two forward model predictions and the expected result is a stable softmax result rather than one that oscillates.

9.2.5 *Competition*

Given that each forward model is imbuing a simulation of an expected external observer, the forward model with the highest probability, and thus the lowest error, is awarded the fixation in a Winner Takes All (WTA) strategy. This results in counts of each forward model which can then be compared across forward models.

9.3 VALIDATION

9.3.1 *Experiments in Simulation*

To test the accuracy of the architecture at classification, a 480×270 pixel image composed of a red and blue circle aligned horizontally with a radius of $50px$ was generated as illustrated in [Figure 9.2](#). Two saliency images

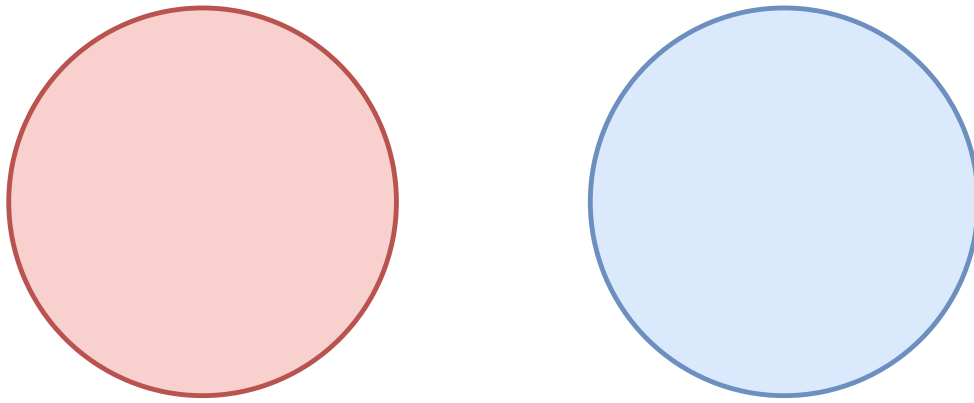


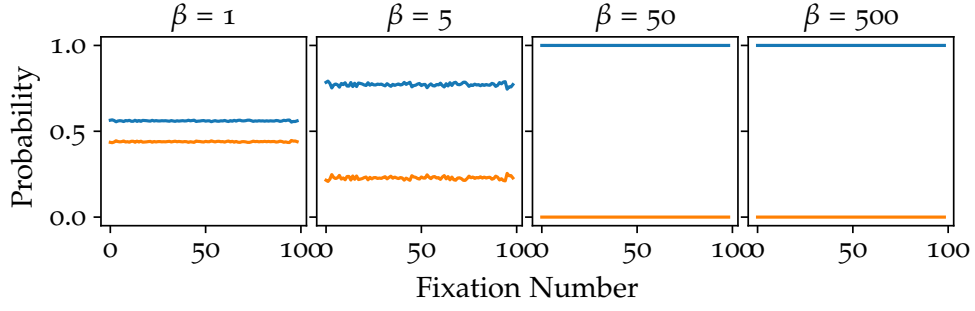
Figure 9.2: Illustration of the scene of the simulated environment under for validation of the architecture using synthetic data. Two experiments were performed, the first is a *constant* task where fixations originated from the red circle, and a *switch* task where fixation started in the red circle and then switched to the blue.

were calculated from that image, one places all support under the red circle whereas the other places it under the blue circle.

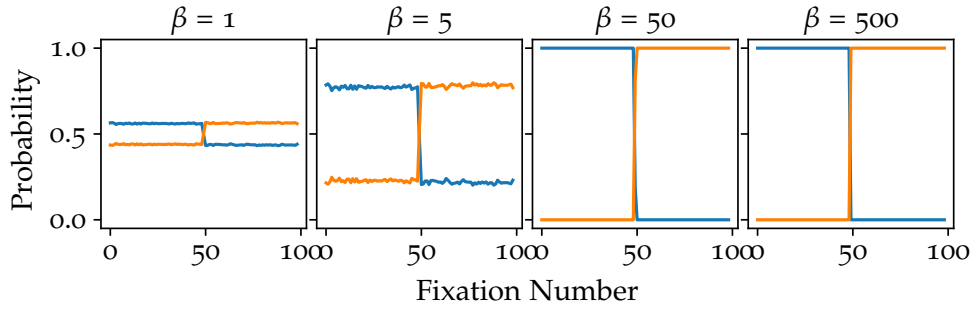
Fixation data was then simulated to originate from the centre of one of the circles; in the first experiment, the gaze was simulated to originate from the red circle constantly while in a second experiment, the gaze was set to originate from the red circle and then switch to the blue circle. For each experiment, a noise term to simulate gaze measurement error. Field of view is set to equal the field of view of a human's; 210° horizontally and 150° vertically.

Saccade biases were modelled empirically as a Gaussian distribution with a full width at half maximum of 2° . Inhibition of return is modelled as per [Section 9.2.3](#).

[Figure 9.3](#) demonstrates the results. The *constant* task in [Figure 9.3a](#) is reliably and consistently classified with high probability. The second task, [Figure 9.3b](#), where the gaze would *switch* between two forward models is also detected with a sharp and immediate switch in probabilities. The *ripple* in probabilities is due to the added noise term to the gaze. This experiment



(a) Results for the biologically inspired architecture as a solution to the inverse Yarbus problem experiments. In this experiment, fixations are simulated to originate from the red circle for 100 fixations.



(b) Results for the biologically inspired architecture as a solution to the inverse Yarbus problem experiments. In this experiment, fixations are simulated to originate from the red circle for 50 fixations and then switch to the blue circle for the remaining 50 fixations.

Figure 9.3: Simulation results for the architecture as a solution to the inverse Yarbus problem. Two eye movement strategies are presented, a fixed strategy, and a switch strategy. Various β parameters are explored aiming to maximise the range of the softmax operator in Equation 9.4. Small values of β result in encroachment of the two simulations' probabilities while larger values maximise classification distance but lose sensitivity to small changes.

also served as a method to grid-search a suitable value for β ; a value of 50 was chosen for all subsequent experiments.

9.3.2 Visual Search Dataset

Zhang et al. 's work on zero-shot invariant visual search produced a dataset of fixations from participants looking for a target – a cartoon character that is often hidden in a complex scene (Zhang et al., 2018). We use the *oracle* subset of this dataset which truncates the list of fixations across subjects such that the last fixation is the one closest to the target. This serves as the ground truth

– the last fixation should be near target in sufficient proximity for the subjects to declare that they have located the target. The objective of this experiment is to ascertain the performance of the architecture in correctly classifying the fixations on the target during a visual search task where multiple switches between forward models occur.

Five saliency models were utilised that aim to explain the fixation strategies:

1. The ground truth bounding box around the target – this represents the object in the scene that the participants are searching for.
2. A synthetic *Not-Target* dataset that is the binary opposite of the ground truth supplied – this would represent *any* other search target for all other objects
3. Itti’s saliency model aims to explain fixations in free-viewing in a biologically plausible manner (Itti, Koch, and Niebur, 1998)
4. Spectral Residual Approach saliency model that uses aggregate image data from an information gain perspective to highlight proto-objects (Hou and Zhang, 2007).
5. Centre Bias model that places a two dimensional Gaussian distribution on the centre – this model is used to demonstrate a control.

Rather than placing a restrictive prior on the task through modelling specific saccade biases, we have chosen to model the saccade biases using a non-informative Gaussian prior. This places the least restriction on the possible eye movements that could occur. Inhibition of return is modelled as per [Section 9.2.3](#).

Figure 9.5 demonstrates the fixation and the outputs of the architecture for all participants in the dataset viewing the same image. As the participants search for the target, the score of each forward models change, all culminating in the score of the target model being the highest – in line with the ground

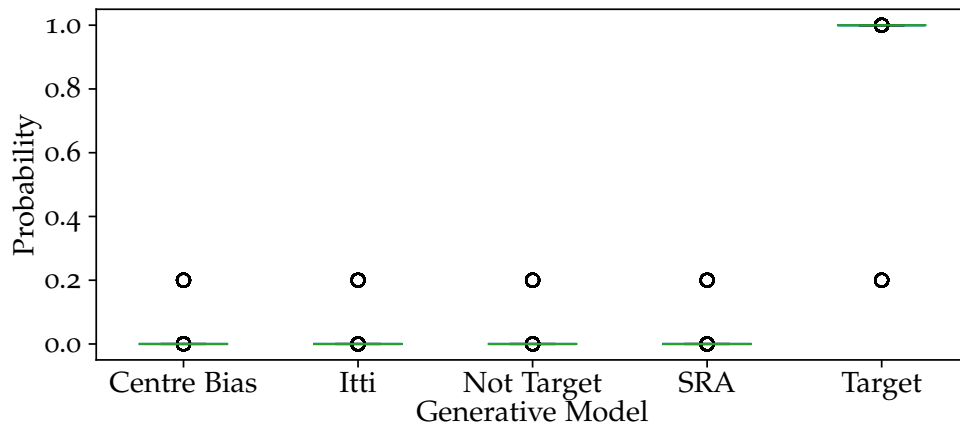


Figure 9.4: A box-plot of the last fixation scores across all subjects across all images. Outliers presented with an open circle. Last fixation of the Target is correctly classified across all images and all subjects leading to the collapse of the box-plot into a single line.

truth. Figure 9.4 demonstrates the probabilities of the last fixation where the expectation is that the target saliency model is the highest as the ground truth dictates that this should be the case.

9.3.3 Saliency Models for Delirium Understanding

Five saliency models were used, namely Itti, Spectral Residual Approach for Saliency Detection (SRA), an Objects map, a Persons map, and finally a Motion map. These correspond to a pyramidal increase in visual processing complexity. Specifically, Itti's saliency map was chosen as it proposes a neurologically inspired approach that aims to explain which features in the images draw the attention and has been found to map to a similar neuronal representation in the primate superior colliculus (Itti, Koch, and Niebur, 1998; White et al., 2017). Increasing in visual processing complexity, SRA was chosen as it focuses on frequency deviations in spectral space and is thus a purely image-centric approach and has been shown to encompass proto-objects (Hou and Zhang, 2007). Lastly, FasterRCNN is a deep-neural network that detects objects and their corresponding bounding boxes in the

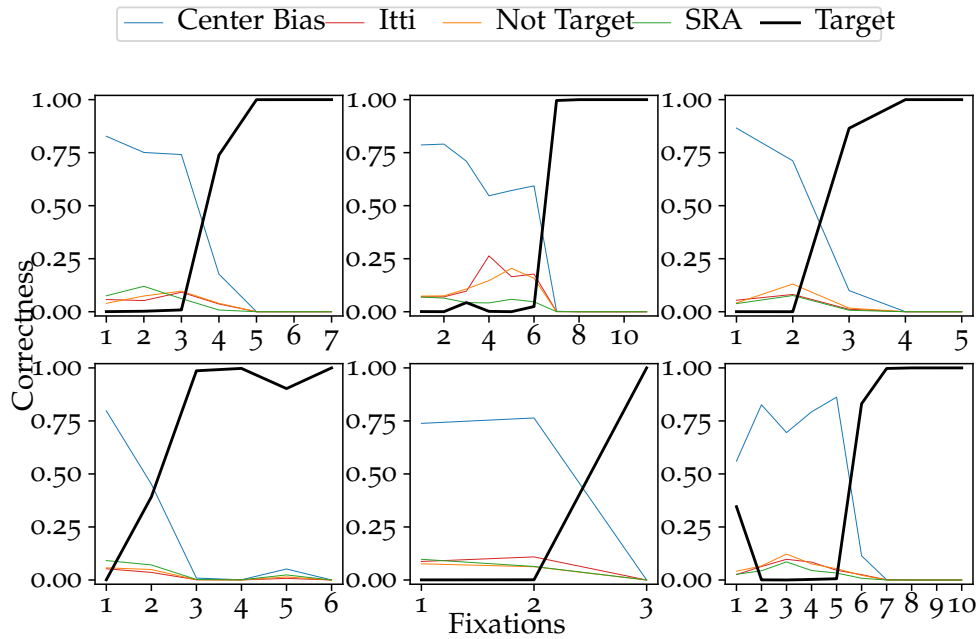


Figure 9.5: Forward model probabilities for the subjects viewing the same image. All fixations at the end of the search land on the target where the forward model for the target correctly classifies the viewing patterns.

scene and thus represents the most complicated visual-processing saliency map (Ren et al., 2015). The object detection saliency map consists of using FasterRCNN as a backbone of object detection, where, object detections are then rendered as a Gaussian blob centred on the centre of the object with a full-width-half-maximum of the diagonal of the object's bounding box.

The final two saliency maps detected are persons and motion respectively as those are known to have specific neuronal encodings.

Examples of saliency alongside original images are presented in Figure 9.6.

9.4 FORWARD MODELS EXPLAIN UNDERLYING ATTENTION MECHANISMS IN DELIRIUM

This section serves to outline how the proposed architecture was used for the analysis of the data recorded as per Chapter 7. We explore the results and discuss how the results fit with pre-existing work in delirium.

As per [Chapter 4](#), and as outlined in [Figure 9.6](#), images from the scene camera were used to calculate saliency maps. Five saliency maps were chosen based on a hierarchy of increasing visual processing complexity – low-level saliency that which places density on lines, edges, and areas of high contrast; mid-level saliency which places density on shapes and proto-objects; and high-



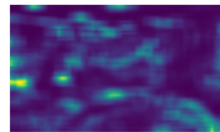
(a) Demonstration of the setup around the bed-space, the scene camera located behind the bed-space and the head-camera is facing the demonstrator. Saliency models are calculated from the scene-camera images.



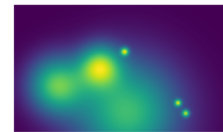
(b) View from the Scene Camera



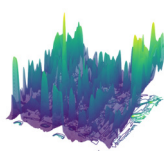
(c) Itti saliency



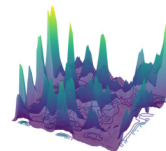
(e) SRA saliency



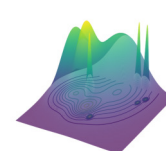
(g) Object saliency



(d) Itti saliency presented as a probability mass function demonstrating high frequency details.



(f) SRA saliency as a probability mass function demonstrating medium frequency details.



(h) Object saliency as a probability mass function demonstrating low frequency details.

Figure 9.6: Original scene image and corresponding saliency maps as calculated by Itti, SRA, and Object detection respectively. Note that the scene images are blurred for privacy but saliency calculations occur at full resolution without blurring. Note the finer saliency details of Itti's saliency model compared to SRA and Object respectively. SRA saliency produces proto-images, shapes that can be objects or part of objects whilst Object saliency produces only objects.

level saliency which places density on complete objects (Hou and Zhang, 2007; Itti, Koch, and Niebur, 1998; Ren et al., 2015). The final two saliency maps detected persons and motion representing further higher-level functional processing.

Each saliency map is then element-wise multiplied by the saccade bias parametrised by the patient's current fixation. The saccade biases are task-dependent and form part of the top-down signal (Meur and Coutrot, 2016). To avoid placing restrictive priors on the expected task that patients might be performing per recording, a non-informative Gaussian prior was chosen that has its centre at the fixation point with a full-width-half-maximum set to 2° as demonstrated in Figure 9.1a. Lastly, to encourage scene exploration, an inhibition of return module (Figure 9.1b) is used to avoid the exploration of a scene that has already been viewed recently. The final result of the pipeline is a set of possible fixations and their probabilities encoded as a two-dimensional map.

A Monte-Carlo scheme samples from the predicted fixation map and calculates the Euclidean distance between the sampled point and the patient's fixation and the expectation over those samples is then used as the accuracy of the forward models' prediction. Thus, the forward model with the highest score is most in-line with the current viewing behaviour of the participant; the pipeline is presented in Figure 9.7.

For each fixation of each recording session per participant, the forward model's probability is tallied in a WTA scheme. A table stratified by the label of the recording session is then created. Null hypothesis testing using the χ^2 test was used where the null hypothesis is that the expected frequencies of fixations do not differ between delirious and non-delirious patients with significance set to $P < 0.05$.

Table 9.1 demonstrates the results of applying the forward models on our dataset. Stratifying by delirious and non delirious episodes, all visual features

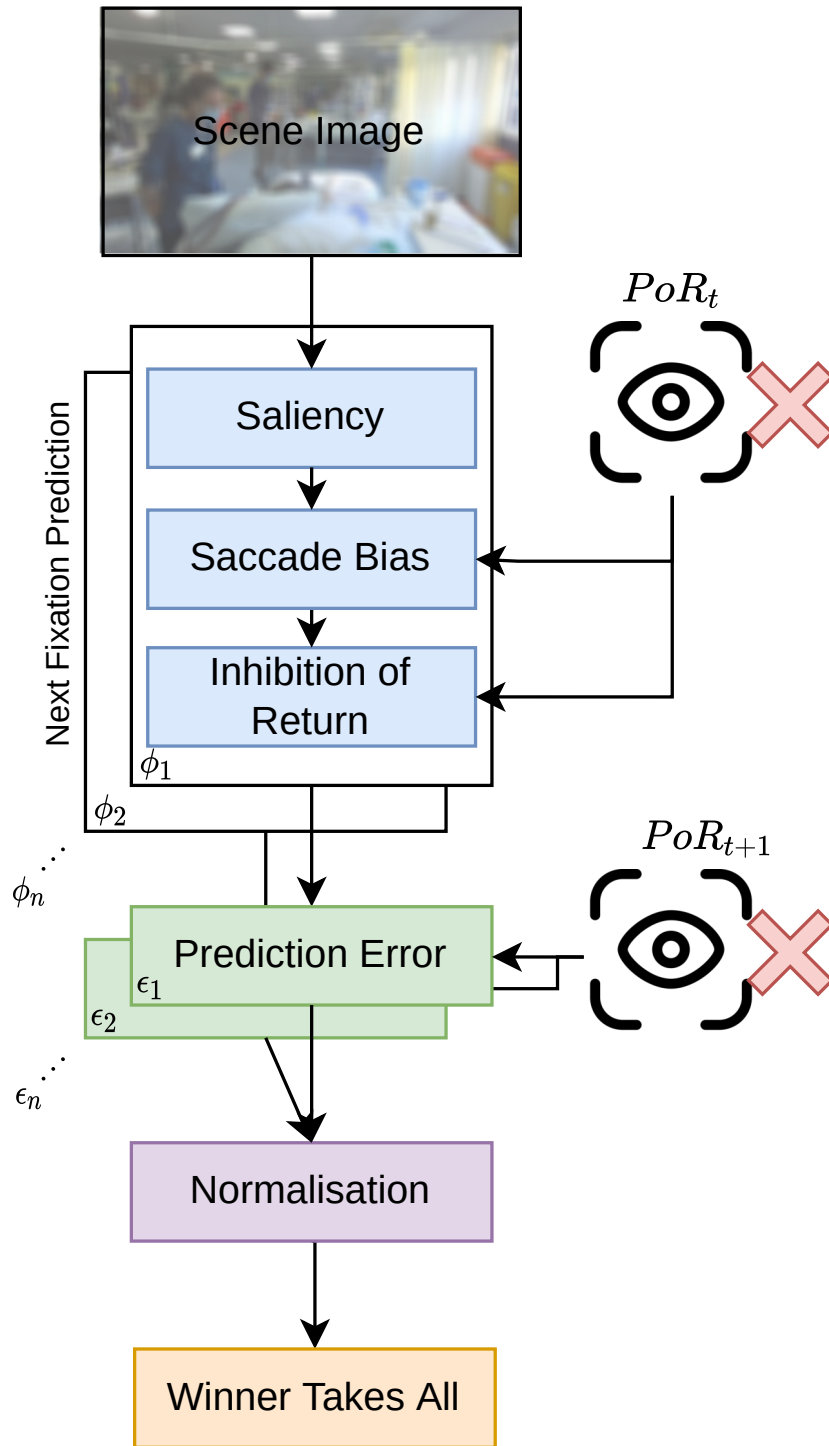


Figure 9.7: A schematic of the proposed architecture as a solution to the inverse Yarbus problem. Multiple forward models are stacked, each containing a saliency, saccade bias, and inhibition of return modules which are parametrised by the current fixation at time t . Each forward model's correctness is then measured at the next fixation at time $t + 1$. The prediction error is then normalised using a modified soft-max as per Equation (9.4). Following normalisation, the forward model with the highest probability is deemed the winner and allocated a point.

received fewer fixations in the delirium patients compared to non-delirium patients in a statistically significant manner. Testing across forward models results in $\chi^2 = 9.80$ ($p = 4.40 \times 10^{-2}$).

Following on from [Chapter 8](#) where we demonstrated that delirious and non-delirious episodes have different eye-movement characteristics driven, at least in-part, by extrinsic scene information, the question remained as to *what* visual differences exist between delirious and non-delirious episodes. In order to understand this, we developed a novel architecture that simulates future fixations given the current scene and fixation. The forward models demonstrated that delirious episodes are punctuated by visual *in-attention* to many parts of the visual processing hierarchy. This finding is in keeping with Electroencephalogram (EEG) findings of loss of high frequency power and predominant low frequency (theta and delta bands) power in patients with delirium (Koponen et al., 1989a; Numan et al., 2017; van der Kooi et al., 2015). Our findings are also consistent with the network disconnectivity hypothesis of delirium where, previously connected networks become disconnected thus limiting the transfer of information between areas of the brain (Maldonado, 2013). Our results are also consistent with pre-existing Magnetic Resonance Image (MRI) and Computed Tomography (CT)-Angiography literature on delirium where occipital, parietal, and temporal lobe pathologies exist in keeping with pathology in the visual processing hierarchy (Fong et al., 2006; Koponen et al., 1989b; Yokota et al., 2003). Other radiological findings of periventricular and deep white matter hyperintensities, and cortical atrophy have not been accounted for by our results as our forward models were restricted to the visual processing hierarchy (Figiel et al., 1989).

Table 9.1: The tallied fixations for each forward model stratified by delirium and non-delirium episodes. Statistical significance testing using the χ^2 test between the tallied probabilities of delirium vs non-delirium with the null hypothesis being of no difference.

Forward Model	Delirium Fixations	No Delirium Fixations	χ^2	df	P value
Low-Level	11282	14014	294	1	$< 2.2 \times 10^{-16}$
Mid-Level	11282	14011	295	1	$< 2.2 \times 10^{-16}$
High-Level	13729	16441	244	1	$< 2.2 \times 10^{-16}$
Persons	12596	15126	231	1	$< 2.2 \times 10^{-16}$
Motion	12531	15532	321	1	$< 2.2 \times 10^{-16}$
Total	61453	75162	9.80	4	$= 4.40 \times 10^{-2}$

9.5 CONCLUSION

This chapter focused on our proposed architecture for the inverse Yarbus problem. By generating multiple hypotheses with saliency maps specified *a-priori*, we can then simulate future fixations and then score the accuracy of each forward model's *correctness*. The forward model with the highest score is most in-line with the current cognitive processes.

We validated this architecture in under simulation where we also found optimal hyper-parameters. The hyper-parameters are then used for further validation experiments from a visual-search dataset where the proposed architecture correctly identified the task in all participants with high accuracy. These experiments validated the architecture's prowess for the understanding the cognitive processes in patients with delirium.

The forward models were then used to analyse the eye movement data gathered and where the results demonstrate that delirium episodes are punctuated by global visual in-attention to low, mid, and high level visual features, but also to persons and motion.

CONCLUSIONS AND FUTURE WORK

In this chapter, we will summarise our contributions, discuss the limitations of our findings and provide direction for future work.

10.1 OVERVIEW AND CONTRIBUTIONS OF THE THESIS

In this thesis, we outlined the clinical problem of delirium in Adult Intensive Care Unit (AICU), the ramifications related to the development of delirium, and how delirium has been changing throughout the years as intensive care, as a speciality, has evolved.

We demonstrated that delirium can be classified from routinely collected healthcare data of demographic, observational, and pathological variables. However, the resulting model was merely an association engine without significant explanatory power.

Thus, with the finding that in-attention is used in clinical scoring systems, we sought to automate this process. This led to the development of a non-invasive eye tracking platform that can regress the gaze-vector of a patient, whether their eyes are open, and estimate the Point of Regard (PoR) in the scene. This is all performed at a frequency that meets the Nyquist criteria, is calibration free, and is clinically safe. We validated the accuracy and precision of the platform compared to empiric requirements showing that it exceeded requirements.

To do this, we developed a novel curriculum learning method that led to state-of-the-art accuracy and performance in blink detection. We also

developed a novel screen-space PoR estimation technique that is state-of-the-art under rigorous conditions.

We then deployed the platform across two AICU, namely Chelsea & Westminster Hospital (CWH) and West Middelsex Hospital (WMH), where we collected daily recordings of recruited patients which were labelled with the current delirium state of the patient. Then, in a supervised learning scheme, trained two Temporal Convolutional Networks (TCNs) – one classified delirium purely based on eye movements whilst the second used the estimated PoR. These two networks demonstrated that eye-movements differ between episodes of delirium and non-delirium which are driven, in-part, by scene information.

To understand the results of these classifiers, we developed a novel architecture that uses a biologically inspired approach to fixation prediction. The architecture consists of multiple competing forward models that are ran in parallel and each output a distribution of possible next fixation which the actual fixation is then measured against. The correctness of each generative model is then tallied in a Winner Takes All (WTA) scheme.

The architecture was validated under stimulation, and, following hyperparameter tuning, was also validated using a visual search dataset. We then used the architecture on the collected AICU dataset where we demonstrated that episodes of delirium are punctuated by visual in-attention to low-level, mid-level, high-level, persons and motion in a statistically significant manner.

10.2 LIMITATIONS

In this section, we outlined the limitations of this thesis.

10.2.1 Hypothesis

The definition of delirium requires the presence of cognitive and memory dysfunction but scoring systems have been built and validated which replace cognition and memory with consciousness and *in-attention*. Whilst these scoring systems have been shown to be accurate, reliable, and easy to perform, our entire hypothesis of *visual* in-attention is shouldered by the link of task in-attention to visual in-attention. Anecdotally, and based on our clinical experience, this hypothesis holds – many patients on AICU who are delirious are also *visually* in-attentive (Exton and Leonard, 2009).

10.2.2 Longitudinal Healthcare Data

We responded to a call for research into mental health diseases as result of the COVID-19 pandemic (Holmes et al., 2020). We thus used routinely collected healthcare data to train a Long-Short Term Memory (LSTM) neural network for the classification of delirium. The homogeneity of the disease demonstrated the feasibility of approach. We used the first wave of COVID-19 at CWH to train the model and the second wave to perform validation. The networks validated well with excellent performance and calibration. However, the training of the network was limited by the relatively small number of patients. A larger number of patients would increase the confidence of the findings and facilitate prospective evaluation of the model.

10.2.3 Eye Tracking Platform

The eye tracking platform developed was accurate and precise, far exceeding empirical requirements in AICU. However, its use was limited to facially *normal* patients owing to the use of a face-detector and a facial-landmark extractor

that was trained on an annotated dataset of celebrity faces. The faces of patients in AICU are vastly different compared to the face of celebrities owing to the medical equipment used on patients in AICU that change the appearance of faces. Oxygen tubing, invasive venous access lines, and feeding tubes, were all found to interfere with the platform either through an intermittent failure to detect a face, and/or, incorrect facial landmark extraction. This limitation was mitigated by tuning the facial-detection threshold prior to every-recording but did result in the exclusion of a small number of recording sessions due to failure of the system to track them adequately. This was checked daily and recordings resumed when the platform was deemed to be accurate again.

10.2.4 *Delirium Classification using Temporal Convolutional Networks*

Following the development and validation of our eye-tracking platform, we deployed the platform across two hospitals and gathered recordings of delirious and non-delirious patients. The dataset was used to train a TCN from data collected from CWH and use data gathered from WMH for validation.

The number of recording sessions is a limitation to the external validity of our models. Whilst one use would be the continuous classification of delirium from eye-movement data, they served a purpose to understand if there is a difference in eye-movements between delirious and non-delirious recordings and the biological *source* of the difference.

In order to fully utilise the networks developed here, they would require further recordings and prospective evaluation.

10.2.5 *Delirium Classification using Competing Forward Models*

In order to explain the differences in visual-processing between delirious and non-delirious episodes, we developed and validated a competing forward models architecture ensuring a consistent baseline between participants. We took inspiration from what is currently known about the neuronal representations of the visual processing hierarchy in humans. The biological inspiration was not a simulation, but a computational representation of the functional units in the visual hierarchy. Thus, the immediate applicability of the findings of our competing forward models architecture to neurology and neuronal-circuitry is not possible. This would require further development of computational models that further span higher cortical centres.

10.3 FUTURE DIRECTIONS

There are several directions in which this work can be built upon. In this section, we aim to highlight a few.

10.3.1 *Prospective Evaluation*

The neural networks trained for the classification of delirium from eye-movement or scene data could be prospectively evaluated. This would require appropriate regulatory approvals, the on-boarding of several research and development departments across several hospitals that have an [AICU](#). This would provide further external validity of the networks for the ultimate goal of using the platform for the automated diagnosis of delirium.

10.3.2 *Neuro-computational models*

Given the success of the competing forward models at identifying differences between delirious and non-delirious episodes, a finer grained approach that is closer to the biological neuronal representation would facilitate a deeper and fuller understanding of visual attention in delirium. This would identify specific neuronal machinery that could be used as a therapeutic target.

10.4 EPILOGUE

Delirium, a disturbance of cognition and consciousness, operationalised to be a disturbance of memory and attention, has significant clinical consequences and thus the prompt diagnosis is of importance. In this thesis, we highlighted avenues for the classification of delirium and developed a novel platform for the non-invasive acquisition of eye-movements from patients in [AICU](#). Despite the limitations described above, we believe that the work presented describes a path towards improved delirium detection for the ultimate goal of improving patient care.

AUTHOR'S PUBLICATIONS

This section presents all peer-reviewed publications that resulted from this thesis. For each publication, we provide a summary, the full reference, and indicate how each publication relates to the thesis.

Al-Hindawi A, Vizcaychipi MP, Demiris Y. Faster, Better Blink Detection through Curriculum Learning by Augmentation. In 2022 Symposium on Eye Tracking Research and Applications 2022 Jun 8 (pp. 1-7).

[doi:10.1145/3517031.3529617](https://doi.org/10.1145/3517031.3529617)

In this article, we present an algorithm for training of a blinking classifier that combines Curriculum Learning and Data Augmentation into a training regime called Curriculum Learning by Augmentation ([CLbA](#)). The training regime results in a state-of-the-art blink detector whilst using a smaller, and thus more performant, neural network.

This article contributed to [Chapter 5](#)

Al-Hindawi A, Vizcaychipi MP, Demiris Y. What Is The Patient Looking At? Robust Gaze-Scene Intersection Under Free-Viewing Conditions. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) 2022 May 23 (pp. 2430-2434). IEEE.

[doi:10.1109/ICASSP43922.2022.9746437](https://doi.org/10.1109/ICASSP43922.2022.9746437)

In this paper, we present a dichotomisation scheme for the estimation of the Point of Regard ([PoR](#)) from a quasi-egocentric perspective. This scheme reduced the number of point-ray inter-

section calculations required whilst also being more accurate and robust to sparse depth maps.

This paper contributed to [Chapter 6](#)

Al-Hindawi A, Vizcaychipi MP, Demiris Y. Continuous Non-Invasive Eye Tracking In Intensive Care. In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) 2021 Nov 1 (pp. 1869-1873). IEEE.

[doi:10.1109/EMBC46164.2021.9630999](https://doi.org/10.1109/EMBC46164.2021.9630999)

In this paper, the eye-tracking platform for use in Adult Intensive Care Unit ([AICU](#)) is presented alongside empiric evaluations of its accuracy and performance. Posterior predictive checks on a sample of patients in [AICU](#) demonstrate that the platform produces distribution of data in-keeping with the training distributions.

The paper contributed to [Chapter 4](#).

11.1 PENDING

The following articles have been submitted and are currently under peer-review:

Al-Hindawi A, Vizcaychipi MP, Demiris Y. Non-Invasive Eye Tracking for the Diagnosis of Delirium in Intensive Care. Submitted to Nature Medicine 2022 Jun 14.

In this paper, we present the findings of our Temporal Convolutional Networks ([TCNs](#)) in classifying delirium from eye-movement data. We also present our theory of mind inspired hierarchical competing forward models architecture which we demonstrate

that episodes of delirium are punctuated by visual inattention to a statistically significant degree.

The paper contributed to [Chapters 8](#) and [9](#).

SCANPATH METRICS

This appendix reports on a preliminary experiment in trying to ascertain similarities between an observer and a baseline as a possible solution to the inverse yabus problem presented in [Chapter 9](#). The models outlined here were ultimately not used but could serve initial experiments for further understanding of eye movements, specifically scan-paths.

Two main metrics have been created for the comparisons of scan paths but challenges remain due to the complexities of what a scanpath represents. Recently, an algorithm called MultiMatch was developed that fuses different characteristics of scanpath metrics into a single metric. The aim of this appendix is to compare and contrast MultiMatch with ScanMatch, a more established scanpath comparison metric.

A.1 SCANMATCH ALGORITHM

The ScanMatch algorithm is a variant on the Levenshtein distance metric where the image is segmented into Regions Of Interests ([ROIs](#)) ([Cristino et al., 2010](#)). Fixations are then encoded in terms of the 'hit' on [ROI](#). Durations of fixation are encoded as multiple recurrent sequences of the [ROI](#) - the 'temporal bins' have a fixed duration. Each [ROI](#) is coded as a string and string comparison algorithms are then used to find alignment between two strings.

The Needleman-Wunsch algorithm globally aligns strings by using local optimal alignment of sub-sequences to guarantee the best overall alignment. The core of the algorithm is a substitution matrix that places a cost on the substitution of one string for another; this matrix can be composed of any

metric but the inverse of the Euclidean distance of the ROIs is often used; bins are formulated based on the distance with positive scores given when the distance is below threshold and a negative score when above. Differing length strings are matched in length by the insertion of "gaps" - these have a penalty score which can be inserted if they improve the overall global matching. The score is then normalised.

The main drawbacks of ScanMatch has been its reliance on ROIs where two fixations near boundaries of ROIs are given differing scores despite their similarity. While this could be remedied by improved construction of the substitution matrix, this requires foresight into the underlying processes that created the scanpaths.

ScanMatch's algorithm is also scan path shape invariant, a property which is not ideal. Recent neurophysiological evidence suggests that the shape of the scan path is highly relevant to the underlying cognitive processes (Johansson, Holsanova, and Holmqvist, 2006).

A.2 MULTIMATCH

In trying to remedy the reliance of ScanMatch on ROIs, MultiMatch was developed to addresses ScanMatch's weaknesses by using a vector based approach instead of ROI (Dewhurst et al., 2012). It also compares scan paths along multiple dimensions of interest that would be sensitive to differing processes; namely that of shape, position, length, direction and duration.

The algorithm is outlined in Figure A.1; the scan path vectors are compared along a metric, usually Euclidean distance. An $N \times M$ Matrix is formed where the elements of the matrix represent the metric (euclidean distance). The matrix is then transformed to Adjacency Matrix $A(k, l)$ and Dijkstra's algorithm is ran on A resulting in the shortest path. While Dijkstra's algorithm has been proven to find the shortest path in a graph (and in this case, a

Table A.1: Distance metrics used in the evaluation of a scan path.

Dimension	Distance Definition
Shape	The vector difference between aligned saccade pairs
Length	The difference in length between the end points of saccade vectors
Direction	Angular difference between the saccade vectors as a measure of similarity in shape when saccadic amplitudes are different
Position	The difference in position between aligned fixations using euclidean distance
Duration	This is the difference in fixation durations between aligned fixations - a measure of processing time.

globally aligned scanpath), it does not guarantee optimal local matching making the comparison of long sequences less reliable.

The matrix is normalised by the diagonal of the viewable area resulting in a path that is normalised between 0 and 1 where 0 is a perfect match. [Table A.1](#) summarises the metrics.

A.3 DYNAMIC TIME WARPING

Dynamic Time Warping ([DTW](#)) is an algorithm that globally aligns signals that are often of differing lengths using dynamic programming (Berndt and Clifford, 1994). The intuition is that locality based metrics often miss better matching if the phases of the paths were slightly shifted - i.e. global matching would be improved if one string would be ‘warped’ into the next. The amount of warping necessary is the [DTW](#) distance. The algorithm starts by the creation of an $N \times M$ matrix of distances (usually Euclidean) between the two points. Dynamic programming then recursively matches the signals according to the path of least resistance. This open strategy can lead to spurious matches where the search space is continuously searched for a better match due to its expansive nature, thus constraints have been imposed

on the dynamic programming nature of the algorithm to ensure accurate and speedy matching.

A.4 COMBINING DTW AND MULTIMATCH

One of the strength's of Dynamic Time Warping (DTW) is it's ability to constrain the search space to maintain temporal coherency. The intuition here is that it is not often ideal to match extremes of signals at the end of one signal with the beginning of the other. Thus algorithms have been developed to optimally constrain the search space for signal matching (Ratanamahatana and Keogh, 2004).

This is in contrast to MultiMatch's algorithm that relies on Dijkstra's algorithm for temporal coherency; while forward consistency is maintained, matching is not performed to maintain locally optimality but finds the global optimal shortest path. This has the effect that in scanpaths of differing lengths, the matching fails to maintain local temporal coherency in favour for an otherwise global match. This is due to DTW solving small repeated smaller sub-problems and finding the optimal solution to this sub-problem rather than Dijkstra's finding the best fit globally.

We propose a solution to combine the strengths of MultiMatch with Dynamic Time Warping (DTW). Scanpaths are aligned within DTW using the 'shape' distance metric from MultiMatch in an initialisation step. Once initialisation is complete, aligned scan paths are then compared using the 3 other MultiMatch metrics without further alignment. This ensures local consistency and has the side effect of being a simpler algorithm.

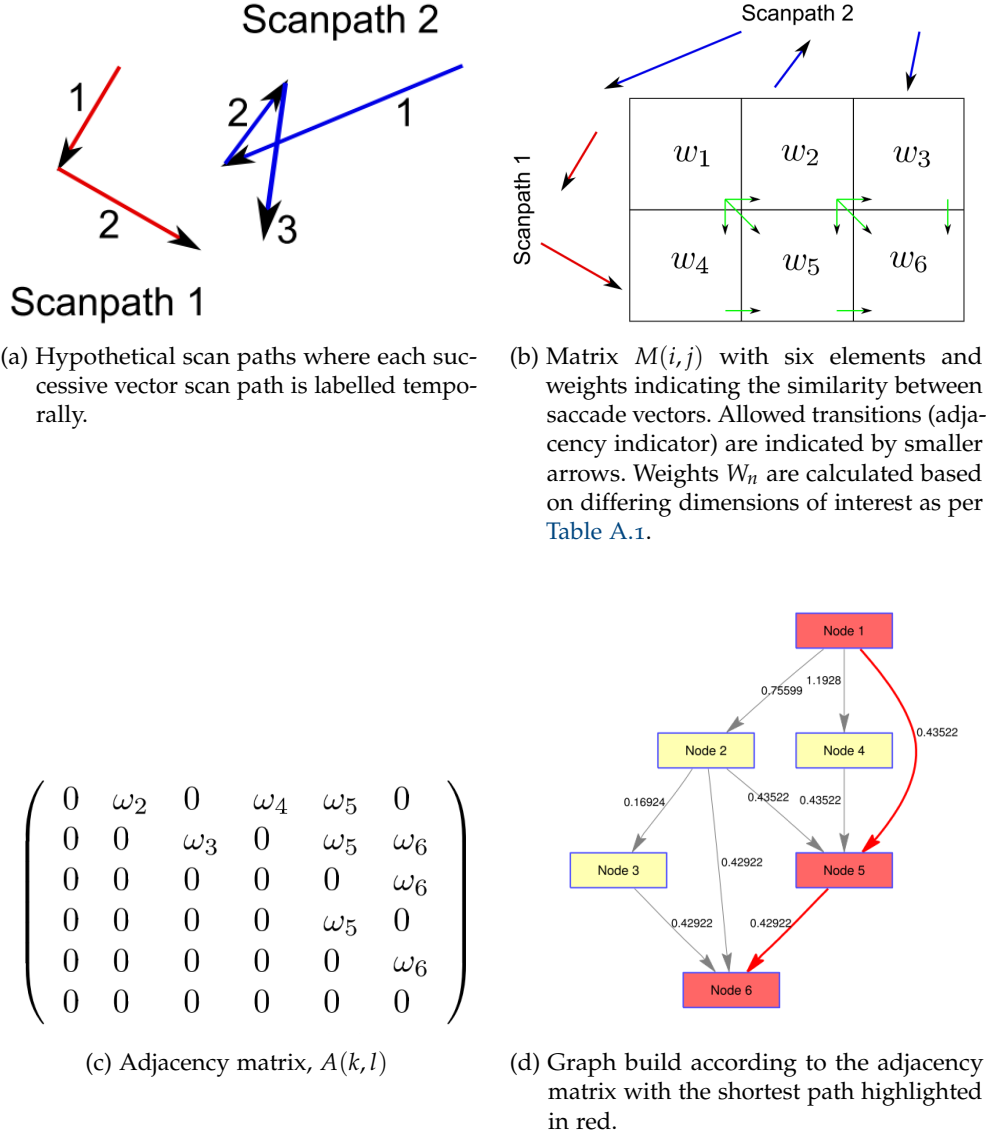


Figure A.1: Multimatch Algorithm where Subfigure A.1a is the scan paths to compare. Figure A.1b is the comparison matrix according to a predefined metric. Subfigure A.1c is the adjacency matrix calculated from the comparator matrix in Figure A.1b. A graph is then constructed from top left to bottom right and Dijkstra's algorithm is used to find the shortest path as per Figure A.1d (Jarodzka, Holmqvist, and Nyström, 2010).

A.4.1 Validation of MultiMatch DTW

The modification of MultiMatch through the use of DTW has since been validated in a simulation of scan-paths. A total of 6000 sets of scan-paths were generated where each set contains two scan paths that were randomly

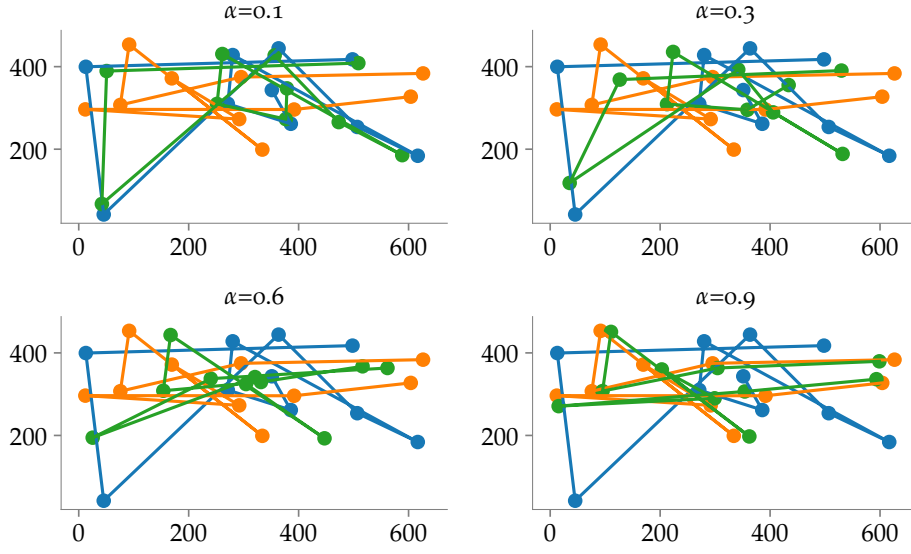


Figure A.2: Example of an interpolation between two scan-paths, $sp1$ (blue) and $sp2$ (orange). $sp3$ (green) is being interpolated with the corresponding α values from $sp1$ into $sp2$.

generated using a uniform distribution with an emulated size of 640×480 pixels. A third scan-path was interpolated between the first and the second scan-paths using:

$$sp3 = (1 - \alpha) sp1 + \alpha sp2 \quad (A.1)$$

where α is the interpolation constant that ranged from 0.1 to 0.9 in 9 steps; [Figure A.2](#) demonstrates an example of this.

The metrics of position, direction, shape and length were calculated for the comparison using $S(sp1, sp3)$ and $S(sp2, sp3)$ where S is a function that returns the above referenced scan-path metrics. $sp3$ was labelled as having been correctly classified if $S(sp1, sp3) > S(sp2, sp3)$. [Figure A.3](#) demonstrates the validation results. The percentage of correctly identified scan-paths is demonstrated for $\alpha < 0.5$ which demonstrates high confidence of correct classification.

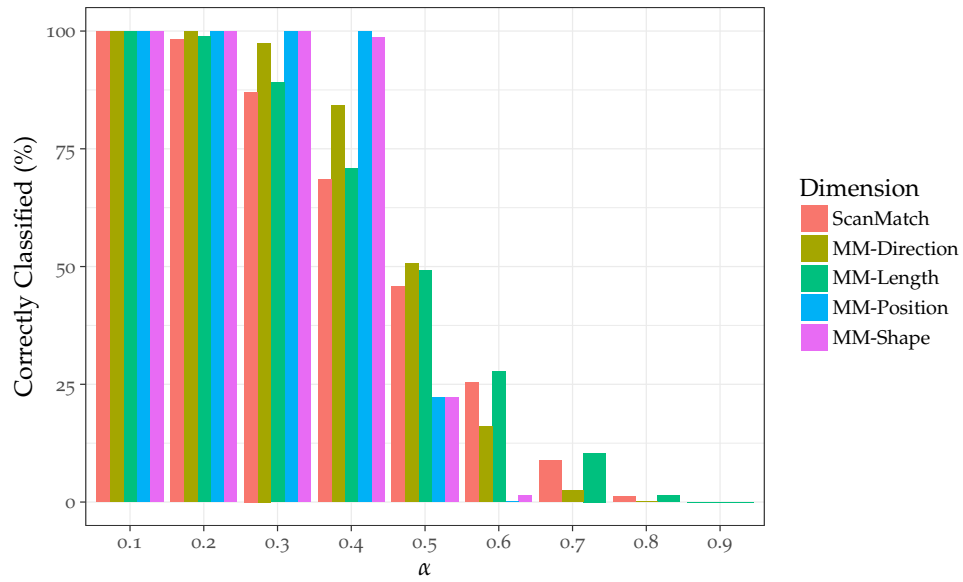


Figure A.3: Validation result of combining MultiMatch with DTW. As $sp1$ is interpolated to $sp2$ as $sp3$ using the interpolation constant α the classification accuracy stays relatively accurate and only decreases steeply once the $sp3$ is ambiguous. ScanMatch fails linearly in comparison to MultiMatch.

TRIAL DOCUMENTS

B.1 PROTOCOL

Included in this section is the protocol for Continuous Non-Invasive Eye Tracking for the Diagnosis of Delirium ([CONfuSED](#)). The trial's protocol has been reviewed and approved by Research & Ethics Committee ([REC](#)) and Health Research Authority ([HRA](#)) on the 27th of July 2020. The study gained Sponsorship by Chelsea and Westminster Hospital NHS Foundation Trust and was subsequently adopted by National Institute of Health Research ([NIHR](#)).

The following documents are part of the Site File – working documents for clinical trials across the UK. Due to regulatory requirements, they were authored in Microsoft Word; PDFs of the final versions are thus appended here.

Short Title/Acronym: CONfuSED

IRAS Ethics Ref: 264759

Date: 27/07/2020

Version No: 3.0

Study Title: Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit

Chief Investigator: Dr Marcela Vizcaychipi, Consultant in Anaesthesia, Chelsea and Westminster Hospital NHS Foundation Trust
marcela.vizcaychipi@chelwest.nhs.uk

Investigators: Dr Ahmed Al-Hindawi, Research Fellow in Anaesthetics, Chelsea and Westminster Hospital NHS Foundation Trust
ahmed.al-hindawi@nhs.net

Sponsor: Chelsea and Westminster Hospital NHS Foundation Trust
Damon Foster, Research Operations Manager, Chelsea and Westminster
damon.foster@chelwest.nhs.uk; 020 3315 6887

Funder: Westminster Medical Research Fund
Chelsea & Westminster Hospital NHS Foundation Trust

**Chief Investigator
Signature:**



Confidentiality Statement

This document contains confidential information that must not be disclosed to anyone other than the Sponsor, the Investigator Team, host organisation, and members of the Research Ethics Committee, unless authorised to do so.

Table of Contents

1 Synopsis	4
2 Abbreviations	5
3 Background & Rationale	6
4 Objectives and Outcome Measures/Endpoints	8
5 Study Design	9
6 Participants.....	10
6.1 Participant Identification	10
6.2 Inclusion Criteria.....	10
6.3 Exclusion Criteria	10
7 Study Procedures	11
7.1 Recruitment	11
7.2 Consent	11
7.3 Risks & Benefits.....	11
7.4 Screening & Eligibility Assessment	12
7.5 Ongoing Assessments	12
7.6 Discontinuation/Withdrawal of Participants from Study	12
7.7 Definition of End of Study	13
8 Statistics & Analysis.....	14
8.1 Description of Statistical Methods	14
8.2 Number of Participants	14
9 Data Management	15
9.1 Access to Data	15
9.2 Data Recording & Record Keeping.....	15
10 Ethical & Regulatory Considerations	16
10.1 Declaration of Helsinki	16
10.2 ICH Guidelines for Good Clinical Practice.....	16
10.3 Approvals.....	16
10.4 Reporting	16
10.5 Participant Confidentiality	16
11 Publication Policy	16
12 Bibliography	17

13 Calculation of E-PRE-DELIRIC	19
14 Acceptability Questionnaire	20

1 Synopsis

Study Title	Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit	
Short Title/Acronym	CONfuSED	
Study Design	Feasibility/Pilot study	
Study Participants	Adult patients admitted to HDU/ICU at high risk of developing Delirium	
Planned Sample Size	30	
Planned Study Period	Eight to Twelve months	
	Objectives	Endpoints
Primary	1. Demonstrate eye gaze correlates with CAM-ICU. 2. Acceptance of the use of cameras and artificial intelligence on ICU	Development of Delirium

2 Abbreviations

CAM-ICU	Confusion Assessment Method in ICU - A scoring system used for the diagnosis of Delirium on Intensive Care
ICU	Intensive Care Unit
E-PREDELIRIC	Early PREdiction of DELIRium in ICu patients - Linear regression model that estimates the probability of developing delirium from baseline characteristics
SVM	Support Vector Machine
ICF	Informed Consent Form
PIS	Participant/Patient Information Sheet
CWFT	Chelsea & Westminster NHS Foundation Trust
ICMJE	International Committee of Medical Journal Editors
LSTM	Long Short-Term Memory – a type of neural network module aimed at learning time related data.
ICH	The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
NOK	Next of Kin
SOP	Standard Operating Procedure
CRF	Case Report Form

3 Background & Rationale

The use of technology in health care is increasing with some implementation even reaching mainstream news media attention. This has had the impact of increasing the amount of data available to health care professionals which has led to the development of sophisticated machine learning frameworks that have significant predictive power [2]. Intensive Care, as a medical speciality, is a highly reliant on technology. Patients in the Intensive Care Unit (ICU) are continuously and invasively monitored throughout their stay leading to the generation of massive amounts of data. Technology has also led to the development of sophisticated treatment modalities such as ventilators and dialysis machines. This speciality is also very welcoming to novel technology; recent work introduced the NAO robot into a subsection of Intensive Care known as the High Dependency Unit [3]. One of the main challenges for the future of intensive care is improving long term outcomes following survival of a critical illness. Many patients suffer with psychological sequelae following their intensive care stay that often result in a dramatic impact on their quality of life. The psychological sequelae often stem from the development of Delirium.

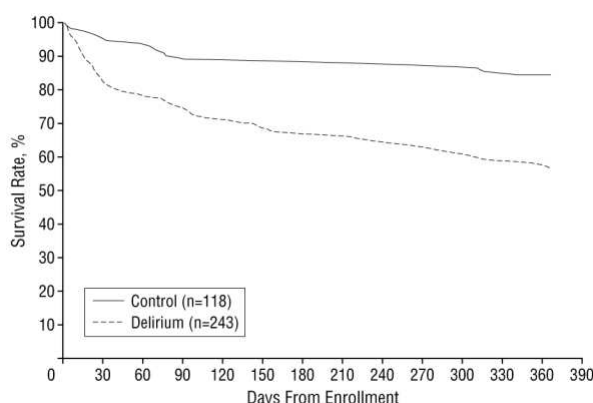


Figure 1: Mortality of patients with and without delirium demonstrating increased mortality of patients who suffer with delirium during the in-patient stay.

Delirium is defined as transient temporally fluctuating disorder of cognition and memory. It can occur at any age, but it occurs more commonly in patients who are elderly and have compromised mental status [4]. This state can occur in any acutely unwell patient and those who are socially isolated but older patients are often at risk as they have lower functional cognitive reserves to deal with pathological insults, such as surgery, sepsis, metabolic derangements and burns [5, 6, 7, 8, 9]. The development of delirium in patients can be thought of in terms of organ failure. This metaphor has implications as new data has suggested that patients who develop delirium during their stay in hospital have increased mortality compared to patients who do not (Figure 1)[10, 11, 1]. A conservative estimate of the incidence of delirium in acutely unwell patients is 20% and has been reported as high as 80% [12, 10]. Thus, the temporaneous diagnosis is of paramount importance to facilitate timely individualised treatment.

Current detection methods of delirium require manual assessment of the patient, usually performed by the bed-side nurse. The assessment takes the form of a question-answer response where the nurse asks questions that probe into the memory and cognition that are then presented as diagnosis. The most widespread scoring system is the Confusion Assessment Method in ICU (CAM-ICU) scoring system. The central tenant of CAM-ICU is 'in-attention' whereby the patient is asked to follow a command repeatedly that tests memory and executive function; the operator asks the patient to squeeze their hand every-time the letter 'A' is said, the operator then would spell 'SAVE A HAART', a squeeze on letter other than 'A' or failed to squeeze on 'A' are considered errors, more than 2 errors classify delirium. This test is intermittent and labour intensive.

Eye tracking has a long history in the diagnosis and monitoring of neuropsychiatric disease. The advent of high frequency eye tracking hardware and appropriate algorithms led to the automatic

measurement and classification of saccades, fixations and smooth pursuit. These three eye movements, in conjunction with eye tracking paradigms, have led to the accurate and

system setup. The contemporaneous diagnosis of schizophrenia and affective disorders [13, 14, 15]. More recently, Zhang et al have developed an automated algorithm that monitors dementia through the use of non-invasive eye tracking technology [16].

Exton *et al* hypothesised that delirium could be diagnosed through eye tracking [17]. The hypothesis hinges on clinical experience backed by earlier work in other domains as outlined above but doesn't shed light on the practical aspects of eye-tracking in delirious patients that are often phenotypically agitated and by definition have cognitive and memory dysfunction. Thus, the diagnosis of delirium through eye tracking is currently withheld by the lack of a robust technical solution that is amenable to the use with delirious patients. The current devices available are mounted on the user in the form of glasses or caps and are unsuitable for the use in patients with delirium. Calibration, a requirement of current technologies is a further hurdle which delirious patients cannot participate in due to their cognitive dysfunction.

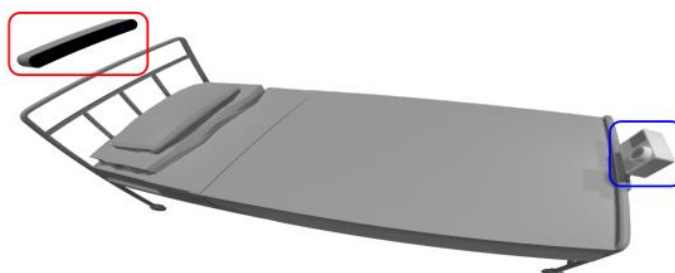


Figure 2: Render of the system demonstrating the two-camera system setup. The Camera placed at the back of the bed (Outlined in Red) is a depth camera while the camera looking at the patient (Outlined in Blue) is a normal 2D camera. Both are connected to a laptop situated below the bed.

To this extent, we have developed a non-invasive, continuous, near real-time, bed camera-based system that measures the gaze of a participant, the setup of which is shown in Figure 2. Two cameras are required, one at the end of the bed (termed the head camera) that looks at the patient and measures their gaze while the other is behind the patient looking at the scene (termed the scene camera). State of the art convolutional neural networks are run for facial recognition and then gaze regression to then be subsequently mapped out in the reference frame of the scene camera. The system performs with superhuman level accuracy at 12-14Hz [18]. In order to understand the participant's gaze behaviour, simulations of gaze behaviour are concurrently run from the scene camera and are then compared to the gaze of the participant. This results in a score of how attentive the participant is to the scene presented to them compared to controls. This document outlines the Protocol for a feasibility study looking into the use of eye-tracking on ICU as a method of diagnosis of delirium.

4 Objectives and Outcome Measures/Endpoints

In this trial, we seek to demonstrate feasibility of the system in continuously detecting 'inattention' as a subset of CAM-ICU in a small representative sample of patients on ICU at Chelsea & Westminster NHS Foundation Trust (CWFT). Table 1 demonstrates the objectives of the study and Figure 3 demonstrates the feasibility trial's process as a flow chart.

Objectives	Outcome Measures/Endpoints
Primary Objective	Demonstrate eye gaze correlates with CAM-ICU
Secondary Objectives	Acceptance of the use of cameras and artificial intelligence on ICU
Tertiary Objectives	None

Table 1: Table demonstrating the objectives of the feasibility study

5 Study Design

Figure 3 demonstrates the study design process of the feasibility trial and the patient's course within it. Inclusion & Exclusion are highlighted as per Section 5 with the study procedures in Section 6.

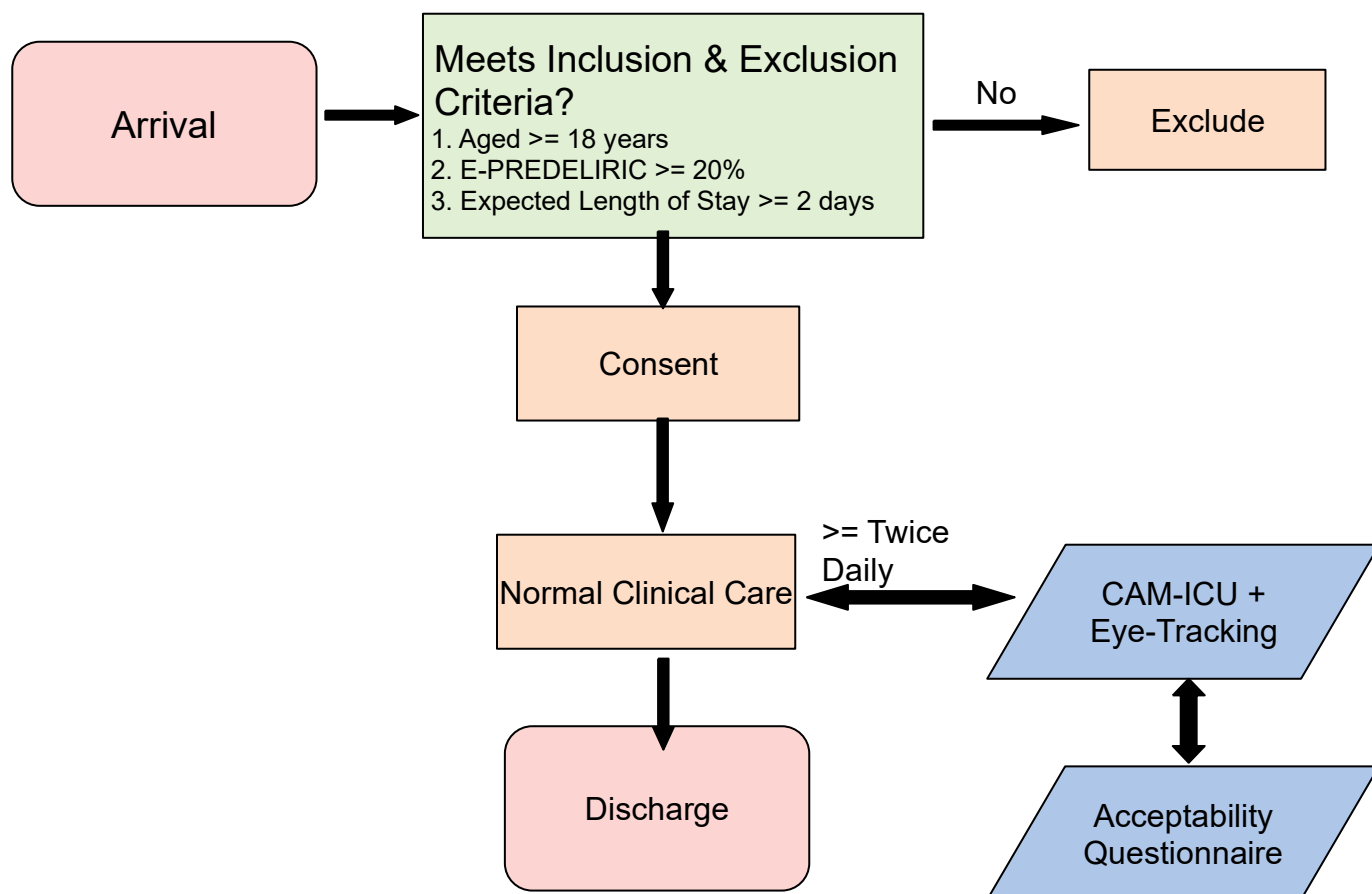


Figure 3: Process of recruiting patients into the clinical trial. Green box represents decision/questions while red boxes indicate start/stop conditions. Blue box indicates the research points within the trial.

Patients arriving to ICU will be highlighted to the research team by the clinical team. The clinical team will seek permission from the patient, if capacitious, to inform the research team. If inclusion criteria are met and none of the exclusion criteria are present, then the patient is approached for consent. If the patient's capacity means that they are not able to consent, then the patient's relatives/friends will be approached for advice or, if not available, a nominated consultee will be appointed. A nominated consultee will be one of the Intensivists at Chelsea & Westminster Hospital who are not involved in the care of the patient at time of advice. If in their opinion, the patient would not raise an objection to participate in the study then the usual care to the patient continues and the system will be set up within the bed space of the patient. The system is only switched on 5 minutes prior and 5 minutes following the assessment of Delirium by CAM-ICU. At any other time, the system is switched off and the lenses covered.

6 Participants

6.1 Participant Identification

We aim to recruit patients into the trial that represent the general population admitted to a general ICU. Thus, we aim to recruit all comers into ICU that meet our inclusion and exclusion criteria.

6.2 Inclusion Criteria

1. Aged >18
2. Predicted Delirium as defined by the Early PREdiction of DELIRium in ICu patients (E-PREDELIRIC) score $\geq 20\%$
3. Expected length of stay ≥ 2 days

6.3 Exclusion Criteria

1. Non-concordant eyes
2. Visual Impairment
3. Dementia
4. Inability for facial recognition and eye tracking to be performed reliably
5. Lack of consent
6. In-ability to perform CAM-ICU reliably

7 Study Procedures

7.1 Recruitment

Patients will be identified as potential study subjects in ICU by the clinical team referring to the research team. If found to be eligible for inclusion once evaluated with the relevant inclusion and exclusion criteria, the patient is to be consented using the standardised Informed Consent Form (ICF) by one of the registered study personnel.

7.2 Consent

Consent to enter this study is to be obtained in one of two ways; as the trial studies delirium, a capacity losing state, the assessment of capacity and thus consent is of paramount importance. Once the patient has been flagged as a potential study participant, the research team will perform an assessment of the patient's capacity and if present, fully informed consent will be gained from the patient directly. This is done following a full verbal and written account of the study's nature, purpose, risks and potential benefits. A Participant/Patient Information Sheet (PIS) will be provided and the patient will have the opportunity to deliberate for 24 hours. It will be clearly stated that the patient is free to withdraw from the study at any time for any reason without prejudice to future care, and with no obligation to provide the reason for withdrawal. Should the patient feel that further deliberation will not lead to a change in their decision and the researcher seeking consent is satisfied that the patient has retained, understood and deliberated on the information provided, then consent can be taken. Before any study specific procedures are performed, patients will personally sign and date the ICF. One copy of the signed ICF will be given to the patient, one copy of the signed ICF will be placed in the patient's medical notes and the original signed ICF will be retained at the study site in the Investigator File.

Should the participant's mental state, competence and capacity, mean that they are unable to provide consent, the patient's relatives/friends will be approached for advice using the same procedure as above.

Should the a relative/friend not be available, then a nominated consultee will be sought. They will be one of the Intensivists at Chelsea & Westminster Hospital who are not involved in the care of the patient at time of advice.

Should the patient recover their capacity, fully informed consent will be undertaken to ensure that their participation in the study is in-line with their wishes.

7.3 Risks & Benefits

As this is a purely non-invasive feasibility study, the risks to patients are minimal and surround privacy and no direct patient contact will occur as a result of the research equipment. However, as the camera equipment will surrounding the patient and could be deemed intrusive, two techniques will be employed to minimise privacy intrusion:

1. Cameras will be turned off and lenses physically covered when not recording.

2. Recording happens at the discretion of the bedside nurse around the time their assessment of the patient's mental state occurs as part of routine clinical practice. This ensures patient's privacy and integrity are maintained throughout the recording.
3. No direct recording of the images is performed, but the mathematical representation of the comparison between simulations and eye gaze behaviour is recorded instead. This avoids any intrusion on patient's privacy and maintains integrity.

7.4 Screening & Eligibility Assessment

Patients' eligibility will be decided upon prior to obtaining consent and a completed checklist for each patient will be attached to the consent form in the study folder.

7.5 Ongoing Assessments

Patient's usual clinical care will not be affected by the assessments performed by this study. In our institution, CAM-ICU is performed at least twice daily and more so if there are changes to mental status. The system will be switched on 5 minutes prior to performing CAM-ICU and will record for a duration of 10 minutes following which, the system will be turned off and the cameras covered. This procedure will occur every time CAM-ICU is performed. Once per day, due to nursing scheduling, a questionnaire (Outlined in Section 13) will be given and filled by the current nurse. The sum of which will be analysed using qualitative methods.

7.6 Discontinuation/Withdrawal of Participants from Study

Each patient has the right to withdraw from the study at any time. In addition, the Investigator may discontinue a participant from the study at any time if the Investigator considers it necessary for any reason including:

- Ineligibility (either arising during the study or retrospectively having been overlooked at screening)
- Significant non-compliance with study requirements
- Withdrawal of consent
- Other reasons outlined here

If a patient withdraws from the trial, data already collected with consent will be retained and used in the study. However, the patient will be given the option, but encouraged not to, remove their data from the study. No further data will be collected or any other research procedures carried out on or in relation to the patient. The reason for withdrawal will be recorded in the Case Report Form (CRF).

Furthermore, the clinical and/or research team can stop the system in the event of any unexpected complication.

The system (cameras, laptop and cables) will be cleaned in accordance with their CE marking and by the recommended methods as per the manufacturers.

7.7 Definition of End of Study

The end of the study will be when 30 patients have been recruited and completed their participation in the trial.

8 Statistics & Analysis

8.1 Description of Statistical Methods

Eye gaze data from delirious and non-delirious episodes will be labelled accordingly and machine learning approaches aimed at classifying time-series data will be used. These include Support Vector Machines (SVMs) and Long-Short Term Memory (LSTM) neural networks.

Questionnaire responses will be analysed using qualitative descriptive statistical methods such as bar graphs, median and interquartile ranges.

8.2 Number of Participants

Due to the novel nature of the system, no prior study exists to infer sample size and thus this is a feasibility study. The suggested sample size is thus 30 patients split into two groups as per [19].

In order to maximise the chance of delirious patients being recruited, we aim to use the E-PRE-DELIRIC score, a linear regression model that aims to predict the development of delirium in ICU. The ICU at Chelsea & Westminster Hospital's a median E-PRE-DELIRIC score is 20%. Thus, by only including patients above this threshold, the number of patients who develop delirium is maximised. Some of those "high-risk" patients will develop delirium, some of them won't. The patients who will develop delirium will then form the positive cohort of patients and the ones that don't become the experimental control group. As the patient's order of coming into ICU is stochastically random, confounders will be distributed between the patients who will become delirious and not delirious thus matching the two groups as best as possible.

Recruiting patients with a higher than median E-PRE-DELIRIC score of 20% yields a theoretical minimum of 6/30 delirious patients. However, as the system operates at 10-20Hz, our recording speed is sufficient to overcome this small sample size.

Each patient will have at least twice daily CAM-ICUs with 5-minute recordings either side (10 minutes/recording = 20 minutes/day). Chelsea & Westminster Hospital's ICU's median length of stay is 3 days giving circa 60 minutes of recordings/patient. The system runs at 10-20Hz generating at least ~40,000 data points/patient which are then labelled according to whether the patient was CAM-ICU positive or not.

The ~240,000 data points generated from the potential 6 delirious patients is sufficient as that will saturate the best time series machine learning algorithms available with the other 24 patients serving as controls.

9 Data Management

9.1 Access to Data

Direct access will be granted to authorised representatives from the Sponsor or host institution for monitoring and/or audit of the study to ensure compliance with regulations. The trial may be monitored, or audited in accordance with the current approved protocol, ICH GCP, relevant regulations and standard operating procedures.

9.2 Data Recording & Record Keeping

Data will be stored on secure encrypted laptop at CWFT. The laptop will be stored in a locker that is locked using a 4-digit key code that is only known to the principal investigator. The locker is situated behind secured doors that are only accessible using preauthorised personnel using ID badges supplied under the security of CWFT. The room is situated in a corridor also secured using ID badges. Both the room and the corridor's access are monitored continuously and access is recorded centrally.

10 Ethical & Regulatory Considerations

10.1 Declaration of Helsinki

The Investigator will ensure that this study is conducted in accordance with the principles of the Declaration of Helsinki.

10.2 ICH Guidelines for Good Clinical Practice

The Investigator will ensure that this study is conducted in full conformity with relevant regulations and with The International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH) Guidelines for Good Clinical Practice.

10.3 Approvals

All relevant documentation will be submitted to the NHS REC and host institution R&D departments. The Investigator will submit and, where necessary, obtain approval from the above parties for all substantial amendments to the original approved documents.

10.4 Reporting

An End of Study notification and final report will be submitted to the same parties. Trials will be registered to clinicaltrials.gov.

10.5 Participant Confidentiality

All CRFs will contain a unique study identification number but not the patients name or other personal identifiable data, to prevent identification of subjects involved in the study, and the Sponsor Subject ID Log will be used to cross reference patient's identifiable information. Of note, access to data will be in-line with Sponsor Standard Operating Procedure (SOP) and will abide to the Guide to the General Data Protection Regulation (GDPR) 2018 principles / be carried out in accordance to GDPR 2018 principles.

11 Publication Policy

After the study has ended, all results, once analysed and interpreted, will be used to draft a manuscript for submission and potential publication to an appropriate scientific journal. The Investigators will be involved in reviewing drafts of the manuscripts, abstracts, press releases and any other publications arising from the trial. Authorship will be determined in accordance with the International Committee of Medical Journal Editors (ICMJE) guidelines and other contributors will be acknowledged.

12 Bibliography

- [1] J. McCusker, M. Cole, M. Abrahamowicz, F. Primeau, E. Belzile, Delirium Predicts 12-Month Mortality, *Archives of Internal Medicine* 162 (2002) 457.
- [2] R. Miotto, L. Li, B. A. Kidd, J. T. Dudley, P. Agarwal, Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records, *Scientific Reports* 6 (2016) 26094.
- [3] M. Sarabia, Y. Demiris, A Humanoid Robot Companion for Wheelchair Users, *Social Robotics* 8239 (2013) 432–441.
- [4] A. P. A. A. P. Association, et al., DSM-IV-TR: Diagnostic and Statistical Manual of Mental Disorders, Washington, DC: American psychiatric association (1994) 143–147.
- [5] J. Holt-Lunstad, T. B. Smith, J. B. Layton, Social Relationships and Mortality Risk: A Meta-analytic Review, *PLoS Med* 7 (2010) e1000316.
- [6] R. Palmu, Mental Disorders among Burn Patients, *Burns* 36 (2011) 1072–1079.
- [7] M. Lundström, B. Olofsson, M. Stenvall, S. Karlsson, L. Nyberg, U. Englund, B. Borssén, O. Svensson, Y. Gustafson, Postoperative delirium in old patients with femoral neck fracture: A randomized intervention study., *Aging clinical and experimental research* 19 (2007) 178–186.
- [8] W. Meier-Ruge, O. Hunziker, P. Iwangoff, Senile Dementia: A Threshold Phenomenon of Normal Aging?., *Annals of the New York Academy of Sciences* 621 (1991) 104–118.
- [9] D. J. Meagher, Delirium: Optimising management., *BMJ (Clinical research ed.)* 322 (2001) 144–9.
- [10] S.-M. Lin, C.-y. Liu, C.-h. Wang, H.-c. Lin, C.-D. Huang, P.-y. Huang, Y.-f. Fang, M.-h. Shieh, H.-p. Kuo, The impact of delirium on the survival of mechanically ventilated patients., *Critical care medicine* 32 (2004) 2254–9.
- [11] S. Ouimet, B. P. Kavanagh, S. B. Gottfried, Y. Skrobik, Incidence, risk factors and consequences of ICU delirium, *Intensive Care Medicine* 33 (2007) 66–73.
- [12] M. A. Pisani, K. L. B. Araujo, P. H. Van Ness, Y. Zhang, E. W. Ely, S. K. Inouye, A research algorithm to improve detection of delirium in the intensive care unit., *Critical care (London, England)* 10 (2006) R121.
- [13] S. A. Beedie, D. M. St.Clair, P. J. Benson, Atypical scanpaths in schizophrenia: Evidence of a trait- or state dependent phenomenon?, *Journal of Psychiatry and Neuroscience* 36 (2011) 150–164.
- [14] C. M. Loughland, L. M. Williams, E. Gordon, Schizophrenia and affective disorder show different visual scanning behavior for faces: A trait versus state-based distinction?, *Biological Psychiatry* 52 (2002) 338–348.
- [15] P. Trillenber, R. Lencer, W. Heide, Eye movements and psychiatric disease., *Current opinion in neurology* 17 (2004) 43–7.

- [16] Y. Zhang, T. Wilcockson, K. I. Kim, T. Crawford, H. Gellersen, P. Sawyer, Monitoring Dementia with Automatic Eye Movements Analysis, in: I. Czarnowski, A. M. Caballero, R. J. Howlett, L. C. Jain (Eds.), *Intelligent Decision Technologies 2016*, volume 57, Springer International Publishing, Cham, 2016, pp. 299–309.
- [17] C. Exton, M. Leonard, Eye tracking technology: A fresh approach in delirium assessment?, *International Review of Psychiatry* 21 (2009) 8–14. [18] T. Fischer, H. Jin Chang, Y. Demir, RT-GENE: Real-Time Eye Gaze Estimation in Natural Environments, in: *Proceedings of the European Conference on Computer Vision (ECCV)*, pp. 334–352.
- [19] G. A. Lancaster, S. Dodd, P. R. Williamson, Design and analysis of pilot studies: recommendations for good practice, *Journal of Evaluation in Clinical Practice* 10 (2004) 307–312.
- [20] A. Wassenaar, M. van den Boogaard, T. van Achterberg, A. J. C. Slooter, M. A. Kuiper, M. E. Hoogendoorn, K. S. Simons, E. Maseda, N. Pinto, C. Jones, A. Luetz, A. Schandl, W. Verbrugghe, L. M. Aitken, F. M. P. van Haren, A. R. T. Donders, L. Schoonhoven, P. Pickkers, Multinational development and validation of an early prediction model for delirium in ICU patients, *Intensive Care Medicine* 41 (2015) 1048–1056.
- [21] M. MacMurchy, S. Stemler, M. Zander, C. P. Bonafide, Acceptability, Feasibility, and Cost of Using Video to Evaluate Alarm Fatigue, *Biomedical Instrumentation & Technology* 51 (2017) 25–33.

13 Calculation of E-PRE-DELIRIC

Table 2 lists the variables and regression coefficients for the accurate calculation of the E-PREDELIRIC linear regression scoring system [20].

Variable	Regression Coefficient
Age	0.025
History of Cognitive Impairment	0.878
History of Alcohol Abuse	0.505
Admission Category	
Surgery	0.000
Medical	0.370
Trauma	1.219
Neurology/Neurosurgery	0.504
Urgent Admission	0.612
MAP at time of ICU admission	-0.006
Use of corticosteroids	0.283
Respiratory Failure	0.982
BUN at time of ICU admission	0.018

Table 2: Variables and regression coefficients used in the E-PREDELIRIC model, reproduced from [20].

In order to calculate the risk of delirium for each patient, the applicable variables' regression coefficients are summed into a logit as per the following equations:

$$\text{logit} = \sum (\text{Table 2 Variables})$$

$$\text{Risk of Delirium} = \frac{1}{1 + e^{+3.907 + \text{logit}}}$$

14 Acceptability Questionnaire

This questionnaire is designed to ascertain the acceptability of the use of cameras within this study on day to day activities of the nursing staff on ICU. This questionnaire is based on work by MacMurchy et al [21].

Question	Answer
Did participating in this study affect your ability to care for your patients?	Yes/No
Did participating in this study affect your interactions with patients/families?	Yes/No
Did participating in this study affect your interactions with other nurses?	Yes/No
Did participating in this study affect your interactions with doctors?	Yes/No

B.2 DECLARATION FORMS

This section includes the declaration forms for personal and nominated consultees as approved by [REC](#) and [HRA](#).

Declaration form – For Nominated Consultee

Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit (CONfuSED)

The nominated consultee should complete the whole of this sheet him/herself.

The study has been explained to me by:

Prof/Dr/Mr/Mrs/Ms _____

I am the nominated consultee for:

Please initial each box:

I have read the Information Sheet version number _____ date _____.

☐

I have been given the opportunity to ask questions and discuss this study.

☐

I have received satisfactory answers to all my questions and received enough information about the study.

☐

I understand that once the person, for whom I am being asked to act as a nominated consultee, has regained capacity they will be given a participant information sheet and asked for their consent.

☐

I understand that I can request he/she is withdrawn from the study at any time, without giving any reason and without his/her care or legal rights being affected.

☐

I believe the person, for whom I am being asked to act as a nominated consultee, would not object to the access of their records by responsible individuals from the site, from the sponsor, and from regulatory authorities, where it is relevant to taking part in this research.

☐

I believe that _____ would have no objection to taking part in this study.

☐

Name of
Nominated Consultee

Date

Signature

Person undertaking consultation:

Investigator

Date

Signature

1 copy will be given to the consultee; 1 copy will be kept in the electronic medical notes and the original will be added to the investigator site file.

Declaration form – For Relative/Friend

Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit (CONfuSED)

The relative/friend should complete the whole of this sheet him/herself.

The study has been explained to me by:

Prof/Dr/Mr/Mrs/Ms _____

I am the (please circle) relative/friend of

Please initial each box:

I have read the Information Sheet version number _____ date _____.

☐

I have been given the opportunity to ask questions and discuss this study.

☐

I have received satisfactory answers to all my questions and received enough information about the study.

☐

I understand that once my relative/friend has regained capacity they will be given the information sheet and asked for their consent.

☐

I believe that my relative/friend would not object to the access of their records by responsible individuals from the site, from the sponsor, and from regulatory authorities, where it is relevant to my relative/friend taking part in this research.

☐

I believe that _____ would wish to take part in this study

☐

Signed _____ Date _____

NAME IN BLOCK CAPITALS _____

Investigator's signature _____ Date _____

NAME IN BLOCK CAPITALS _____

1 copy will be given to the consultee; 1 copy will be kept in the electronic medical notes and the original will be added to the investigator site file.

B.3 CONSENT FORMS

This section includes the consent forms for patients with capacity and for when a patient regains capacity as approved by [REC](#) and [HRA](#).



Patient Consent Form for Participants when they regain capacity

Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit (CONfuSED)

The participant should complete the whole of this sheet him/herself.

The study has been explained to me by:

Prof/Dr/Mr/Mrs/Ms _____

Please initial each statement as it applies to you.

I have read the Information Sheet version n. _____ date _____.

☐

I have been given the opportunity to ask questions and discuss this study.

☐

I have received satisfactory answers to all my questions and received enough information about the study.

☐

I understand that I am free to withdraw from the study at any time, without having to give a reason for withdrawing and without affecting my future medical care.

☐

I understand and agree to include the data already collected in the study.

☐

I understand my medical notes may be looked at by responsible individuals from the site, from the sponsor, and from regulatory authorities, where it is relevant to my taking part in this research.

☐

I give permission for these individuals to access my records that are relevant to this research.

I give permission to other researchers associated with the project to access study data in an anonymised format.

I give consent to remain in this study.

☐

Signed _____ Date _____

NAME IN BLOCK CAPITALS _____

Investigator's signature _____ Date _____

NAME IN BLOCK CAPITALS _____

1 copy will be given to the participant; 1 copy will be added to the investigator site file and 1 (original) will be kept in the medical notes.

Participant Consent Form – Patients with Capacity

Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit (CONfuSED)

The participant should complete the whole of this sheet him/herself.

The study has been explained to me by:

Prof/Dr/Mr/Mrs/Ms _____

Please initial each statement as it applies to you.

I have read the Information Sheet version n. _____ date _____.

☐

I have been given the opportunity to ask questions and discuss this study.

☐

I have received satisfactory answers to all my questions and received enough information about the study.

☐

I understand that I am free to withdraw from the study at any time, without having to give a reason for withdrawing and without affecting my future medical care.

☐

I understand that data generated by my participation will be stored in a secure electronic database.

☐

I understand that sections of any of my medical notes may be looked at by responsible individuals from either Chelsea and Westminster Hospital or West Middlesex University Hospital or from regulatory authorities where it is relevant to my taking part in this research. I give permission for these individuals to access my records that are relevant to this research. I give permission to other researchers associated with the project to access study data in anonymized format

☐

I agree to take part in this study.

☐

Signed _____ Date _____

NAME IN BLOCK CAPITALS _____

Investigator's signature _____ Date _____

NAME IN BLOCK CAPITALS _____

B.4 INFORMATION SHEETS

This section includes the information sheets to be given to nominated and personal consultees as approved by [REC](#) and [HRA](#).

Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit (CONfuSED)

Information Sheet – Guidance for Nominated Consultees

Invitation

We feel that the patient is unable to decide for himself/herself whether to participate in this research.

To help decide if the patient should join the study, we'd like to ask your opinion on whether or not you think they would have any objection to taking part. We'd ask you to consider what you know of their wishes and feelings, and to consider their interests. Please let us know of any advance decisions they may have made about participating in research. These should take precedence.

If you decide the patient would have no objection to taking part, we will ask you to read and sign the consultee declaration form. We'll then give you a copy to keep. We will keep you fully informed during the study so you can let us know if you have any concerns or you think the patient should be withdrawn.

If you decide that the patient would not wish to take part, it will not affect the standard of care they receive in any way.

If you are unsure about taking the role of consultee you may seek independent advice.

You do not have to act as a consultee, and we will understand if you do not want to take on this responsibility.

What is the purpose of the study?

Delirium, or confusion, is unfortunately very common in Intensive Care. This is in part due to the disease process, the new environment that the patient is in but also due to personal factors like age. If it occurs, it can have an impact on the treatment, length of stay in hospital, and ultimately the development cognitive impairment following discharge from the hospital. It is therefore very important that we diagnose delirium as promptly as possible to facilitate prompt treatment of its underlying cause.

The usual method of diagnosing delirium is through the bedside nurse asking the patient a question and the patient giving an answer. We then score the patient's answers and see whether the patient is delirious. We carry out these tests twice a day. This is very sensitive at diagnosing delirium but is manual, laborious, and only picks up the possibility of delirium once it's performed and thus can miss episodes of delirium.

Our aim in this research study is to find out whether eye tracking, in the form of two cameras placed around the bedspace, can automatically diagnose delirium.

The cameras around the bedspace do not record the patient. They measure out what the patient is looking at and compare that against simulations of non-delirious patients. We hypothesise that patients who are looking at things in the scene in-line with the simulations are

not confused and ones that are not in-line with the simulations are confused; but we don't know for certain, hence the study.

It is important to say that patients' individual results from this study will not be/have not been shared with members of the clinical team who are not involved in the study. The clinical care is separate from this study and the study has no impact on the care the patient will receive/has received.

Why they are eligible?

They have been chosen because they are 18 years of age or older, are in a high dependency/intensive care unit and that we think they have a more than a 20% (one in five) chance of developing delirium.

Does my patient have to take part?

You are under no obligation to say the patient should take part and, even if you do, you are free to withdraw them at any time without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care they receive in any way.

Once the patient is able to give consent themselves, we will seek their consent to remain in the study and consent to use the data already collected.

What will happen to the patient if they take part?

The participation in the study is for the duration of the patient's stay at the intensive care unit. If you agree for them to take part, they will receive the usual clinical care which is not affected by this study at all. This includes twice daily assessments of delirium through the question-answer paradigm. In addition, we will add two cameras to the bed-space, one behind the bedspace looking out to what the patient is seeing and one in-front of them looking at them, both connected to a laptop.

The cameras will be turned off and the lenses covered most of the day. Five minutes prior to the manual assessment of delirium, the cameras will be uncovered and the system switched on. The assessment will then take place and the cameras will continue for a further 5 minutes afterwards.

The cameras do not record. They form a mathematical representation of the eyes and compare what the patient is seeing with what simulations are expecting and we record the similarity.

What do the cameras do?

The camera behind the bed (the scene-camera) looks at the scene in a similar manner to what the patient is seeing. This allows us to run simulations of what most people would look at given the same scene – this is called saliency.

The camera in-front of the patient will look at the patient's face, finds their eyes and measures the direction of their gaze.

Combining the two cameras, we know where they are looking at. Comparing what they are looking at with the simulation is the goal of the system. We do not record the patient at all.

What do they have to do?

The patient will not need to do anything during the course of the study.

What are the possible benefits of taking part?

The analysis of eye tracking in intensive care may help form a continuous measure of delirium for future patients.

What if there is a problem?

Complaints

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions.

Will the patient's involvement in this study be kept confidential?

Chelsea and Westminster Hospital NHS Foundation Trust is the sponsor for this study based in the United Kingdom. We will be using information from you and the patient in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your and the patient's information and using it properly. Chelsea and Westminster Hospital NHS Foundation Trust will keep identifiable information about you and the patient for 5 years after the study has finished.

Your and the patient's rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. To safeguard your and the patient's rights, we will use the minimum personally-identifiable information possible.

How will we use information about the patient?

We will need to use information from the patient for this research project. This information will include their name, surname, date of birth, NHS number, hospital number and contact details. People will use this information to do the research or to check their records to make sure that the research is being done properly.

People who do not need to know who they are will not be able to see their name or contact details. Their data will have a code number instead. We will keep all information about them safe and secure. Once we have finished the study, we will keep some of the data so we can check the results. We will write our reports in a way that no-one can work out that they took part in the study.

What are your choices about how the patient's information is used?

You can stop the patient being part of the study at any time, without giving a reason, but we will keep information about them that we already have. We need to manage their records in specific ways for the research to be reliable. This means that we won't be able to let you, or the patient see or change the data we hold about them.

Where can you find out more about how your information is used?

You can find out more about how we use the patient's information by asking one of the research team (details of the principal investigator are at the bottom of the page)

- at www.hra.nhs.uk/information-about-patients/
- at <https://www.chelwest.nhs.uk/about-us/research-development/gdpr-and-research> or
- by contacting the Data Protection Officer at DPO.Chelwest@nhs.net.

Will the patient's General Practitioner be informed?

As this study is purely a data collection observational study of the patient's time in intensive care, we will not inform the patient's General Practitioners of the patients' participation in this study.

What will happen to the results of the research study?

If you or the patient wish to find out what the final results of the study are, we will be able to contact you and the patient once we have completed it. It does tend to take some time to finalise the results from a study and so you or the patient may not hear from us for over a year. Our plan is to publish the results in an appropriate scientific journal so that others can read about and learn from the trial. Please note that you or the patient will not be identifiable in any publication related to this research.

Who is organising and funding the research?

It has been funded by Westminster Medical School Research Trust and it is being organised by Dr Marcela Vizcaychipi, Consultant in Anaesthesia at Chelsea and Westminster Hospital NHS Foundation Trust, and Dr Ahmed Al-Hindawi, Research Fellow in Anaesthetics also at Chelsea and Westminster Hospital NHS Foundation Trust.

Who has reviewed the study?

It has been peer reviewed by two Consultants in Intensive Care, both of whom have recommended the study be undertaken.

Further information and contact details

If you would like any further information about this trial, please contact the relevant study personnel as per the details below:

Chelsea and Westminster Hospital NHS Foundation Trust

Name: Dr Ahmed Al-Hindawi

Position: Intensive Care Registrar, Magill Research Fellow

Email: ahmed.al-hindawi@nhs.net

Telephone, via Anila Qato, Anaesthetic Secretary: 02033158026

Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit (CONfuSED)

Information Sheet – Guidance for Relative/Friend

Invitation

We feel your relative/friend is unable to decide for himself/herself whether to participate in this research.

To help decide if they should join the study, we'd like to ask your opinion on whether or not they would have any objection to taking part'. We'd ask you to consider what you know of their wishes and feelings, and to consider their interests. Please let us know of any advance decisions they may have made about participating in research. These should take precedence.

If you decide your relative/friend would have no objection to taking part, we will ask you to read and sign the consultee declaration form. We'll then give you a copy to keep. We will keep you fully informed during the study so you can let us know if you have any concerns or you think your relative/friend should be withdrawn.

If you decide that your friend/relative would not wish to take part, it will not affect the standard of care they receive in any way.

If you are unsure about taking the role of consultee you may seek independent advice.

You do not have to act as a consultee, and we will understand if you do not want to take on this responsibility.

What is the purpose of the study?

Delirium, or confusion, is unfortunately very common in Intensive Care. This is in part due to the disease process, the new environment that your relative/friend is in but also due to personal factors, like age. If it occurs, it can have an impact on the treatment, length of stay in hospital, and ultimately the development cognitive impairment following discharge from the hospital. It is therefore very important that we diagnose delirium as promptly as possible to facilitate prompt treatment of its underlying cause.

The usual method of diagnosing delirium is through scoring the answers your friend/relative gives to some questions that their bedside nurse would ask them twice a day as part of their standard of care. This is very sensitive at diagnosing delirium but is manual, laborious, and only picks up the possibility of delirium once it's performed and thus can miss episodes of delirium.

Our aim in this research study is to find out whether eye tracking, in the form of two cameras placed around the bedspace, can automatically diagnose delirium.

The cameras around the bedspace do not record your friend/relative. They measure out what they are looking at and compare that against simulations of non-delirious patients. We hypothesise that patients who are looking at things in the scene in-line with the simulations are

not confused and ones that are not-inline with the simulations are confused; but we don't know for certain, hence the study.

It is important to say that patients' individual results from this study will not be/have not been shared with members of the clinical team who are not involved in the study. The clinical care is separate from this study and the study has no impact on the care that your relative/friend will receive/has received.

Why they are eligible?

They have been chosen because they are 18 years of age or older, are in a high dependency/intensive care unit and have a more than 20% (one in five) chance of developing delirium.

Does my relative/friend have to take part?

You are under no obligation to say your relative/friend should take part and, even if you do, you are free to withdraw them at any time without giving a reason. A decision to withdraw at any time, or a decision not to take part, will not affect the standard of care they receive in any way.

Once your relative/friend is able to give consent themselves we will seek their consent to remain in the study and consent to use the data already collected.

What will happen to my relative/friend if they take part?

The participation in the study is for the duration of your relative's/friend's stay at the intensive care unit. If you agree for them to take part, they will receive the usual clinical care which is not affected by this study at all. This includes twice daily assessments of delirium through the question-answer paradigm. In addition, we will add two cameras to the bed-space, one behind the bedspace looking out to what your relative/friend is seeing and one in-front of them looking at them, both connected to a laptop.

The cameras will be turned off and the lenses covered most of the day. Five minutes prior to the manual assessment of delirium, the cameras will be uncovered and the system switched on. The assessment will then take place and the cameras will continue for a further 5 minutes afterwards.

The cameras do not record. They form a mathematical representation of the eyes and compare what your relative/friend is seeing with what simulations are expecting and we record the similarity.

What do the cameras do?

The camera behind the bed (the scene-camera) looks at the scene in a similar manner to what your friend/relative is seeing. This allows us to run simulations of what most people would look at given the same scene – this is called saliency.

The camera in-front will look at your relative/friend's face, find their eyes and measure the direction of their gaze.

Combining the two cameras, we know what and where your relative/friend is looking at. Comparing what they are looking at with the simulation is the goal of the system. We do not record them at all.

What do they have to do?

Your relative/friend will not need to do anything during the course of the study.

What are the possible benefits of taking part?

The analysis of eye tracking in intensive care may help form a continuous measure of delirium for future patients.

What if there is a problem?

Complaints

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this by contacting the relevant Patient Advice and Liaison Service (PALS) as detailed below:

Chelsea and Westminster Hospital NHS Foundation Trust
Email: cwpals@chelwest.nhs.uk
Telephone: 02033156727

Harm

If your relative/friend is harmed during the research and this is due to someone's negligence, then they may have grounds for a legal action for compensation but they may have to pay their legal costs. The normal National Health Service complaints mechanisms will still be available to them.

Will my relative/friend's involvement in this study be kept confidential?

Chelsea and Westminster Hospital NHS Foundation Trust is the sponsor for this study based in the United Kingdom. We will be using information from your relative/friend and their medical records in order to undertake this study and we will act as the data controller for this study. This means that we are responsible for looking after their information and using it properly. Chelsea and Westminster Hospital NHS Foundation Trust will keep identifiable information about you for 5 years after the study has finished.

Your relative/friend's rights to access, change or move your information are limited, as we need to manage their information in specific ways in order for the research to be reliable and accurate. To safeguard your rights, we will use the minimum personally-identifiable information possible.

How will we use information about relative/friend?

We will need to use information from your relative/friend for this research project. This information will include their name, surname, date of birth, NHS number, hospital number and contact details. People will use this information to do the research or to check their records to make sure that the research is being done properly.

People who do not need to know who they are will not be able to see their name or contact details. Their data will have a code number instead. We will keep all information about them safe and secure. Once we have finished the study, we will keep some of the data so we can check the results. We will write our reports in a way that no-one can work out that they took part in the study.

What are your choices about how your relative/friend information is used?

You can stop your relative/friend being part of the study at any time, without giving a reason, but we will keep information about them that we already have. We need to manage their records in specific ways for the research to be reliable. This means that we won't be able to let you, or your relative/friend see or change the data we hold about them.

Where can you find out more about how your information is used?

You can find out more about how we use your relative/friend's information by asking one of the research team (details of the principal investigator are at the bottom of the page)

- at www.hra.nhs.uk/information-about-patients/
- at <https://www.chelwest.nhs.uk/about-us/research-development/gdpr-and-research> or
- by contacting the Data Protection Officer at DPO.Chelwest@nhs.net.

Will my relative/friend's General Practitioner be informed?

As this study is purely a data collection observational study of your relative/friend's time in intensive care, we will not inform their General Practitioner of their participation in this study.

What will happen to the results of the research study?

If your relative/friend wishes to find out what the final results of the study are, we will be able to contact them once we have completed it. It does tend to take some time to finalise the results from a study and so they may not hear from us for over a year. Our plan is to publish the results in an appropriate scientific journal so that others can read about and learn from the trial. Please note that they will not be identifiable in any publication related to this research.

Who is organising and funding the research?

It has been funded by the Chelsea & Westminster Hospital NHS Foundation Trust and Westminster Research Fund.

Who has reviewed the study?

It has been peer reviewed by two Consultants in Intensive Care, both of whom have recommended the study be undertaken.

Further information and contact details

If you would like any further information about this trial, please contact the relevant study personnel as per the details below:

Chelsea and Westminster Hospital NHS Foundation Trust
Name: Dr Ahmed Al-Hindawi
Position: Intensive Care Registrar, Magill Research Fellow
Email: ahmed.al-hindawi@nhs.net
Telephone, via Anila Qato, Anaesthetic Secretary: 02033158026

B.5 PATIENT INFORMATION SHEETS

This section includes the patient information sheets as approved by [REC](#) and [HRA](#).

Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit (CONfuSED)

Information Sheet – Guidance for Patients Regaining Capacity

Invitation

You are currently taking part in a new research study. Earlier in your admission we felt that you were unable to decide for yourself whether to participate in this research. As such, we asked your relative/friend for their opinion and they advised us that, in their opinion, they thought you would have no objection to taking part.

Before you decide whether to continue, it is important for you to understand why the research is being done and what it involves. Please take time to read the following information carefully and discuss it with a friend or relative if you wish. Ask us if there is anything that is not clear or if you would like more information. Please take time to decide whether or not you wish to continue taking part.

What is the purpose of the study?

Delirium, or confusion, is unfortunately very common in Intensive Care. This is in part due to your disease process, the new environment that you are in but also personal factors, like age. If it occurs, it can have an impact on your treatment, length of stay in hospital, and ultimately whether you develop cognitive impairment following your discharge from the hospital. It is therefore very important that we diagnose delirium as promptly as possible to facilitate prompt treatment of its underlying cause.

The usual method of diagnosing delirium is through scoring the answers you give to some questions that your bedside nurse will ask you twice a day as part of your standard of care.. This is very sensitive at diagnosing delirium but is manual, laborious, and only picks up the possibility of delirium once it's performed and thus can miss episodes of delirium.

Our aim in this research study is to find out whether eye tracking, in the form of two cameras placed around the bedspace, can automatically diagnose delirium.

The cameras around your bedspace do not record you. They measure out what you are looking at around you and compare what you are looking at against simulations of non-delirious patients. We hypothesise that patients who are looking at things in the scene in-line with the simulations are not confused and the ones that are looking at things which are not in-line with the simulations are confused; but we don't know for certain, hence the study.

It is important to say that patients' individual results from this study will not be/have not been shared with members of your clinical team who are not involved in the study. Your clinical care is separate from this study and the study has had no impact on the care you have received.

Why have I been invited?

You have been chosen because you are 18 years of age or older, are in a high dependency /intensive care unit and we think you have a more than a 20% (one in five) chance of developing delirium.

We gave an information sheet, similar to this one, to your relative/friend earlier in your admission to hospital. On your behalf, your relative/friend advised that you would have wished to enter this study. If no relative was found, a nominated consultee was asked for their advice on your views and feelings about taking part in the study. In either case, this was recorded in a Consultee Declaration Form

Do I have to continue to take part?

At this point, it is entirely your decision whether or not you continue to take part. You are under no obligation to do so and you are free to withdraw at any time without giving a reason. A decision to withdraw at any time, or a decision not to continue taking part, will not affect the standard of care you receive.

If you decide that you would like to continue to take part, you will be given this information sheet to keep and be asked to sign a consent form after you have discussed any questions you may still have with one of the investigators.

What will happen to me if I continue to take part?

The participation in the study is for your stay at the intensive care unit. We already taken some non-invasive measurements of your eyes and what you are looking at using the two cameras described above in a purely observational way.

If you agree to continue to take part, we will continue to collect data, and use the data previously collected within the study.

What do the cameras do?

The camera behind the bed (the scene-camera) looks at the scene in a similar manner to what you're seeing. This allows us to run simulations of what most people would look at given the same scene – this is called saliency.

The camera in-front of you will look at your face, find your eyes and measure the direction of your gaze.

Combining the two cameras, we know where you are looking at. Comparing what you are looking at with the simulation is the goal of the system. We do not record you at all.

What do I have to do?

You will not need to do anything during the course of the study.

What are the possible benefits of taking part?

The analysis of eye tracking in intensive care may help form a continuous measure of delirium for future patients.

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence.

What will happen if I do not want to carry on with the study?

If you decide to withdraw from the study, you can do so and the standard of medical care you receive will not be affected. We will stop the study and no further measurements will occur. You may withdraw the data already collected while you were lacking capacity, but we encourage you not to as it is important for the results of the study.

What if there is a problem?

Complaints

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this by contacting the relevant Patient Advice and Liaison Service (PALS) as detailed below:

Chelsea and Westminster Hospital NHS Foundation Trust

Email: cwpals@chelwest.nhs.uk

Telephone: 02033156727

Harm

If you are harmed during the research and this is due to someone's negligence, then you may have grounds for a legal action for compensation but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Will my taking part in this study be kept confidential?

Chelsea and Westminster Hospital NHS Foundation Trust is the sponsor for this study based in the United Kingdom. We will be using information from you and your medical records in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly. Chelsea and Westminster Hospital NHS Foundation Trust will keep identifiable information about you for at least 5 years after the study has finished.

Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. To safeguard your rights, we will use the minimum personally-identifiable information possible.

How will we use information about you?

We will need to use information from you for this research project. This information will include your name, surname, date of birth, NHS number, hospital number and contact details. People will use this information to do the research or to check your records to make sure that the research is being done properly.

People who do not need to know who you are will not be able to see your name or contact details. Your data will have a code number instead. We will keep all information about you safe and secure. Once we have finished the study, we will keep some of the data so we can check the results. We will write our reports in a way that no-one can work out that you took part in the study.

What are your choices about how your information is used?

You can stop being part of the study at any time, without giving a reason, but we will keep information about you that we already have. We need to manage your records in specific ways for the research to be reliable. This means that we won't be able to let you see or change the data we hold about you.

Where can you find out more about how your information is used?

You can find out more about how we use your information by asking one of the research team (details of the principal investigator are at the bottom of the page)

- at www.hra.nhs.uk/information-about-patients/
- at <https://www.chelwest.nhs.uk/about-us/research-development/gdpr-and-research>
- by contacting the Data Protection Officer at DPO.Chelwest@nhs.net.

Will my General Practitioner be informed?

As this study is purely a data collection study of your time in intensive care, we will not inform your General Practitioner of your participation in this study.

What will happen to the results of the research study?

If you wish to find out what the final results of the study are, we will be able to contact you once we have completed it. It does tend to take some time to finalise the results from a study and so you may not hear from us for over a year. Our plan is to publish the results in an appropriate scientific journal so that others can read about and learn from the trial. Please note that you will not be identifiable in any publication related to this research.

Who is organising and funding the research?

It has been funded by Westminster Medical School Research Trust and it is being organised by Dr Marcela Vizcaychipi, Consultant in Anaesthesia at Chelsea and Westminster Hospital NHS Foundation Trust, and Dr Ahmed Al-Hindawi, Research Fellow in Anaesthetics also at Chelsea and Westminster Hospital NHS Foundation Trust.

Who has reviewed the study?

It has been peer reviewed by two Consultants in Intensive Care, both of whom have recommended the study be undertaken.

How will we use information about you?

We will need to use information from your medical records for this research project. This information will include your NHS number held by Chelsea & Westminster Hospital NHS Foundation Trust. People will use this information to do the research or to check your records to make sure that the research is being done properly. People who do not need to know who you are will not be able to see your name or contact details. Your data will have a code number instead. We will keep all information about you safe and secure. Once we have finished the study, we will keep some of the data so we can check the results. We will write our reports in a way that no-one can work out that you took part in the study.

Further information and contact details

If you would like any further information about this trial, please contact the relevant study personnel as per the details below:

Chelsea and Westminster Hospital NHS Foundation Trust
Name: Dr Ahmed Al-Hindawi
Position: Intensive Care Registrar, Magill Research Fellow
Email: ahmed.al-hindawi@nhs.net
Telephone, via Anila Qato, Anaesthetic Secretary: 02033158026

Continuous Non-Invasive Eye Tracking for the Early Detection of Delirium on the Intensive Care Unit (CONfuSED)

Participant Information Sheet – Adults with Capacity

Invitation

It is our pleasure to invite you to participate in a research study that is being undertaken at Chelsea and Westminster Hospital NHS Foundation Trust where we are hoping to find out how to best diagnose delirium in Intensive Care.

Before you decide whether to participate or not, we would like you to understand why the research is being done and what it would involve for you. One of our team will go through the information sheet with you. Please discuss it with others if you wish to do so, feel free to ask us any questions and take the time to decide whether or not you want to be involved.

Part 1

What is the purpose of the study?

Delirium, or confusion, is unfortunately very common in Intensive Care. This is in part due to your disease process, the new environment that you are in but also personal factors, like age. If it occurs, it can have an impact on your treatment, length of stay in hospital, and ultimately whether you develop cognitive impairment following your discharge from the hospital. It is therefore very important that we diagnose delirium as promptly as possible to facilitate prompt treatment of its underlying cause

The usual method of diagnosing delirium is through scoring the answers you give to some questions that your bedside nurse will ask you twice a day as part of your standard of care. This is very sensitive at diagnosing delirium but is manual, laborious, and only picks up the possibility of delirium once it's performed and thus can miss episodes of delirium.

Our aim in this research study is to find out whether eye tracking, in the form of two cameras placed around the bedspace, can automatically diagnose delirium.

The cameras around your bedspace do not record you. They measure out what you are looking at around you and compare what you are looking at against simulations of non-delirious patients. We hypothesise that patients who are looking at things in the scene in-line with the simulations are not confused and ones that are not-inline with the simulations are confused; but we don't know for certain, hence the study.

It is important to say that patients' individual results from this study will not be shared with members of your clinical team who are not involved in the study. Your clinical care is separate from this study and the study will have no impact on the care you receive.

Why have I been invited?

You have been chosen because you are 18 years of age or older, are in a high dependency /intensive care unit and that we think you have a more than a 20% (one in five) chance of developing delirium.

Do I have to take part?

Your participation in this research is completely voluntary and it is up to you to decide whether or not to take part. We will describe the study and go through this information sheet with you. If you agree to take part, we will then ask you to sign a consent form. You are free to withdraw at any time, without giving a reason. If you would prefer not to participate or withdraw from the trial at a later date, the standard of medical care you receive will not be affected.

What will happen to me if I take part?

You will receive the usual clinical care which is not affected by this study at all. This includes twice daily assessments of delirium (which consists of questions asked by your bedside nurse). In addition, we will add two cameras to the bed-space, one behind you looking out to what you're seeing and one in-front of you looking at you, both connected to a laptop.

The cameras will be turned off and the lenses covered most of the day. Five minutes prior to the manual assessment of delirium, the cameras will be uncovered and the system switched on. The assessment will then take place and the cameras will continue for a further 5 minutes afterwards. The cameras do not record. They form a mathematical representation of your eyes and compare what you are seeing with what simulations are expecting and we record the similarity.

If you lose capacity during the study (for example because you developed delirium), we will ask a personal consultee or a nominated consultee to make sure that it is still your wish to continue taking part in the study.

What do the cameras do?

The camera behind the bed (the scene-camera) looks at the scene in a similar manner to what you're seeing. This allows us to run simulations of what most people would look at given the same scene – this is called saliency.

The camera in-front of you will look at your face, find your eyes and measure the direction of your gaze.

Combining the two cameras, we know where you are looking at. Comparing what you are looking at with the simulation is the goal of the system. We do not record you at all.

What will I have to do?

If you agree to participate in this trial, there is no work to be done on your part – the cameras will be setup in the bed space and will be present till your discharge or if you wish to no longer take part in the study.

If you do become delirious during your stay in ICU and thus lose capacity, we will continue to use the cameras throughout your stay in conjunction with normal clinical tests.

What are the possible benefits of taking part?

The analysis of eye tracking in intensive care may help form a continuous measure of delirium for future patients.

What if there is a problem?

Any complaint about the way you have been dealt with during the study or any possible harm you might suffer will be addressed. Detailed information on this is included in Part 2.

Will my taking part in the study be kept confidential?

Yes. We will follow ethical and legal practice and all information about you will be handled in confidence. Details are provided in Part 2.

This completes Part 1. If the information in Part 1 has interested you and you are considering participation, please read the additional information in Part 2 before making any decision.

Part 2

What if relevant new information becomes available?

Sometimes we receive new information about the treatment being studied. If this happens, your research doctor will tell you and discuss whether you should continue in the study. If you decide not to carry on, your research doctor will make arrangements for your care to continue. If you decide to continue in the study, he or she may ask you to sign an agreement outlining the discussion.

What will happen if I do not want to carry on with the study?

If you decide to withdraw from the study, you can do so and the standard of medical care you receive will not be affected. Information collected up until the point you withdraw from the study will still be used.

What if there is a problem?

Complaints

If you have a concern about any aspect of this study, you should ask to speak to the researchers who will do their best to answer your questions. If you remain unhappy and wish to complain formally, you can do this by contacting the relevant Patient Advice and Liaison Service (PALS) as detailed below:

Chelsea and Westminster Hospital NHS Foundation Trust
Email: cwpals@chelwest.nhs.uk
Telephone: 02033156727

Harm

If you are harmed during the research and this is due to someone's negligence, then you may have grounds for a legal action for compensation but you may have to pay your legal costs. The normal National Health Service complaints mechanisms will still be available to you.

Will my taking part in this study be kept confidential?

Chelsea and Westminster Hospital NHS Foundation Trust is the sponsor for this study based in the United Kingdom. We will be using information from you and your medical records in order to undertake this study and will act as the data controller for this study. This means that we are responsible for looking after your information and using it properly. Chelsea and Westminster Hospital NHS Foundation Trust will keep identifiable information about you for 5 years after the study has finished.

Your rights to access, change or move your information are limited, as we need to manage your information in specific ways in order for the research to be reliable and accurate. To safeguard your rights, we will use the minimum personally-identifiable information possible.

How will we use information about you?

We will need to use information from you for this research project. This information will include your name, surname, date of birth, NHS number, hospital number and contact details. People will use this information to do the research or to check your records to make sure that the research is being done properly.

People who do not need to know who you are will not be able to see your name or contact details. Your data will have a code number instead. We will keep all information about you safe and secure. Once we have finished the study, we will keep some of the data so we can check the results. We will write our reports in a way that no-one can work out that you took part in the study.

What are your choices about how your information is used?

You can stop being part of the study at any time, without giving a reason, but we will keep information about you that we already have. We need to manage your records in specific ways for the research to be reliable. This means that we won't be able to let you see or change the data we hold about you.

Where can you find out more about how your information is used?

You can find out more about how we use your information by asking one of the research team (details of the principal investigator are at the bottom of the page)

- at www.hra.nhs.uk/information-about-patients/
- at <https://www.chelwest.nhs.uk/about-us/research-development/gdpr-and-research> or
- by contacting the Data Protection Officer at DPO.Chelwest@nhs.net.

Will my General Practitioner be informed?

As this study is purely a data collection study of your time in intensive care, we will not inform your General Practitioner of your participation in this study.

What will happen to the results of the research study?

If you wish to find out what the final results of the study are, we will be able to contact you once we have completed it. It does tend to take some time to finalise the results from a study and so you may not hear from us for over a year. Our plan is to publish the results in an appropriate scientific journal so that others can read about and learn from the trial. Please note that you will not be identifiable in any publication related to this research.

Who is organising and funding the research?

It has been funded by Westminster Medical School Research Trust and it is being organised by Dr Marcela Vizcaychipi, Consultant in Anaesthesia at Chelsea and Westminster Hospital NHS Foundation Trust, and Dr Ahmed Al-Hindawi, Research Fellow in Anaesthetics also at Chelsea and Westminster Hospital NHS Foundation Trust.

Who has reviewed the study?

It has been peer reviewed by two Consultants in Intensive Care, both of whom have recommended the study be undertaken.

Further information and contact details

If you would like any further information about this trial, please contact the relevant study personnel as per the details below:

Chelsea and Westminster Hospital NHS Foundation Trust
Name: Dr Ahmed Al-Hindawi
Position: Intensive Care Registrar, Magill Research Fellow
Email: ahmed.al-hindawi@nhs.net
Telephone, via Anila Qato, Anaesthetic Secretary: 02033158026

BIBLIOGRAPHY

- Abdulaal, Ahmed, Aatish Patel, Esmiya Charani, Sarah Denny, Nabeela Mughal, and Luke Moore (2020). "Prognostic Modeling of COVID-19 Using Artificial Intelligence in the United Kingdom: Model Development and Validation." In: *Journal of Medical Internet Research* 22.8, e20259. DOI: [10.2196/20259](https://doi.org/10.2196/20259). URL: <https://www.jmir.org/2020/8/e20259/> (visited on 10/09/2020) (cit. on p. 32).
- Abel, L. A., S. Levin, and P. S. Holzman (June 1992). "Abnormalities of Smooth Pursuit and Saccadic Control in Schizophrenia and Affective Disorders." In: *Vision Research* 32.6, pp. 1009–1014. ISSN: 0042-6989. DOI: [10.1016/0042-6989\(92\)90002-Z](https://doi.org/10.1016/0042-6989(92)90002-Z). URL: <https://www.sciencedirect.com/science/article/pii/004269899290002Z> (visited on 04/20/2022) (cit. on p. 20).
- Al-Hindawi, A, Jagdish Sokhi, Alfred Hill, Alice Sisson, and Marcela Vizcaychipi (Dec. 2020). "Delirium Prediction in COVID-19 Intensive Care Unit: A Bayesian Model with Informed Priors (DELIPIOUS)." In: *Intensive Care Medicine Experimental*. 12th ser. 8.S2, p. 62. ISSN: 2197-425X. DOI: [10.1186/s40635-020-00354-8](https://doi.org/10.1186/s40635-020-00354-8). URL: <https://icm-experimental.springeropen.com/articles/10.1186/s40635-020-00354-8> (visited on 03/15/2021) (cit. on p. 28).
- Al-Hindawi, Ahmed, Marcela P Vizcaychipi, and Yiannis Demiris (Nov. 2021). "Continuous Non-Invasive Eye Tracking In Intensive Care." In: *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pp. 1869–1873. DOI: [10.1109/EMBC46164.2021.9630999](https://doi.org/10.1109/EMBC46164.2021.9630999) (cit. on pp. 73, 82, 88).
- Amadori, Pierluigi Vito, Tobias Fischer, Ruohan Wang, and Yiannis Demiris (Sept. 2020). "Decision Anticipation for Driving Assistance Systems." In:

- 2020 *IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, pp. 1–7. DOI: [10.1109/ITSC45102.2020.9294216](https://doi.org/10.1109/ITSC45102.2020.9294216) (cit. on pp. 41, 73).
- An, Gwon Hwan, Siyeong Lee, Min-Woo Seo, Kugjin Yun, Won-Sik Cheong, and Suk-Ju Kang (Dec. 2018). “Charuco Board-Based Omnidirectional Camera Calibration Method.” In: *Electronics* 7.12, p. 421. DOI: [10.3390/electronics7120421](https://doi.org/10.3390/electronics7120421). URL: <https://www.mdpi.com/2079-9292/7/12/421> (visited on 07/01/2020) (cit. on p. 43).
- Archibald, Neil K., Sam B. Hutton, Michael P. Clarke, Urs P. Mosimann, and David J. Burn (Mar. 2013). “Visual Exploration in Parkinson’s Disease and Parkinson’s Disease Dementia.” In: *Brain* 136.3, pp. 739–750. ISSN: 0006-8950. DOI: [10.1093/brain/awt005](https://doi.org/10.1093/brain/awt005). URL: <https://doi.org/10.1093/brain/awt005> (visited on 06/17/2021) (cit. on p. 19).
- Bai, Shaojie, J. Zico Kolter, and Vladlen Koltun (Apr. 2018). “An Empirical Evaluation of Generic Convolutional and Recurrent Networks for Sequence Modeling.” In: *arXiv:1803.01271 [cs]*. arXiv: [1803.01271 \[cs\]](https://arxiv.org/abs/1803.01271). URL: <http://arxiv.org/abs/1803.01271> (visited on 09/21/2021) (cit. on pp. 23, 97, 99).
- Bakker, Robbert C., Robert Jan Osse, Joke H.M. Tulen, A. Pieter Kappetein, and Ad J.J.C. Bogers (Mar. 2012). “Preoperative and Operative Predictors of Delirium after Cardiac Surgery in Elderly Patients.” In: *European Journal of Cardio-Thoracic Surgery* 41.3, pp. 544–549. ISSN: 1010-7940. DOI: [10.1093/ejcts/ezr031](https://doi.org/10.1093/ejcts/ezr031). URL: <https://doi.org/10.1093/ejcts/ezr031> (visited on 06/16/2022) (cit. on p. 14).
- Belkin, Mikhail, Daniel Hsu, Siyuan Ma, and Soumik Mandal (Aug. 2019). “Reconciling Modern Machine-Learning Practice and the Classical Bias–Variance Trade-Off.” In: *Proceedings of the National Academy of Sciences* 116.32, pp. 15849–15854. ISSN: 0027-8424, 1091-6490. DOI: [10.1073/pnas.1903070116](https://doi.org/10.1073/pnas.1903070116). URL: <https://www.pnas.org/content/116/32/15849> (visited on 01/25/2022) (cit. on p. 38).

- Bengio, Yoshua, Jérôme Louradour, Ronan Collobert, and Jason Weston (2009). "Curriculum Learning." In: *Proceedings of the 26th Annual International Conference on Machine Learning - ICML '09*. Montreal, Quebec, Canada: ACM Press, pp. 1–8. ISBN: 978-1-60558-516-1. DOI: [10.1145/1553374.1553380](https://doi.org/10.1145/1553374.1553380). URL: <http://portal.acm.org/citation.cfm?doid=1553374.1553380> (visited on 03/26/2021) (cit. on p. 56).
- Bergstra, James and Yoshua Bengio (Feb. 2012). "Random Search for Hyper-Parameter Optimization." In: *The Journal of Machine Learning Research* 13.null, pp. 281–305. ISSN: 1532-4435 (cit. on pp. 33, 100).
- Berndt, Donald J. and James Clifford (July 1994). "Using Dynamic Time Warping to Find Patterns in Time Series." In: *Proceedings of the 3rd International Conference on Knowledge Discovery and Data Mining*. AAAIWS'94. Seattle, WA: AAAI Press, pp. 359–370 (cit. on p. 139).
- Bertalmio, M., A.L. Bertozzi, and G. Sapiro (2001). "Navier-Stokes, Fluid Dynamics, and Image and Video Inpainting." In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. CVPR 2001. Vol. 1. Kauai, HI, USA: IEEE Comput. Soc, pp. I–355–I–362. ISBN: 978-0-7695-1272-3. DOI: [10.1109/CVPR.2001.990497](https://doi.org/10.1109/CVPR.2001.990497). URL: <http://ieeexplore.ieee.org/document/990497/> (visited on 07/19/2021) (cit. on p. 78).
- Boisvert, Jonathan F.G. and Neil D.B. Bruce (Sept. 2016). "Predicting Task from Eye Movements: On the Importance of Spatial Distribution, Dynamics, and Image Features." In: *Neurocomputing* 207, pp. 653–668. ISSN: 09252312. DOI: [10.1016/j.neucom.2016.05.047](https://doi.org/10.1016/j.neucom.2016.05.047). URL: <https://www.sciencedirect.com/science/article/pii/S0925231216304131?via%3Dihub> (visited on 12/13/2017) (cit. on p. 108).
- Borji, Ali and Laurent Itti (2013). "State-of-the-Art in Visual Attention Modeling." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.1, pp. 185–207. ISSN: 01628828. DOI: [10.1109/TPAMI.2012.89](https://doi.org/10.1109/TPAMI.2012.89) (cit. on p. 108).

- Borji, Ali, Andreas Lennartz, and Marc Pomplun (2015). "What Do Eyes Reveal about the Mind?. Algorithmic Inference of Search Targets from Fixations." In: *Neurocomputing* 149.PB, pp. 788–799. ISSN: 18728286. DOI: [10.1016/j.neucom.2014.07.055](https://doi.org/10.1016/j.neucom.2014.07.055). URL: <http://dx.doi.org/10.1016/j.neucom.2014.07.055> (cit. on p. 108).
- Brier, G (1950). "Verification of Forecasts Expressed in Terms of Probability." In: *Monthly Weather Report* 78, pp. 1–3. DOI: [10.1175/1520-0493\(1950\)078<0001:V0FEIT>2.0.CO;2](https://doi.org/10.1175/1520-0493(1950)078<0001:V0FEIT>2.0.CO;2) (cit. on p. 24).
- Bristow, Davina, John-Dylan Haynes, Richard Sylvester, Christopher D. Frith, and Geraint Rees (July 2005). "Blinking Suppresses the Neural Response to Unchanging Retinal Stimulation." In: *Current Biology* 15.14, pp. 1296–1300. ISSN: 0960-9822. DOI: [10.1016/j.cub.2005.06.025](https://doi.org/10.1016/j.cub.2005.06.025). URL: <https://www.sciencedirect.com/science/article/pii/S0960982205006603> (visited on 03/26/2021) (cit. on p. 65).
- Brockmann, Dirk and Theo Geisel (June 2000). "The Ecology of Gaze Shifts." In: *Neurocomputing* 32–33, pp. 643–650. ISSN: 0925-2312. DOI: [10.1016/S0925-2312\(00\)00227-7](https://doi.org/10.1016/S0925-2312(00)00227-7). URL: <https://www.sciencedirect.com/science/article/pii/S0925231200002277> (visited on 07/28/2021) (cit. on p. 73).
- Browne, Michael W (Mar. 2000). "Cross-Validation Methods." In: *Journal of Mathematical Psychology* 44.1, pp. 108–132. ISSN: 0022-2496. DOI: [10.1006/jmps.1999.1279](https://doi.org/10.1006/jmps.1999.1279). URL: <http://www.sciencedirect.com/science/article/pii/S0022249699912798> (visited on 12/22/2020) (cit. on pp. 32, 100).
- Cerejeira, Joaquim, Horácio Firmino, Adriano Vaz-Serra, and Elizabeta B Mukaetova-Ladinska (June 2010). "The Neuroinflammatory Hypothesis of Delirium." In: *Acta Neuropathologica* 119.6, pp. 737–754. ISSN: 0001-6322. DOI: [10.1007/s00401-010-0674-1](https://doi.org/10.1007/s00401-010-0674-1). URL: <http://www.ncbi.nlm.nih.gov/pubmed/20309566> (visited on 07/11/2016) (cit. on p. 28).

- Chambers, J. M. (Dec. 2017). *Graphical Methods for Data Analysis*. New York: Chapman and Hall/CRC. ISBN: 978-1-351-07230-4. DOI: [10.1201/9781351072304](https://doi.org/10.1201/9781351072304) (cit. on p. 84).
- Chan, Cynthia Y. H., Antoni B. Chan, Tatia M. C. Lee, and Janet H. Hsiao (Dec. 2018). "Eye-Movement Patterns in Face Recognition Are Associated with Cognitive Decline in Older Adults." In: *Psychonomic Bulletin & Review* 25.6, pp. 2200–2207. ISSN: 1069-9384, 1531-5320. DOI: [10.3758/s13423-017-1419-0](https://doi.org/10.3758/s13423-017-1419-0). URL: <http://link.springer.com/10.3758/s13423-017-1419-0> (visited on 04/19/2022) (cit. on p. 19).
- Collette, Andrew (2013). *Python and HDF5: Unlocking Scientific Data*. "O'Reilly Media, Inc." (cit. on p. 95).
- Cortacero, Kevin, Tobias Fischer, and Yiannis Demiris (2019). "RT-BENE: A Dataset and Baselines for Real-Time Blink Estimation in Natural Environments." In: *Proceedings of the IEEE International Conference on Computer Vision Workshops*. URL: https://openaccess.thecvf.com/content_ICCVW_2019/html/GAZE/Cortacero_RT-BENE_A_Dataset_and_Baselines_for_Real-Time_Blink_Estimation_in_ICCVW_2019_paper.html (visited on 07/01/2020) (cit. on pp. 46, 55, 58, 65, 68, 73).
- Crawford, Trevor J. (2015). "The Disengagement of Visual Attention in Alzheimer's Disease: A Longitudinal Eye-Tracking Study." In: *Frontiers in Aging Neuroscience* 7. ISSN: 1663-4365. DOI: [10.3389/fnagi.2015.00118](https://doi.org/10.3389/fnagi.2015.00118). URL: <https://www.frontiersin.org/articles/10.3389/fnagi.2015.00118/full> (visited on 06/17/2021) (cit. on p. 19).
- Cristino, Filipe, Sebastiaan Mathôt, Jan Theeuwes, and Iain D. Gilchrist (Aug. 2010). "ScanMatch A Novel Method for Comparing Fixation Sequences." In: *Behavior Research Methods* 42.3, pp. 692–700. ISSN: 1554-351X. DOI: [10.3758/BRM.42.3.692](https://doi.org/10.3758/BRM.42.3.692). URL: <http://www.springerlink.com/index/10.3758/BRM.42.3.692> (visited on 02/14/2017) (cit. on p. 137).

- Cubuk, Ekin D., Barret Zoph, Jonathon Shlens, and Quoc V. Le (2020). "Randaugment: Practical Automated Data Augmentation With a Reduced Search Space." In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, pp. 702–703. URL: https://openaccess.thecvf.com/content_CVPRW_2020/html/w40/Cubuk_Randaugment_Practical_Automated_Data_Augmentation_With_a_Reduced_Search_Space_CVPRW_2020_paper.html (visited on 03/25/2021) (cit. on p. 55).
- Daffner, K. R., L.F.M. Scinto, S. Weintraub, J. E. Guinessey, and M. M. Mesulam (Feb. 1992). "Diminished Curiosity in Patients with Probable Alzheimer's Disease as Measured by Exploratory Eye Movements." In: *Neurology* 42.2, pp. 320–320. ISSN: 0028-3878, 1526-632X. DOI: [10.1212/WNL.42.2.320](https://doi.org/10.1212/WNL.42.2.320). URL: <http://www.neurology.org/lookup/doi/10.1212/WNL.42.2.320> (visited on 06/17/2021) (cit. on p. 19).
- Dao, Tri, Albert Gu, Alexander Ratner, Virginia Smith, Chris De Sa, and Christopher Re (May 2019). "A Kernel Theory of Modern Data Augmentation." In: *Proceedings of the 36th International Conference on Machine Learning*. PMLR, pp. 1528–1537. URL: <https://proceedings.mlr.press/v97/dao19b.html> (visited on 09/28/2021) (cit. on p. 56).
- Dawson-Howe, Kenneth M. and David Vernon (1994). "Simple Pinhole Camera Calibration." In: *International Journal of Imaging Systems and Technology* 5.1, pp. 1–6. ISSN: 1098-1098. DOI: [10.1002/ima.1850050102](https://doi.org/10.1002/ima.1850050102). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/ima.1850050102> (visited on 07/28/2021) (cit. on pp. 78, 81).
- Demiris, Yiannis and Bassam Khadhour (May 2006). "Hierarchical Attentive Multiple Models for Execution and Recognition of Actions." In: *Robotics and Autonomous Systems* 54.5, pp. 361–369. ISSN: 09218890. DOI: [10.1016/j.robot.2006.02.003](https://doi.org/10.1016/j.robot.2006.02.003). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0921889006000169> (cit. on p. 109).

- Dewhurst, Richard, Marcus Nyström, Halszka Jarodzka, Tom Foulsham, Roger Johansson, and Kenneth Holmqvist (2012). "It Depends on How You Look at It: Scanpath Comparison in Multiple Dimensions with MultiMatch, a Vector-Based Approach." In: *Behavior Research Methods*, pp. 1079–1100. ISSN: 1554-3528. DOI: [10.3758/s13428-012-0212-2](https://doi.org/10.3758/s13428-012-0212-2) (cit. on p. 138).
- Ely, E Wesley, Brenda Truman, Donna J Manzi, Jeffrey C Sigl, Ayumi Shintani, and Gordon R Bernard (May 2004). "Consciousness Monitoring in Ventilated Patients: Bispectral EEG Monitors Arousal Not Delirium." In: *Intensive care medicine* 30.8, pp. 1537–43. ISSN: 0342-4642. DOI: [10.1007/s00134-004-2298-1](https://doi.org/10.1007/s00134-004-2298-1). URL: <http://www.ncbi.nlm.nih.gov/pubmed/15127189> (cit. on pp. 12, 13).
- Ely, E. Wesley et al. (Dec. 2001a). "Delirium in Mechanically Ventilated Patients: Validity and Reliability of the Confusion Assessment Method for the Intensive Care Unit (CAM-ICU)." In: *JAMA* 286.21, p. 2703. ISSN: 0098-7484. DOI: [10.1001/jama.286.21.2703](https://doi.org/10.1001/jama.286.21.2703). URL: <http://jama.jamanetwork.com/article.aspx?doi=10.1001/jama.286.21.2703> (visited on 05/20/2019) (cit. on pp. 2, 29, 88).
- Ely, E., S. Gautam, R. Margolin, J. Francis, L. May, T. Speroff, B. Truman, R. Dittus, G. Bernard, and S. Inouye (Dec. 2001b). "The Impact of Delirium in the Intensive Care Unit on Hospital Length of Stay." In: *Intensive Care Medicine* 27.12, pp. 1892–1900. ISSN: 1432-1238. DOI: [10.1007/s00134-001-1132-2](https://doi.org/10.1007/s00134-001-1132-2). URL: <https://doi.org/10.1007/s00134-001-1132-2> (visited on 05/22/2020) (cit. on pp. 1, 14).
- European Delirium Association and American Delirium Society (Sept. 2014). "The DSM-5 Criteria, Level of Arousal and Delirium Diagnosis: Inclusiveness Is Safer." In: *BMC Medicine* 12.1, p. 141. ISSN: 1741-7015. DOI: [10.1186/s12916-014-0141-2](https://doi.org/10.1186/s12916-014-0141-2). URL: <https://doi.org/10.1186/s12916-014-0141-2> (visited on 03/26/2022) (cit. on pp. 1, 2, 11, 12).

- Exton, C. and M. Leonard (2009). "Eye Tracking Technology: A Fresh Approach in Delirium Assessment?" In: *International Review of Psychiatry* 21.1, pp. 8–14. ISSN: 1369-1627. DOI: 908706022[pil]\r10.1080/09540260802675106. URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=19219708 (visited on 06/29/2016) (cit. on pp. 41, 129).
- Fang, Yi, Jiapeng Tang, Wang Shen, Wei Shen, Xiao Gu, Li Song, and Guangtao Zhai (2021). "Dual Attention Guided Gaze Target Detection in the Wild." In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11390–11399. URL: https://openaccess.thecvf.com/content/CVPR2021/html/Fang_Dual_Attention_Guided_Gaze_Target_Detection_in_the_Wild_CVPR_2021_paper.html (visited on 07/24/2021) (cit. on pp. 74, 85).
- Feng, Gary (Mar. 2006). "Eye Movements as Time-Series Random Variables: A Stochastic Model of Eye Movement Control in Reading." In: *Cognitive Systems Research* 7.1, pp. 70–95. ISSN: 13890417. DOI: 10.1016/j.cogsys.2005.07.004. URL: <http://linkinghub.elsevier.com/retrieve/pii/S1389041705000276> (cit. on p. 42).
- Figiel, Gary S., K. Ranga Krishnan, John C. Breitner, and Charles B. Nemeroff (1989). "Radiologic Correlates of Antidepressant-Induced Delirium: The Possible Significance of Basal-Ganglia Lesions." In: *The Journal of Neuropsychiatry and Clinical Neurosciences* 1.2, pp. 188–190. ISSN: 1545-7222. DOI: 10.1176/jnp.1.2.188 (cit. on p. 123).
- Fischer, Tobias, Hyung Jin Chang, and Yiannis Demiris (2018). "RT-GENE: Real-Time Eye Gaze Estimation in Natural Environments." In: *Computer Vision – ECCV 2018*. Ed. by Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss. Vol. 11214, pp. 339–357. DOI: 10.1007/978-3-030-01249-6_21. URL: http://link.springer.com/10.1007/978-3-030-01249-6_21 (visited on 06/30/2021) (cit. on p. 46).

- Fischer, Tobias and Yiannis Demiris (May 2016). "Markerless Perspective Taking for Humanoid Robots in Unconstrained Environments." In: *2016 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 3309–3316. DOI: [10.1109/ICRA.2016.7487504](https://doi.org/10.1109/ICRA.2016.7487504) (cit. on p. 73).
- Fischler, Martin A. and Robert C. Bolles (June 1981). "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography." In: *Communications of the ACM* 24.6, pp. 381–395. ISSN: 0001-0782. DOI: [10.1145/358669.358692](https://doi.org/10.1145/358669.358692). URL: <https://doi.org/10.1145/358669.358692> (visited on 07/01/2020) (cit. on p. 45).
- Fong, Tamara G., Sidney T. Jr. Bogardus, Aditya Daftary, Eliza Auerbach, Hal Blumenfeld, Sharada Modur, Linda Leo-Summers, John Seibyl, and Sharon K. Inouye (Dec. 2006). "Cerebral Perfusion Changes in Older Delirious Patients Using 99mTc HMPAO SPECT." In: *The Journals of Gerontology: Series A* 61.12, pp. 1294–1299. ISSN: 1079-5006. DOI: [10.1093/gerona/61.12.1294](https://doi.org/10.1093/gerona/61.12.1294). URL: <https://doi.org/10.1093/gerona/61.12.1294> (visited on 03/16/2022) (cit. on p. 123).
- Gal, Yarín and Zoubin Ghahramani (2016). "A Theoretically Grounded Application of Dropout in Recurrent Neural Networks." In: *Advances in Neural Information Processing Systems* 29, pp. 1019–1027. URL: <https://papers.nips.cc/paper/2016/hash/076a0c97d09cf1a0ec3e19c7f2529f2b-Abstract.html> (visited on 12/22/2020) (cit. on pp. 33, 100).
- Gallagher, Helen L. and Christopher D. Frith (Feb. 2003). "Functional Imaging of 'Theory of Mind'." In: *Trends in Cognitive Sciences* 7.2, pp. 77–83. ISSN: 1364-6613. DOI: [10.1016/S1364-6613\(02\)00025-6](https://www.sciencedirect.com/science/article/pii/S1364661302000256). URL: <https://www.sciencedirect.com/science/article/pii/S1364661302000256> (visited on 06/01/2022) (cit. on p. 109).
- Gaudreau, Jean-David, Pierre Gagnon, François Harel, Annie Tremblay, and Marc-André Roy (Apr. 2005). "Fast, Systematic, and Continuous Delirium Assessment in Hospitalized Patients: The Nursing Delirium Screening

- Scale." In: *Journal of Pain and Symptom Management* 29.4, pp. 368–375. ISSN: 0885-3924, 1873-6513. DOI: [10.1016/j.jpainsymman.2004.07.009](https://doi.org/10.1016/j.jpainsymman.2004.07.009). URL: [https://www.jpasmjournal.com/article/S0885-3924\(05\)00053-9/fulltext](https://www.jpasmjournal.com/article/S0885-3924(05)00053-9/fulltext) (visited on 03/07/2022) (cit. on p. 2).
- Gers, Felix A., Jürgen Schmidhuber, and Fred Cummins (Oct. 2000). "Learning to Forget: Continual Prediction with LSTM." In: *Neural Computation* 12.10, pp. 2451–2471. ISSN: 0899-7667. DOI: [10.1162/089976600300015015](https://doi.org/10.1162/089976600300015015). URL: <https://doi.org/10.1162/089976600300015015> (visited on 04/11/2022) (cit. on p. 22).
- Girshick, Ross (2015). "Fast R-CNN." In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 1440–1448. URL: https://openaccess.thecvf.com/content_iccv_2015/html/Girshick_Fast_R-CNN_ICCV_2015_paper.html (visited on 05/12/2022) (cit. on p. 46).
- Glassner, Andrew S. (Oct. 1984). "Space Subdivision for Fast Ray Tracing." In: *IEEE Computer Graphics and Applications* 4.10, pp. 15–24. ISSN: 1558-1756. DOI: [10.1109/MCG.1984.6429331](https://doi.org/10.1109/MCG.1984.6429331) (cit. on p. 44).
- Greene, Michelle R., Tommy Liu, and Jeremy M. Wolfe (2012). "Reconsidering Yarbus: A Failure to Predict Observers' Task from Eye Movement Patterns." In: *Vision Research* 62, pp. 1–8. ISSN: 00426989. DOI: [10.1016/j.visres.2012.03.019](https://doi.org/10.1016/j.visres.2012.03.019). URL: <http://dx.doi.org/10.1016/j.visres.2012.03.019> (cit. on pp. 107, 108).
- Hacohen, Guy and Daphna Weinshall (May 2019). "On The Power of Curriculum Learning in Training Deep Networks." In: *International Conference on Machine Learning*. PMLR, pp. 2535–2544. URL: <https://proceedings.mlr.press/v97/hacohen19a.html> (visited on 08/26/2021) (cit. on pp. 56, 57).
- Haji-Abolhassani, Amin and James J. Clark (Oct. 2014). "An Inverse Yarbus Process: Predicting Observers' Task from Eye Movement Patterns." In: *Vision Research* 103, pp. 127–142. ISSN: 00426989. DOI: [10.1016/j.visres.2014.10.009](https://doi.org/10.1016/j.visres.2014.10.009).

- 2014.08.014. URL: <http://linkinghub.elsevier.com/retrieve/pii/S0042698914002004> (cit. on p. 108).
- Han, Jin H., Svetlana Eden, Ayumi Shintani, Alessandro Morandi, John Schnelle, Robert S. Dittus, Alan B. Storrow, and E. Wesley Ely (2011). "Delirium in Older Emergency Department Patients Is an Independent Predictor of Hospital Length of Stay." In: *Academic Emergency Medicine* 18.5, pp. 451–457. ISSN: 1553-2712. DOI: [10.1111/j.1553-2712.2011.01065.x](https://doi.org/10.1111/j.1553-2712.2011.01065.x). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1553-2712.2011.01065.x> (visited on 06/16/2022) (cit. on p. 14).
- Haruno, Masahiko, Daniel M. Wolpert, and Mitsuo Kawato (Oct. 2001). "MO-SAIC Model for Sensorimotor Learning and Control." In: *Neural Computation* 13.10, pp. 2201–2220. ISSN: 0899-7667, 1530-888X. DOI: [10.1162/089976601750541778](https://doi.org/10.1162/089976601750541778). URL: <https://direct.mit.edu/neco/article/13/10/2201-2220/6463> (visited on 05/15/2022) (cit. on p. 109).
- (Oct. 2003). "Hierarchical MOSAIC for Movement Generation." In: *International Congress Series. Cognition and Emotion in the Brain. Selected Topics of the International Symposium on Limbic and Association Cortical Systems* 1250, pp. 575–590. ISSN: 0531-5131. DOI: [10.1016/S0531-5131\(03\)00190-0](https://doi.org/10.1016/S0531-5131(03)00190-0). URL: <https://www.sciencedirect.com/science/article/pii/S0531513103001900> (visited on 05/15/2022) (cit. on p. 109).
- He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun (2015). "Deep Residual Learning for Image Recognition." In: *Arxiv.Org* 7.3, pp. 171–180. ISSN: 1664-1078. DOI: [10.3389/fpsyg.2013.00124](https://doi.org/10.3389/fpsyg.2013.00124). arXiv: [1512.03385](https://arxiv.org/abs/1512.03385) (cit. on pp. 23, 99).
- Hesch, Joel A. and Stergios I. Roumeliotis (Nov. 2011). "A Direct Least-Squares (DLS) Method for PnP." In: *2011 International Conference on Computer Vision*, pp. 383–390. DOI: [10.1109/ICCV.2011.6126266](https://doi.org/10.1109/ICCV.2011.6126266) (cit. on p. 45).
- Hirsch, J., G. DePalma, T. T. Tsai, L. P. Sands, and J. M. Leung (Sept. 2015). "Impact of Intraoperative Hypotension and Blood Pressure Fluctuations on

- Early Postoperative Delirium after Non-Cardiac Surgery." In: *British Journal of Anaesthesia* 115.3, pp. 418–426. ISSN: 14716771. DOI: [10.1093/bja/aeu458](https://doi.org/10.1093/bja/aeu458). URL: <http://bja.oxfordjournals.org/lookup/doi/10.1093/bja/aeu458> (cit. on p. 28).
- Hochreiter, Sepp, S Hochreiter, Jürgen Schmidhuber, and J Schmidhuber (1997). "Long Short-Term Memory." In: *Neural computation* 9.8, pp. 1735–80. ISSN: 0899-7667. DOI: [10.1162/neco.1997.9.8.1735](https://doi.org/10.1162/neco.1997.9.8.1735). arXiv: [1206.2944](https://arxiv.org/abs/1206.2944). URL: <http://www.ncbi.nlm.nih.gov/pubmed/9377276> (cit. on p. 22).
- Holmes, Emily A. et al. (June 2020). "Multidisciplinary Research Priorities for the COVID-19 Pandemic: A Call for Action for Mental Health Science." In: *The Lancet Psychiatry* 7.6, pp. 547–560. ISSN: 2215-0366, 2215-0374. DOI: [10.1016/S2215-0366\(20\)30168-1](https://doi.org/10.1016/S2215-0366(20)30168-1). URL: [https://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366\(20\)30168-1/abstract](https://www.thelancet.com/journals/lanpsy/article/PIIS2215-0366(20)30168-1/abstract) (visited on 07/13/2020) (cit. on p. 129).
- Holt-Lunstad, Julianne, Timothy B Smith, and J Bradley Layton (Nov. 2010). "Social Relationships and Mortality Risk: A Meta-analytic Review." In: *PLoS Med* 7.7, e1000316. ISSN: 1549-1676. DOI: [10.1371/journal.pmed.1000316](https://doi.org/10.1371/journal.pmed.1000316). URL: <http://dx.doi.org/10.1371/journal.pmed.1000316> (cit. on p. 1).
- Hou, Xiaodi and Liqing Zhang (June 2007). "Saliency Detection: A Spectral Residual Approach." In: *2007 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–8. DOI: [10.1109/CVPR.2007.383267](https://doi.org/10.1109/CVPR.2007.383267) (cit. on pp. 117, 118, 121).
- Irwin, David E (1992). "Memory for Position and Identity across Eye Movements." In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 18.2, p. 307 (cit. on p. 95).
- Itti, Laurent and Christof Koch (Mar. 2001). "Computational Modelling of Visual Attention." In: *Nature Reviews Neuroscience* 2.3, pp. 194–203. ISSN: 1471-0048. DOI: [10.1038/35058500](https://doi.org/10.1038/35058500). URL: <https://www.nature.com/articles/35058500> (visited on 06/09/2020) (cit. on p. 108).

- Itti, Laurent, Christof Koch, and Ernst Niebur (1998). "A Model of Saliency-Based Visual Attention for Rapid Scene Analysis." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20.11, pp. 1254–1259. ISSN: 01628828. DOI: [10.1109/34.730558](https://doi.org/10.1109/34.730558). arXiv: [1011.1669v3](https://arxiv.org/abs/1011.1669v3). URL: <http://ieeexplore.ieee.org/document/730558/> (visited on 02/11/2017) (cit. on pp. [108](#), [117](#), [118](#), [121](#)).
- Jackson, James C., Sharon M. Gordon, Robert P. Hart, Ramona O. Hopkins, and E. Wesley Ely (June 2004). "The Association Between Delirium and Cognitive Decline: A Review of the Empirical Literature." In: *Neuropsychology Review* 14.2, pp. 87–98. ISSN: 1573-6660. DOI: [10.1023/B:NERV.0000028080.39602.17](https://doi.org/10.1023/B:NERV.0000028080.39602.17). URL: <https://doi.org/10.1023/B:NERV.0000028080.39602.17> (visited on 06/16/2022) (cit. on pp. [1](#), [15](#)).
- Jarodzka, Halszka, Kenneth Holmqvist, and Marcus Nyström (2010). "A Vector-Based, Multidimensional Scanpath Similarity Measure." In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications - ETRA '10* 1.212, p. 211. DOI: [10.1145/1743666.1743718](https://doi.org/10.1145/1743666.1743718). URL: <http://portal.acm.org/citation.cfm?doid=1743666.1743718> (cit. on pp. [xxvi](#), [141](#)).
- Johansson, Roger, Jana Holsanova, and Kenneth Holmqvist (Nov. 2006). "Pictures and Spoken Descriptions Elicit Similar Eye Movements During Mental Imagery, Both in Light and in Complete Darkness." In: *Cognitive Science* 30.6, pp. 1053–1079. ISSN: 03640213. URL: http://doi.wiley.com/10.1207/s15516709cog0000_86 (visited on 02/15/2017) (cit. on p. [138](#)).
- Kanan, Christopher, Nicholas a. Ray, Dina N. F. Bseiso, Janet H. Hsiao, and Garrison W. Cottrell (2014). "Predicting an Observer's Task Using Multi-Fixation Pattern Analysis." In: *Proceedings of the Symposium on Eye Tracking Research and Applications - ETRA '14*, pp. 287–290. DOI: [10.1145/2578153.2578208](https://doi.org/10.1145/2578153.2578208). URL: <http://dl.acm.org/citation.cfm?doid=2578153.2578208> (cit. on p. [108](#)).

- Karras, Tero, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila (2020). "Training Generative Adversarial Networks with Limited Data." In: *Thirty-Fourth Conference on Neural Information Processing Systems*. URL: <https://research.aalto.fi/en/publications/training-generative-adversarial-networks-with-limited-data> (visited on 09/28/2021) (cit. on p. 59).
- Kingma, Diederik P. and Jimmy Ba (Jan. 2017). "Adam: A Method for Stochastic Optimization." In: *arXiv:1412.6980 [cs]*. arXiv: 1412.6980 [cs]. URL: <http://arxiv.org/abs/1412.6980> (visited on 12/22/2020) (cit. on p. 63).
- Klok, F. A. et al. (Apr. 2020). "Confirmation of the High Cumulative Incidence of Thrombotic Complications in Critically Ill ICU Patients with COVID-19: An Updated Analysis." In: *Thrombosis Research*. ISSN: 0049-3848. DOI: 10.1016/j.thromres.2020.04.041. URL: <http://www.sciencedirect.com/science/article/pii/S0049384820301572> (visited on 05/25/2020) (cit. on p. 29).
- Koch, Christof and Shimon Ullman (1987). "Shifts in Selective Visual Attention: Towards the Underlying Neural Circuitry." In: *Matters of Intelligence: Conceptual Structures in Cognitive Neuroscience*. Ed. by Lucia M. Vaina. Synthese Library. Dordrecht: Springer Netherlands, pp. 115–141. ISBN: 978-94-009-3833-5. DOI: 10.1007/978-94-009-3833-5_5. URL: https://doi.org/10.1007/978-94-009-3833-5_5 (visited on 10/31/2020) (cit. on p. 108).
- Koponen, H, J Partanen, A Pääkkönen, E Mattila, and P J Riekkinen (Aug. 1989a). "EEG Spectral Analysis in Delirium." In: *Journal of neurology, neurosurgery, and psychiatry* 52.8, pp. 980–5. ISSN: 0022-3050. DOI: 10.1136/jnnp.52.8.980. URL: <http://www.ncbi.nlm.nih.gov/pubmed/2795067> (visited on 06/07/2018) (cit. on p. 123).
- Koponen, Hannu, Leevi Hurri, Ulla Stenbäck, Esa Mattila, Hilka Soininen, and Paavo J. Riekkinen (1989b). "Computed Tomography Findings in

- Delirium." In: *Journal of Nervous and Mental Disease* 177.4, pp. 226–231. ISSN: 1539-736X. DOI: [10.1097/00005053-198904000-00006](https://doi.org/10.1097/00005053-198904000-00006) (cit. on p. 123).
- Kotval, Xerxes P. and Joseph H. Goldberg (Oct. 1998). "Eye Movements and Interface Component Grouping: An Evaluation Method." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 42.5, pp. 486–490. ISSN: 2169-5067. DOI: [10.1177/154193129804200509](https://doi.org/10.1177/154193129804200509). URL: <https://doi.org/10.1177/154193129804200509> (visited on 04/02/2022) (cit. on p. 95).
- Królak, Aleksandra and Paweł Strumiłło (Nov. 2012). "Eye-Blink Detection System for Human–Computer Interaction." In: *Universal Access in the Information Society* 11.4, pp. 409–419. ISSN: 1615-5297. DOI: [10.1007/s10209-011-0256-6](https://doi.org/10.1007/s10209-011-0256-6). URL: <https://doi.org/10.1007/s10209-011-0256-6> (visited on 03/26/2021) (cit. on p. 55).
- Kumar, Ayush, Anjul Tyagi, Michael Burch, Daniel Weiskopf, and Klaus Mueller (June 2019). "Task Classification Model for Visual Fixation, Exploration, and Search." In: *Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications*, pp. 1–4. DOI: [10.1145/3314111.3323073](https://arxiv.org/abs/1907.12635). arXiv: [1907.12635](https://arxiv.org/abs/1907.12635). URL: <http://arxiv.org/abs/1907.12635> (visited on 09/14/2020) (cit. on p. 108).
- Lancaster, Gillian A., Susanna Dodd, and Paula R. Williamson (May 2004). "Design and Analysis of Pilot Studies: Recommendations for Good Practice." In: *Journal of Evaluation in Clinical Practice* 10.2, pp. 307–312. ISSN: 13561294. DOI: [10.1111/j..2002.384.doc.x](https://doi.org/10.1111/j..2002.384.doc.x). URL: <http://doi.wiley.com/10.1111/j..2002.384.doc.x> (visited on 06/07/2018) (cit. on p. 90).
- Le Meur, Olivier and Antoine Coutrot (2016). "Introducing Context-Dependent and Spatially-Variant Viewing Biases in Saccadic Models." In: *Vision Research* 121, pp. 72–84. ISSN: 18785646. DOI: [10.1016/j.visres.2016.01.005](https://doi.org/10.1016/j.visres.2016.01.005). URL: <http://dx.doi.org/10.1016/j.visres.2016.01.005> (cit. on pp. 108, 109).

- Le Meur, Olivier and Zhi Liu (2015). "Saccadic Model of Eye Movements for Free-Viewing Condition." In: *Vision Research* 116, pp. 152–164. ISSN: 18785646. DOI: [10.1016/j.visres.2014.12.026](https://doi.org/10.1016/j.visres.2014.12.026). URL: <http://dx.doi.org/10.1016/j.visres.2014.12.026> (cit. on pp. 109, 112).
- Lehman, Anthony F, Jeffrey A. Lieberman, Lisa B Dixon, and A. Miller (2004). "American Psychiatric Association Practice Guidelines for the Treatment of Psychiatric Disorders: Compendium 2006." In: p. 1600. ISSN: 0002-953X. DOI: [10.1176/appi.ajp.160.8.1531](https://books.google.com/books/about/American_Psychiatric_Association_Practic.html?id=zql0AqtRSrYC). URL: https://books.google.com/books/about/American_Psychiatric_Association_Practic.html?id=zql0AqtRSrYC (cit. on p. 18).
- Leslie, Douglas L. and Sharon K. Inouye (2011). "The Importance of Delirium: Economic and Societal Costs." In: *Journal of the American Geriatrics Society* 59.S2, S241–S243. ISSN: 1532-5415. DOI: [10.1111/j.1532-5415.2011.03671.x](https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1532-5415.2011.03671.x). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1532-5415.2011.03671.x> (visited on 06/16/2022) (cit. on p. 14).
- Lin, Shu-Min, Chien-ying Liu, Chun-hua Wang, Horng-chyuan Lin, Chien-Da Huang, Pei-yao Huang, Yueh-fu Fang, Meng-heng Shieh, and Han-pin Kuo (2004). "The Impact of Delirium on the Survival of Mechanically Ventilated Patients." In: *Critical care medicine* 32.11, pp. 2254–9. ISSN: 0090-3493. DOI: [10.1097/01.CCM.0000145587.16421.BB](https://doi.org/10.1097/01.CCM.0000145587.16421.BB) (cit. on pp. 1, 2).
- Liu, Huiying, Dong Xu, Qingming Huang, Wen Li, Min Xu, and Stephen Lin (2013). "Semantically-Based Human Scanpath Estimation with HMMs." In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 3232–3239. ISSN: 1550-5499. DOI: [10.1109/ICCV.2013.401](https://doi.org/10.1109/ICCV.2013.401) (cit. on p. 109).
- Loshchilov, Ilya and Frank Hutter (Sept. 2018). "Decoupled Weight Decay Regularization." In: *International Conference on Learning Representations*. URL: <https://openreview.net/forum?id=Bkg6RiCqY7> (visited on 01/31/2022) (cit. on pp. 32, 99).

- Loughland, Carmel M., Leanne M. Williams, and Evian Gordon (Aug. 2002). "Schizophrenia and Affective Disorder Show Different Visual Scanning Behavior for Faces: A Trait versus State-Based Distinction?" In: *Biological Psychiatry* 52.4, pp. 338–348. ISSN: 00063223. DOI: [10.1016/S0006-3223\(02\)01356-2](https://doi.org/10.1016/S0006-3223(02)01356-2). URL: <http://linkinghub.elsevier.com/retrieve/pii/S0006322302013562> (cit. on p. 20).
- Lundström, Maria, Birgitta Olofsson, Michael Stenvall, Stig Karlsson, Lars Nyberg, Undis Englund, Bengt Borssén, Olle Svensson, and Yngve Gustafson (July 2007). "Postoperative Delirium in Old Patients with Femoral Neck Fracture: A Randomized Intervention Study." In: *Aging clinical and experimental research* 19.3, pp. 178–186. ISSN: 1594-0667. DOI: [10.1007/BF03324687](https://doi.org/10.1007/BF03324687). URL: <http://link.springer.com/10.1007/BF03324687> (cit. on p. 1).
- MacMurchy, Matt, Shannon Stemler, Mimi Zander, and Christopher P. Bonafide (Jan. 2017). "Acceptability, Feasibility, and Cost of Using Video to Evaluate Alarm Fatigue." In: *Biomedical Instrumentation & Technology* 51.1, pp. 25–33. ISSN: 0899-8205, 1943-5967. DOI: [10.2345/0899-8205-51.1.25](https://doi.org/10.2345/0899-8205-51.1.25). URL: <http://www.aami-bit.org/doi/10.2345/0899-8205-51.1.25> (visited on 04/17/2019) (cit. on p. 49).
- Maldonado, José R. (2013). "Neuropathogenesis of Delirium: Review of Current Etiologic Theories and Common Pathways." In: *American Journal of Geriatric Psychiatry* 21.12, pp. 1190–1222. ISSN: 10647481. DOI: [10.1016/j.jagp.2013.09.005](https://doi.org/10.1016/j.jagp.2013.09.005) (cit. on p. 123).
- Mao, Ling et al. (June 2020). "Neurologic Manifestations of Hospitalized Patients With Coronavirus Disease 2019 in Wuhan, China." In: *JAMA Neurology* 77.6, pp. 683–690. ISSN: 2168-6149. DOI: [10.1001/jamaneurol.2020.1127](https://doi.org/10.1001/jamaneurol.2020.1127). URL: <https://doi.org/10.1001/jamaneurol.2020.1127> (visited on 01/25/2022) (cit. on p. 28).
- McCusker, Jane, Martin Cole, Michal Abrahamowicz, Francois Primeau, and Eric Belzile (Feb. 2002). "Delirium Predicts 12-Month Mortality." In: *Archives*

- of Internal Medicine* 162.4, p. 457. ISSN: 0003-9926. DOI: [10.1001/archinte.162.4.457](https://doi.org/10.1001/archinte.162.4.457). URL: <http://archinte.jamanetwork.com/article.aspx?doi=10.1001/archinte.162.4.457> (visited on 04/04/2017) (cit. on pp. 2, 14).
- Meagher, D J (Jan. 2001). "Delirium: Optimising Management." In: *BMJ (Clinical research ed.)* 322.7279, pp. 144–9. ISSN: 0959-8138. DOI: [10.1136/bmj.322.7302.1602](https://doi.org/10.1136/bmj.322.7302.1602) (cit. on p. 2).
- Mehta, Biren and Stefan Schaal (Aug. 2002). "Forward Models in Visuomotor Control." In: *Journal of Neurophysiology* 88.2, pp. 942–953. ISSN: 0022-3077. DOI: [10.1152/jn.2002.88.2.942](https://doi.org/10.1152/jn.2002.88.2.942). URL: <https://journals.physiology.org/doi/full/10.1152/jn.2002.88.2.942> (visited on 05/15/2022) (cit. on p. 109).
- Meier-Ruge, W., O. Hunziker, and P. Iwangoff (July 1991). "Senile Dementia: A Threshold Phenomenon of Normal Aging?." In: *Annals of the New York Academy of Sciences* 621.1 Physiological, pp. 104–118. ISSN: 0077-8923. DOI: [10.1111/j.1749-6632.1991.tb16973.x](https://doi.org/10.1111/j.1749-6632.1991.tb16973.x). URL: <http://doi.wiley.com/10.1111/j.1749-6632.1991.tb16973.x> (cit. on p. 2).
- Meur, Olivier Le and Antoine Coutrot (Feb. 2016). "How Saccadic Models Help Predict Where We Look during a Visual Task? Application to Visual Quality Assessment." In: *Electronic Imaging* 2016.13, pp. 1–7. DOI: [10.2352/ISSN.2470-1173.2016.13.IQSP-216](https://doi.org/10.2352/ISSN.2470-1173.2016.13.IQSP-216) (cit. on pp. 109, 121).
- Miotto, Riccardo, Li Li, Brian A. Kidd, Joel T. Dudley, and P. Agarwal (Sept. 2016). "Deep Patient: An Unsupervised Representation to Predict the Future of Patients from the Electronic Health Records." In: *Scientific Reports* 6.1, p. 26094. ISSN: 2045-2322. DOI: [10.1038/srep26094](https://doi.org/10.1038/srep26094). URL: <http://www.nature.com/articles/srep26094> (visited on 04/14/2017) (cit. on p. 1).
- Moreno-Barea, Francisco J., José M. Jerez, and Leonardo Franco (Dec. 2020). "Improving Classification Accuracy Using Data Augmentation on Small Data Sets." In: *Expert Systems with Applications* 161, p. 113696. ISSN: 0957-4174. DOI: [10.1016/j.eswa.2020.113696](https://doi.org/10.1016/j.eswa.2020.113696). URL: <https://www.sciencedirect.com>

- [com/science/article/pii/S0957417420305200](https://doi.org/10.1006/j.clinph.2006.09.009) (visited on 06/16/2021) (cit. on p. 55).
- Morimoto, C.H., A. Amir, and M. Flickner (Aug. 2002). "Detecting Eye Position and Gaze from a Single Camera and 2 Light Sources." In: *2002 International Conference on Pattern Recognition*. Vol. 4, 314–317 vol.4. DOI: [10.1109/ICPR.2002.1047459](https://doi.org/10.1109/ICPR.2002.1047459) (cit. on p. 74).
- Munoz, Douglas P. and Stefan Everling (Mar. 2004). "Look Away: The Anti-Saccade Task and the Voluntary Control of Eye Movement." In: *Nature Reviews Neuroscience* 5.3, pp. 218–228. ISSN: 1471-0048. DOI: [10.1038/nrn1345](https://doi.org/10.1038/nrn1345). URL: <https://www.nature.com/articles/nrn1345> (visited on 04/20/2022) (cit. on p. 20).
- Numan, Tianne, Arjen J. C. Slooter, Arendina W. van der Kooi, Annemieke M. L. Hoekman, Willem J. L. Suyker, Cornelis J. Stam, and Edwin van Dellen (June 2017). "Functional Connectivity and Network Analysis during Hypoactive Delirium and Recovery from Anesthesia." In: *Clinical Neurophysiology* 128.6, pp. 914–924. ISSN: 1388-2457. DOI: [10.1016/j.clinph.2017.02.022](https://doi.org/10.1016/j.clinph.2017.02.022). URL: <https://www.sciencedirect.com/science/article/pii/S1388245717300846> (visited on 03/09/2022) (cit. on p. 123).
- Nyström, Marcus, Richard Andersson, Kenneth Holmqvist, and Joost van de Weijer (Mar. 2013). "The Influence of Calibration Method and Eye Physiology on Eyetracking Data Quality." In: *Behavior Research Methods* 45.1, pp. 272–288. ISSN: 1554-3528. DOI: [10.3758/s13428-012-0247-4](https://doi.org/10.3758/s13428-012-0247-4). URL: <https://doi.org/10.3758/s13428-012-0247-4> (visited on 07/28/2021) (cit. on p. 74).
- Ohayon, Shay and E. Rivlin (Aug. 2006). "Robust 3D Head Tracking Using Camera Pose Estimation." In: *18th International Conference on Pattern Recognition (ICPR'06)*. Vol. 1, pp. 1063–1066. DOI: [10.1109/ICPR.2006.999](https://doi.org/10.1109/ICPR.2006.999) (cit. on p. 45).

- Ohm, Christina, Manuel Müller, Bernd Ludwig, and Stefan Bienk (2014). "Where is the Landmark? Eye Tracking Studies in Large-Scale Indoor Environments." In: *2nd International Workshop on Eye Tracking for Spatial Research co-located with the 8th International Conference on Geographic Information Science (GIScience 2014)*. Vol. Vol-1241. Vienna, Austria, pp. 47–51. URL: <http://ceur-ws.org/Vol-1241/> (visited on 09/28/2021) (cit. on p. 73).
- Ouimet, Sébastien, Brian P. Kavanagh, Stewart B. Gottfried, and Yoanna Skrobik (Nov. 2007). "Incidence, Risk Factors and Consequences of ICU Delirium." In: *Intensive Care Medicine* 33.1, pp. 66–73. ISSN: 03424642. DOI: [10.1007/s00134-006-0399-8](https://doi.org/10.1007/s00134-006-0399-8). URL: <http://link.springer.com/article/10.1007/s00134-006-0399-8/fulltext.html> (cit. on p. 2).
- Palmbergen, Wijnand A C, Agnes Van Sonderen, Ali M. Keyhan-Falsafi, Ruud W M Keunen, and Ron Wolterbeek (Sept. 2012). "Improved Perioperative Neurological Monitoring of Coronary Artery Bypass Graft Patients Reduces the Incidence of Postoperative Delirium: The Haga Brain Care Strategy." In: *Interactive Cardiovascular and Thoracic Surgery* 15.4, pp. 671–677. ISSN: 15699293. DOI: [10.1093/icvts/ivs317](https://doi.org/10.1093/icvts/ivs317). URL: <http://icvts.oxfordjournals.org/cgi/doi/10.1093/icvts/ivs317> (cit. on p. 28).
- Palmu, Raimo (2011). "Mental Disorders among Burn Patients." In: *Burns* 36.7, pp. 1072–1079 (cit. on p. 2).
- Pandharipande, P.P. et al. (Oct. 2013). "Long-Term Cognitive Impairment after Critical Illness." In: *New England Journal of Medicine* 369.14, pp. 1306–1316. ISSN: 0028-4793, 1533-4406. DOI: [10.1056/NEJMoa1301372](https://doi.org/10.1056/NEJMoa1301372). URL: <http://www.nejm.org/doi/10.1056/NEJMoa1301372> (visited on 05/22/2020) (cit. on p. 2).
- Paszke, Adam et al. (2019). "PyTorch: An Imperative Style, High-Performance Deep Learning Library." In: *Advances in Neural Information Processing Systems* 32. Ed. by H. Wallach, H. Larochelle, A. Beygelzimer, F. d Alché-Buc, E. Fox, and R. Garnett. Curran Associates, Inc., pp. 8026–8037. URL: <http://proceedings.mlr.press/v32/paszke19.html>

- [//papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf](https://papers.nips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf) (visited on 10/09/2020) (cit. on p. 63).
- Peterson, Josh F., Brenda T. Pun, Robert S. Dittus, Jason W W Thomason, James C. Jackson, Ayumi K. Shintani, and Ewesley Ely (2006). "Delirium and Its Motoric Subtypes: A Study of 614 Critically Ill Patients." In: *Journal of the American Geriatrics Society* 54.3, pp. 479–484. ISSN: 00028614. DOI: [10.1111/j.1532-5415.2005.00621.x](https://doi.org/10.1111/j.1532-5415.2005.00621.x) (cit. on p. 12).
- Pisani, Margaret A, Katy L B Araujo, Peter H Van Ness, Ying Zhang, E W Ely, and Sharon K Inouye (Aug. 2006). "A Research Algorithm to Improve Detection of Delirium in the Intensive Care Unit." In: *Critical care (London, England)* 10.4, R121. ISSN: 1466-609X. DOI: [10.1186/cc5027](https://doi.org/10.1186/cc5027). URL: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1750978&tool=pmcentrez&rendertype=abstract> (cit. on p. 2).
- Prytherch, David R, Gary B Smith, Paul E Schmidt, and Peter I Featherstone (2010). "ViEWS—towards a National Early Warning Score for Detecting Adult Inpatient Deterioration." In: *Resuscitation* 81.8, pp. 932–937 (cit. on p. 2).
- Rahman, A., M. Sirshar, and A. Khan (Dec. 2015). "Real Time Drowsiness Detection Using Eye Blink Monitoring." In: *2015 National Software Engineering Conference (NSEC)*, pp. 1–7. DOI: [10.1109/NSEC.2015.7396336](https://doi.org/10.1109/NSEC.2015.7396336) (cit. on p. 55).
- Ramanauskas, N. (Oct. 2006). "Calibration of Video-Oculographical Eye-Tracking System." In: *Elektronika Ir Elektrotechnika* 72.8, pp. 65–68 (cit. on p. 74).
- Ratanamahatana, Ca and E Keogh (2004). "Everything You Know about Dynamic Time Warping Is Wrong." In: *Third Workshop on Mining Temporal and Sequential Data*, pp. 22–25. ISSN: 00903493. DOI: [10.1097/01.CCM.0000279204.24648.44](https://doi.org/10.1097/01.CCM.0000279204.24648.44) (cit. on p. 140).

- Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun (2015). "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks." In: *Advances in Neural Information Processing Systems*. Vol. 28. Curran Associates, Inc. URL: <https://proceedings.neurips.cc/paper/2015/hash/14bfa6bb14875e45bba028a21ed38046-Abstract.html> (visited on 03/03/2022) (cit. on pp. 119, 121).
- Rogers, Jonathan P, Edward Chesney, Dominic Oliver, Thomas A Pollak, Philip McGuire, Paolo Fusar-Poli, Michael S Zandi, Glyn Lewis, and Anthony S David (May 2020). "Psychiatric and Neuropsychiatric Presentations Associated with Severe Coronavirus Infections: A Systematic Review and Meta-Analysis with Comparison to the COVID-19 Pandemic." In: *The Lancet Psychiatry*. ISSN: 2215-0366. DOI: [10.1016/S2215-0366\(20\)30203-0](https://doi.org/10.1016/S2215-0366(20)30203-0). URL: <http://www.sciencedirect.com/science/article/pii/S2215036620302030> (visited on 06/01/2020) (cit. on p. 28).
- Romero, Nelson, Matheus Gutoski, Leandro Hattori, and Heitor Silvério Lopes (Jan. 2018). "The Effect of Data Augmentation on the Performance of Convolutional Neural Networks." In: *Proceeding XIII Brazilian Congress on Computational Intelligence*. ABRICOM, pp. 1–12. DOI: [10.21528/CBIC2017-51](https://doi.org/10.21528/CBIC2017-51). URL: http://abricom.org.br/eventos/cbic_2017/cbic-paper-51/ (visited on 09/28/2021) (cit. on p. 55).
- Ruiz, Nataniel, Eunji Chong, and James M. Rehg (2018). "Fine-Grained Head Pose Estimation Without Keypoints." In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 2074–2083. URL: https://openaccess.thecvf.com/content_cvpr_2018_workshops/w41/html/Ruiz_Fine-Grained_Head_Pose_CVPR_2018_paper.html (visited on 08/26/2020) (cit. on p. 45).
- Sarabia, Miguel and Yiannis Demiris (2013). "A Humanoid Robot Companion for Wheelchair Users." In: *Social Robotics* 8239, pp. 432–441. DOI: [10.1007/978-3-642-39812-1_34](https://doi.org/10.1007/978-3-642-39812-1_34)

978-3-319-02675-6_43. URL: http://link.springer.com/10.1007/978-3-319-02675-6_43 (visited on 04/17/2019) (cit. on p. 1).

Sarafianos, Nikolaos, Theodoros Giannakopoulos, Christophoros Nikou, and Ioannis A. Kakadiaris (Aug. 2018). "Curriculum Learning of Visual Attribute Clusters for Multi-Task Classification." In: *Pattern Recognition* 80, pp. 94–108. ISSN: 0031-3203. DOI: [10.1016/j.patcog.2018.02.028](https://doi.org/10.1016/j.patcog.2018.02.028). URL: <https://www.sciencedirect.com/science/article/pii/S0031320318300840> (visited on 09/28/2021) (cit. on p. 56).

Seiler, Annina, David Blum, Jeremy Werner Deuel, Caroline Hertler, Markus Schettler, Carl Moritz Zipser, Jutta Ernst, Maria Schubert, Roland von Känel, and Soenke Boettger (June 2021). "Delirium Is Associated with an Increased Morbidity and In-Hospital Mortality in Cancer Patients: Results from a Prospective Cohort Study." In: *Palliative & Supportive Care* 19.3, pp. 294–303. ISSN: 1478-9515, 1478-9523. DOI: [10.1017/S147895152000139X](https://doi.org/10.1017/S147895152000139X). URL: <https://www.cambridge.org/core/journals/palliative-and-supportive-care/article/delirium-is-associated-with-an-increased-morbidity-and-inhospital-mortality-in-cancer-patients-results-from-a-prospective-cohort-study/A7F1A54E872B2C10C93B7204A0DE658A> (visited on 06/16/2022) (cit. on p. 14).

Singer, Mervyn et al. (Feb. 2016). "The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3)." In: *JAMA* 315.8, pp. 801–810. ISSN: 0098-7484. DOI: [10.1001/jama.2016.0287](https://doi.org/10.1001/jama.2016.0287). URL: <https://doi.org/10.1001/jama.2016.0287> (visited on 03/07/2022) (cit. on p. 2).

Špakov, O. and D. Miniotas (Feb. 2007). "Visualization of Eye Gaze Data Using Heat Maps." In: *Elektronika ir Elektrotechnika* 74.2, pp. 55–58. ISSN: 2029-5731. URL: <https://eejournal.ktu.lt/index.php/elt/article/view/10372> (visited on 09/28/2021) (cit. on p. 73).

Tay, Matthew Zirui, Chek Meng Poh, Laurent Rénia, Paul A. MacAry, and Lisa F. P. Ng (June 2020). "The Trinity of COVID-19: Immunity, Inflammation

- and Intervention." In: *Nature Reviews Immunology* 20.6, pp. 363–374. ISSN: 1474-1741. DOI: [10.1038/s41577-020-0311-8](https://doi.org/10.1038/s41577-020-0311-8). URL: <https://www.nature.com/articles/s41577-020-0311-8> (visited on 11/15/2020) (cit. on p. 28).
- Trillenbergh, P, R Lencer, and Wolfgang Heide (Feb. 2004). "Eye Movements and Psychiatric Disease." In: *Current opinion in neurology* 17.1, pp. 43–7. ISSN: 1350-7540. DOI: [10.1097/00019052-200402000-00008](https://doi.org/10.1097/00019052-200402000-00008). URL: <http://www.ncbi.nlm.nih.gov/pubmed/15090876> (cit. on p. 41).
- UK Policy Framework for Health and Social Care Research (Oct. 2020). URL: <https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/uk-policy-framework-health-social-care-research/> (visited on 04/08/2021) (cit. on p. 29).
- Vaidyanathan, P. P. (Sept. 2001). "Generalizations of the Sampling Theorem: Seven Decades after Nyquist." In: *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications* 48.9, pp. 1094–1109. ISSN: 1558-1268. DOI: [10.1109/81.948437](https://doi.org/10.1109/81.948437) (cit. on p. 65).
- Vijjini, Anvesh, Kaveri Anuranjana, and Radhika Mamidi (Apr. 2021). "Analyzing Curriculum Learning for Sentiment Analysis along Task Difficulty, Pacing and Visualization Axes." In: *Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis*. Online: Association for Computational Linguistics, pp. 117–128. URL: <https://aclanthology.org/2021.wassa-1.13> (visited on 09/27/2021) (cit. on p. 56).
- Wang, Ruohan, Pierluigi V. Amadori, and Yiannis Demiris (Oct. 2018). "Real-Time Workload Classification during Driving Using HyperNetworks." In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 3060–3065. DOI: [10.1109/IROS.2018.8594305](https://doi.org/10.1109/IROS.2018.8594305) (cit. on p. 41).
- Wassenaar, A. et al. (June 2015). "Multinational Development and Validation of an Early Prediction Model for Delirium in ICU Patients." In: *Intensive Care Medicine* 41.6, pp. 1048–1056. ISSN: 0342-4642. DOI: [10.1007/s00134-](https://doi.org/10.1007/s00134-015-0948-8)

015-3777-2. URL: <http://link.springer.com/10.1007/s00134-015-3777-2> (visited on 10/17/2016) (cit. on p. 28).

Wassenaar, Annelies, Lisette Schoonhoven, John W Devlin, Frank MP van Haren, Arjen JC Slooter, Philippe G Jorens, Mathieu van der Jagt, Koen S Simons, Ingrid Egerod, Lisa D Burry, et al. (2018). "Delirium Prediction in the Intensive Care Unit: Comparison of Two Delirium Prediction Models." In: *Critical Care* 22.1, p. 114 (cit. on p. 13).

Wei, Ping, Yang Liu, Tianmin Shu, Nanning Zheng, and Song-Chun Zhu (June 2018). "Where and Why Are They Looking? Jointly Inferring Human Attention and Intentions in Complex Tasks." In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6801–6809. DOI: [10.1109/CVPR.2018.00711](https://doi.org/10.1109/CVPR.2018.00711) (cit. on p. 74).

White, Brian J., David J. Berg, Janis Y. Kan, Robert A. Marino, Laurent Itti, and Douglas P. Munoz (Jan. 2017). "Superior Colliculus Neurons Encode a Visual Saliency Map during Free Viewing of Natural Dynamic Video." In: *Nature Communications* 8.1, p. 14263. ISSN: 2041-1723. DOI: [10.1038/ncomms14263](https://doi.org/10.1038/ncomms14263). URL: <https://www.nature.com/articles/ncomms14263> (visited on 06/17/2021) (cit. on p. 118).

White, Ian R., Patrick Royston, and Angela M. Wood (2011). "Multiple Imputation Using Chained Equations: Issues and Guidance for Practice." In: *Statistics in Medicine* 30.4, pp. 377–399. ISSN: 1097-0258. DOI: [10.1002/sim.4067](https://doi.org/10.1002/sim.4067). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1002/sim.4067> (visited on 04/26/2022) (cit. on p. 25).

Wilson, Glenn F. (Jan. 2002). "An Analysis of Mental Workload in Pilots During Flight Using Multiple Psychophysiological Measures." In: *The International Journal of Aviation Psychology* 12.1, pp. 3–18. ISSN: 1050-8414. DOI: [10.1207/S15327108IJAP1201_2](https://doi.org/10.1207/S15327108IJAP1201_2). URL: https://doi.org/10.1207/S15327108IJAP1201_2 (visited on 03/26/2021) (cit. on p. 55).

- Yarbus, Alfred L. (1967). "Eye Movements During Perception of Complex Objects." In: *Eye Movements and Vision*. Ed. by Alfred L. Yarbus. Boston, MA: Springer US, pp. 171–211. ISBN: 978-1-4899-5379-7. DOI: [10.1007/978-1-4899-5379-7_8](https://doi.org/10.1007/978-1-4899-5379-7_8). URL: https://doi.org/10.1007/978-1-4899-5379-7_8 (visited on 03/16/2022) (cit. on p. 107).
- Yokota, Hiroyuki, Sato Ogawa, Akira Kurokawa, and Yasuhiro Yamamoto (2003). "Regional Cerebral Blood Flow in Delirium Patients." In: *Psychiatry and Clinical Neurosciences* 57.3, pp. 337–339. ISSN: 1440-1819. DOI: [10.1046/j.1440-1819.2003.01126.x](https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1440-1819.2003.01126.x). URL: <https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1440-1819.2003.01126.x> (visited on 03/16/2022) (cit. on p. 123).
- Zhang, Chiyuan, Samy Bengio, Moritz Hardt, Benjamin Recht, and Oriol Vinyals (Feb. 2021). "Understanding Deep Learning (Still) Requires Re-thinking Generalization." In: *Communications of the ACM* 64.3, pp. 107–115. ISSN: 0001-0782. DOI: [10.1145/3446776](https://doi.org/10.1145/3446776). URL: <https://doi.org/10.1145/3446776> (visited on 01/25/2022) (cit. on p. 38).
- Zhang, Kaipeng, Zhanpeng Zhang, Zhifeng Li, and Yu Qiao (Oct. 2016). "Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks." In: *IEEE Signal Processing Letters* 23.10, pp. 1499–1503. ISSN: 1558-2361. DOI: [10.1109/LSP.2016.2603342](https://doi.org/10.1109/LSP.2016.2603342) (cit. on p. 45).
- Zhang, Lingyun, Matthew H. Tong, Tim K. Marks, Honghao Shan, and Garrison W. Cottrell (May 2008). "SUN: A Bayesian Framework for Saliency Using Natural Statistics." In: *Journal of Vision* 8.7, pp. 32–32. ISSN: 1534-7362. DOI: [10.1167/8.7.32](http://jov.arvojournals.org/article.aspx?articleid=2297284). URL: <http://jov.arvojournals.org/article.aspx?articleid=2297284> (visited on 12/17/2020) (cit. on p. 108).
- Zhang, Mengmi, Jiashi Feng, Keng Teck Ma, Joo Hwee Lim, Qi Zhao, and Gabriel Kreiman (Sept. 2018). "Finding Any Waldo with Zero-Shot Invariant and Efficient Visual Search." In: *Nature Communications* 9.1, p. 3730. ISSN: 2041-1723. DOI: [10.1038/s41467-018-06217-x](https://www.10.1038/s41467-018-06217-x). URL: <https://www.10.1038/s41467-018-06217-x>.

[nature.com/articles/s41467-018-06217-x](https://www.nature.com/articles/s41467-018-06217-x) (visited on 11/06/2020) (cit. on p. 116).

Zhang, Shifeng, Xiangyu Zhu, Zhen Lei, Hailin Shi, Xiaobo Wang, and Stan Z. Li (2017). "S3FD: Single Shot Scale-Invariant Face Detector." In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 192–201. arXiv: 1708.05237. URL: https://openaccess.thecvf.com/content_iccv_2017/html/Zhang_S3FD_Single_Shot_ICCV_2017_paper.html (visited on 08/01/2021) (cit. on p. 44).

Zhu, Xiangyu, Xiaoming Liu, Zhen Lei, and Stan Z. Li (Jan. 2019). "Face Alignment in Full Pose Range: A 3D Total Solution." In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 41.1, pp. 78–92. ISSN: 1939-3539. DOI: 10.1109/TPAMI.2017.2778152 (cit. on p. 45).

van der Kooi, Arendina W., Irene J. Zaal, Francina A. Klijn, Huiberdina L. Koek, Ronald C. Meijer, Frans S. Leijten, and Arjen J. Slooter (Jan. 2015). "Delirium Detection Using EEG." In: *Chest* 147.1, pp. 94–101. ISSN: 0012-3692. DOI: 10.1378/chest.13-3050. URL: <http://www.sciencedirect.com/science/article/pii/S0012369215302385> (visited on 07/01/2020) (cit. on p. 123).

van der Wilk, Mark, Matthias Bauer, ST John, and James Hensman (2018). "Learning Invariances Using the Marginal Likelihood." In: *Advances in Neural Information Processing Systems*. Vol. 31. Curran Associates, Inc. URL: <https://papers.nips.cc/paper/2018/hash/d465f14a648b3d0a1faa6f447e526c60-Abstract.html> (visited on 09/28/2021) (cit. on p. 56).

