SAMSA: Segment Anything Model enhanced with Spectral Angles for Hyperspectral Interactive Medical Image Segmentation

Alfie Roddan[1], Tobias Czempiel[1], Chi Xu[1], Daniel S. Elson[1], and Stamatia Giannarou[1]

The Hamlyn Centre for Robotic Surgery, Department of Surgery and Cancer
Imperial College London, United Kingdom
agr21@ic.ac.uk

Abstract. Hyperspectral imaging (HSI) provides rich spectral information for medical imaging, yet encounters significant challenges due to data limitations and hardware variations. We introduce SAMSA, a novel interactive segmentation framework that combines an RGB foundation model with spectral analysis. SAMSA efficiently utilizes user clicks to guide both RGB segmentation and spectral similarity computations. The method addresses key limitations in HSI segmentation through a unique spectral feature fusion strategy that operates independently of spectral band count and resolution. Performance evaluation on publicly available datasets has shown 81.0% 1-click and 93.4% 5-click DICE on a neurosurgical and 81.1% 1-click and 89.2% 5-click DICE on an intraoperative porcine hyperspectral dataset. Experimental results demonstrate SAMSA's effectiveness in few-shot and zero-shot learning scenarios and using minimal training examples. Our approach enables seamless integration of datasets with different spectral characteristics, providing a flexible framework for hyperspectral medical image analysis.

Keywords: Hyperspectral · Interactive Segmentation · Generalization

1 Introduction

Hyperspectral imaging (HSI) offers superior intraoperative guidance through its rich spectral information, allowing improved differentiation between visually similar tissues [3, 18]. The diverse range of HSI hardware, with varying spectral ranges and resolutions, creates significant interoperability challenges that impede data standardization [1]. This technical fragmentation, coupled with HSI's limited clinical adoption, has resulted in a shortage of comprehensive datasets, presenting a substantial obstacle for machine learning applications [7]. Despite these challenges, recent advances have demonstrated HSI's potential for intraoperative segmentation in neurosurgery [18, 15, 9] and on porcine organs [17, 19]. However, developing generalized models that account for hardware variations remains unsolved. Classical Spectral Comparison Functions (SCF) such as Spectral Angle (SA) [2] and Pearson's Correlation Coefficient (PCC) [12] offer highly

adaptable approaches for comparing spectra for manual image segmentation. Their untrained nature allows them to generalize to new scenarios without requiring additional data, functioning with any spectral range or number of bands. These methods typically operate by using a reference point (user click) to compare against the rest of the image, making them inherently interactive. However, they face limitations due to the "shading problem" where semantic objects exhibit different spectral signatures, and the challenge of establishing consistent segmentation thresholds within and across images.

Interactive segmentation is particularly valuable in medical imaging, as it leverages expert input to improve performance compared to fully automated methods [23, 20, 21] and enables segmentation of previously unseen tissue classes — a vital capability during surgical procedures where unexpected pathological findings may occur. The shared interactive nature of both classical spectral methods and modern RGB interactive segmentation presents a natural opportunity to combine these approaches, allowing a single user click to serve dual purposes namely, guiding the RGB-based model while simultaneously providing a reference point for spectral comparison. While powerful interactive models like Segment Anything and its successor SAM2 [8, 14] have revolutionized RGB segmentation, these advances cannot be directly applied to HSI due to fundamental differences in data characteristics.

The predominant strategy in existing research has been to transform modalities such as CT or MRI into grayscale representations before feeding them into foundation models like SAM [10, 22]. This approach, however, frequently fails to capitalize on the distinctive advantages that each imaging modality offers. Moving beyond conventional RGB-based representations requires novel integration methodologies, particularly since current multimodal architectures have yet to incorporate spectral fusion techniques. Consequently, established approaches cannot be readily adapted to address these hyperspectral challenges.

In this work, we propose an interactive image segmentation approach by combining SAM2 with spectral analysis techniques to overcome HSI's data limitations. Our approach leverages the advantages of large-scale RGB foundation models and integrating HSI's rich spectral information. Specifically, we contribute: (1) An interactive segmentation framework for HSI, utilizing a dual-input approach that efficiently leverages the same user input (clicks) in two complementary ways: to guide an RGB foundation model and to compute SCF measurements in HSI data, enhancing segmentation performance. (2) We demonstrate effectiveness in both few-shot and zero-shot learning scenarios for tumor classification, showing robust performance even with extremely limited training examples and on unseen test cases. (3) The first HSI machine learning framework that functions independently of HSI band count and wavelength variations, enabling the combination of datasets with different spectral characteristics and semantic classes into a unified training approach.

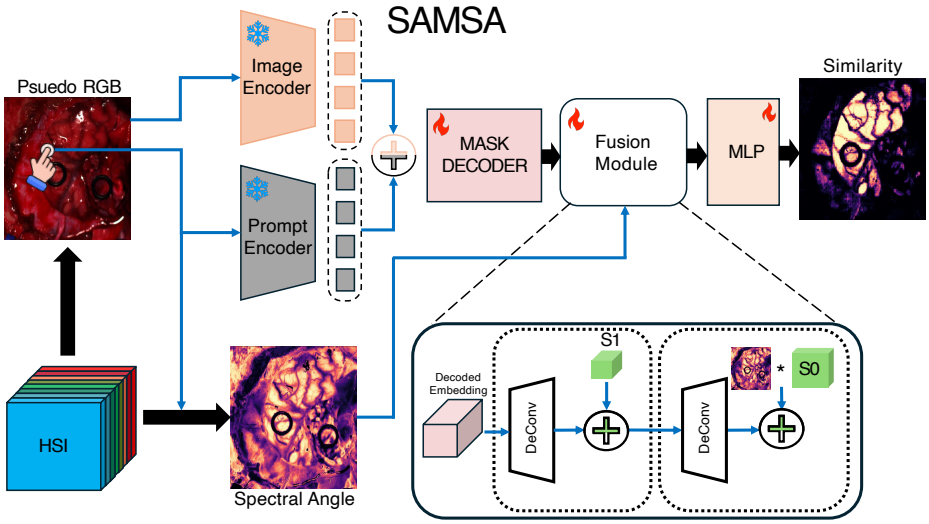


Fig. 1: SAMSA outline - a single click in the pseudo RGB is used to guide both the RGB and spectral branch.

2 Methodology

Given a hyperspectral image $X \in \mathbb{R}^{H \times W \times C}$, where H and W denote the spatial dimensions and C represents the number of spectral channels, our goal is to perform interactive foreground/background segmentation based on user-provided click positions. Additionally, we have available a corresponding pseudo RGB image $X_{rgb} \in \mathbb{R}^{H \times W \times 3}$ and a ground truth label map $Y \in [0, ..., N]^{H \times W}$ with N the number of classes. Let $\mathcal{I} = \{I_{i,j}\}$ be a set of user-provided click positions, where each $I_{i,j} = (i, j)$ corresponds to a pixel location in the image. Each model outputs a foreground/background similarity map $\hat{Y} \in [0, 1]^{H \times W}$ with 0 representing no similarity and 1 representing strong similarity to the clicked pixel(s).

Spectral Comparison Function (SCF). To model the spectral characteristics, we employ a spectral similarity approach based on click positions. Given multiple clicks $\mathcal{I}$ in the region of interest, we compute the similarity for each pixel in the image with respect to all selected spectra $S_{i,j} = X(I_{i,j})$ and assign the highest similarity score. Finally, the SCF outputs a similarity map $\hat{Y}_{SCF} = SCF(X, \mathcal{I})$. SA [2] measures the similarity between two spectra using the angle between them in the spectral space. The spectra can also be compared using PCC [12] which quantifies the linear relationship between the reference spectra and candidate spectra. PCC specifically focuses on modeling negative correlations between spectra distinguishing between positive and negative relationships. While SA and PCC effectively measure similarity between spectral samples, such as those derived from click data, they do not establish decision boundaries since they are not based on learned models. To address this, we employ histogram equalization to maximize the information content of the sim-

ilarity maps by increasing contrast and improving regional separability [6]. We denote this method as $SCF_{Equalized}$.

RGB Interactive Segmentation. To leverage powerful RGB foundational models for HSI, we first generate pseudo RGB images from the HSI data through spectral band selection and combination [4]. SAM2 [14] is utilized as our RGB segmentation backbone due to its state-of-the-art performance in interactive segmentation tasks and its robust generalization capabilities across diverse imaging domains, including medical [13]. SAM2 generates confidence maps $\hat{Y}_{SAM2} = \text{SAM2}(X_{rgb}, \mathcal{I})$ indicating the likelihood of each pixel belonging to the foreground from the pseudo RGB image. $\text{SAM2}_{Base}(\cdot)$ denotes the SAM2-Large model with frozen weights, and $\text{SAM2}_{Tuned}(\cdot)$ denotes the fine-tuned version.

RGB and Spectral Similarity. Both spectral and RGB-based models offer complementary information for image segmentation. While SAM2 processes only RGB information, the combination complements our spectral similarity-based segmentation by capturing spatial and contextual features that may not be evident in pure spectral analysis. To enhance segmentation quality, we explored two initial approaches for fusing the spectral and spatial similarity maps. The first is a simple intersection method where the similarity maps are directly multiplied: $\hat{Y}_{\text{SAM2}_{Intersec.}} = \hat{Y}_{SAM2} \cdot \hat{Y}_{SCF}$. This multiplication produces high values only in regions where both modalities agree, effectively creating a logical AND operation that requires consensus between spectral and spatial information for pixel classification. For a more sophisticated integration, we implement a UNet architecture [16] that takes the similarity maps $\hat{Y}_{SAM2}$ and $\hat{Y}_{SCF}$ as direct inputs to learn optimal fusion strategies: $\hat{Y}_{\text{SAM2}_{UNet}} = \text{SAM2}_{UNet}(\hat{Y}_{SAM2}, \hat{Y}_{SCF})$, where $\text{SAM2}_{UNet}(\cdot)$ represents the trained fusion UNet model. Unlike the deterministic intersection approach, this learnable fusion should uncover complementary spatial relationships between modalities.

SAMSA. To further improve segmentation, we introduce SAMSA, a novel model that fuses spectral similarity with high-resolution spatial features from SAM2. Unlike the aforementioned fusion approaches that combine outputs after segmentation, SAMSA integrates spectral information directly into the upscaling process of the SAM2 mask decoder. A high-level overview of this process is shown in Fig. 1. Given X_{RGB} and $\mathcal{I}$, SAMSA follows the standard SAM2 processing pipeline and additionally integrates the spectral information. The spectral similarity map $\hat{Y}_{SCF}$ is fused with the high-resolution feature maps S_0 extracted from SAM2's encoder, enhancing segmentation decisions based on spectral properties. This allows the model to leverage spectral characteristics that are not visible in pseudo RGB while maintaining SAM2's spatial precision. We freeze the prompt and image encoders from SAM2, fine-tuning only the lightweight mask decoder. This enables SAMSA to generalize to medical datasets with minimal training data while learning how to effectively combine spatial and spectral information. The additional spectral comparison and feature map multiplication introduce minimal computational overhead, with an estimated 1.5 to 2 percent increase in FLOPs compared to SAM2, allowing SAMSA to maintain real-time performance.

Table 1: Macro Results: Performance of models with varying input modalities (Mod.) and user clicks. Bold indicates peak performance per metric and dataset.

Mod.	Model	Heipor			Hib Dataset		
		1 click		5 clicks	1 click		5 clicks
		D@0.5	D@Max	D@0.5	D@0.5	D@Max	D@0.5
HSI	PCC	0.122	0.472	0.117	$0.373^{\pm 0.019}$	$0.885^{\pm 0.034}$	$0.375^{\pm 0.018}$
	SA	0.117	0.489	0.117	$0.374^{\pm 0.019}$	$0.889^{\pm 0.035}$	$0.374^{\pm 0.019}$
	$SA_{Equalized}$	0.205	0.487	0.137	$0.568^{\pm 0.038}$	$0.885^{\pm 0.034}$	$0.482^{\pm 0.033}$
RGB	$SAM2_{Base}$	0.600	0.773	0.643	$0.523^{\pm 0.056}$	$0.727^{\pm 0.043}$	$0.591^{\pm 0.069}$
	$SAM2_{Tuned}$	0.806	**0.864**	0.886	$0.771^{\pm 0.059}$	$0.905^{\pm 0.036}$	$0.912^{\pm 0.025}$
Fusion	$SAM2SA_{Intersec.}$	0.634	0.755	0.647	$0.605^{\pm 0.048}$	$0.832^{\pm 0.033}$	$0.674^{\pm 0.083}$
	$SAM2SA_{UNet}$	0.692	0.798	0.771	$0.650^{\pm 0.115}$	$0.778^{\pm 0.123}$	$0.673^{\pm 0.096}$
	SAMSA (ours)	**0.811**	0.863	**0.892**	$\mathbf{0.810}^{\pm 0.050}$	$\mathbf{0.929}^{\pm 0.028}$	$\mathbf{0.934}^{\pm 0.031}$

3 Experimental Results

For training of the fusion models we mainly follow SAM's optimization procedure [8]. All models are trained with a combined loss function using DICE and cross-entropy loss with equal weighting, excluding any unlabeled regions. Complete implementation details are provided in the accompanying source code, accessible upon acceptance of this manuscript[1].

Datasets. The **HiB dataset** includes hyperspectral and pseudo RGB images from 34 patients, with patient-wise fold splits [9]. It features four labeled classes: Background, Tumor, Healthy, and Vasculature, plus an Unlabeled category. Following preprocessing as in [11], the dataset consists of 128 spectral bands. The **HeiPorSPECTRAL (Heipor)** dataset, collected from 20 porcine subjects at Heidelberg University Hospital, provides HSI data with annotations for 20 distinct organs. Spectral information ranges from 500 nm to 1000 nm, and corresponding RGB images are derived from the HSI data [19].

Evaluation Protocol We evaluate each model on foreground/background segmentation using a single user click, following SAM2's evaluation procedure [8]. For each class, we select a click position at the center of the largest connected component in the foreground region to avoid boundary ambiguity. We report two key metrics: **D@0.5** - The DICE score [5] using the standard decision boundary of 0.5. **D@Max** - The max DICE across all thresholds, representing optimal performance without predefined decision boundaries.

We report macro-averaged (Macro) and per-class results. We also evaluate multi-click performance by placing subsequent clicks on the target foreground class. Finally, for trainable models, we conduct N-shot evaluations (1, 3, 5, 10, and 20 examples) to analyze the relationship between training data availability and segmentation quality.

In our evaluation of spectral similarity functions, SA outperformed PCC with improvements of +0.017 on Heipor and +0.004 on Hib datasets when measured

[1] `https://github.com/CVRS-Hamlyn/SAMSA`

by D@Max (table 1). We further enhanced SA with equalization ($SA_{Equalized}$), improving contrast around the 0.5 threshold to better align with RGB models, and adopted this as our spectral analysis method for subsequent experiments.

For RGB-only performance (table 1), SAM2_{Base} demonstrated reasonable generalization to medical domains, achieving 0.600 Macro D@0.5 on Heipor. However, table 2 reveals significant weaknesses on the Hib dataset's Vascular class (0.335), indicating limited generalization to domain-specific medical structures. Fine-tuning substantially improved performance, with SAM2_{Tuned} achieving 0.757 on Vascular and 0.869 on Background classes.

Our analysis of fusion strategies revealed that late fusion approaches namely, $\text{SAM2SA}_{Intersec.}$ and SAM2SA_{UNet}, underperformed compared to SAM2_{Tuned}, though they improved upon SAM2_{Base}. This suggests spectral information requires earlier integration to enhance segmentation performance, which we implemented in SAMSA.

SAMSA consistently outperformed SAM2_{Tuned} across all classes on D@0.5, with notable improvements of +0.056 for Healthy and +0.06 for Tumor classes. Macro D@0.5 scores increased by +0.039 for Hib and +0.005 for Heipor. The modest gains on Heipor can be attributed to its RGB-oriented annotations and predominance of large, centered objects (fig. 3). These characteristics are particularly favorable for RGB-only models that detect visual boundaries, as evidenced by the strong zero-shot performance of SAM2_{Base} (0.773), which trails the fine-tuned version by only −0.091 D@0.5. For this reason, we focused our per-class metric analysis on the Hib dataset, where spectral information provides more substantial benefits for segmentation.

As expected, additional clicks improved segmentation performance for all fine-tuned models. SAMSA showed significant improvements with 5-click inputs, increasing performance of D@0.5 by +0.081 on Heipor and +0.124 on Hib. In fig. 2 we demonstrate SAMSA's superiority over SAM2_{Tuned} across different click counts on Hib, achieving 0.95 Macro D@0.5 with 5 clicks. Furthermore, with only 20 training examples, SAMSA achieves 0.79 Macro D@0.5 for single-click segmentation. Leveraging foundation models, both SAMSA and SAM2 perform well in limited-data scenarios. Notably, the integration of spectral information consistently enhances the training process, with a clear performance gap between SAMSA and SAM2_{Tuned} emerging at just 5 training examples, highlighting the advantage of spectral information in low-data regimes.

Generalization Results. We conduct a leave-one-class-out experiment on both fine-tuned SAM2 and SAMSA by removing the Tumor class from training while testing across all classes on Hib, simulating real-world scenarios requiring identification of novel structures without prior supervision.

As seen in table 2, when the tumor class is excluded from training, SAM2_{Tuned} performance drops by 0.14, falling below even SAM2_{Base} performance for tumor detection. Despite this, its overall Macro performance remains significantly better (+0.185). Similarly, SAMSA experiences a performance decrease on tumor class (−0.17), but crucially maintains the highest tumor detection capability. Additionally, SAMSA achieves a higher overall Macro result (+0.052), suggest-

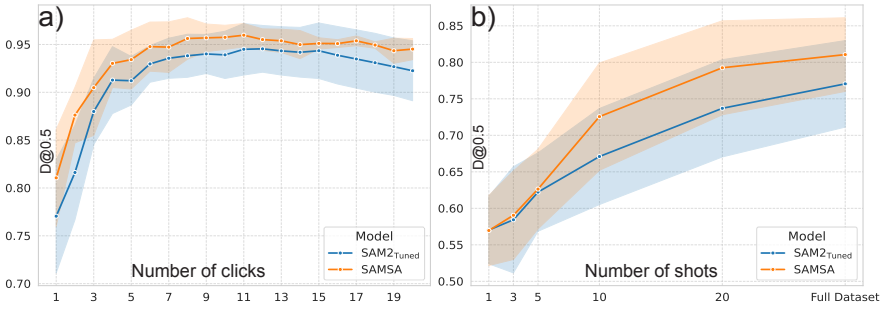


Fig. 2: Performance analysis on Hib dataset: a) Number of clicks and b) Number of shots in training and correlation to model performance.

Table 2: Class results D@0.5 for Hib dataset using 1 click.

Model	Macro	Background	Healthy	Vascular	Tumor
$SA_{Equalized}$	$0.568^{\pm 0.038}$	$0.613^{\pm 0.109}$	$0.815^{\pm 0.093}$	$0.506^{\pm 0.130}$	$0.339^{\pm 0.100}$
$SAM2_{Base}$	$0.523^{\pm 0.056}$	$0.552^{\pm 0.079}$	$0.586^{\pm 0.083}$	$0.335^{\pm 0.095}$	$0.619^{\pm 0.188}$
$SAM2_{Tuned}$	$0.771^{\pm 0.059}$	$0.869^{\pm 0.045}$	$0.778^{\pm 0.081}$	$0.757^{\pm 0.106}$	$0.678^{\pm 0.098}$
SAMSA(ours)	$\mathbf{0.810}^{\pm 0.050}$	$\mathbf{0.881}^{\pm 0.039}$	$\mathbf{0.834}^{\pm 0.066}$	$\mathbf{0.790}^{\pm 0.117}$	$\mathbf{0.738}^{\pm 0.041}$
0-shot case - excluded Tumor class from train					
$SAM2_{Tuned}$	$0.708^{\pm 0.055}$	$0.853^{\pm 0.063}$	$0.735^{\pm 0.080}$	$0.704^{\pm 0.077}$	$0.538^{\pm 0.125}$
SAMSA (ours)	$\mathbf{0.760}^{\pm 0.053}$	$\mathbf{0.881}^{\pm 0.048}$	$\mathbf{0.821}^{\pm 0.081}$	$\mathbf{0.763}^{\pm 0.095}$	$\mathbf{0.576}^{\pm 0.072}$

ing that incorporating spectral information provides meaningful advantages for generalizing to unseen classes.

Secondly, our approach uniquely enables training across datasets with different spectral properties by collapsing spectral information to a single channel regardless of band count or resolution. In table 3, cross-dataset generalization (training on one dataset, testing on another) performs poorly even below the zero-shot $SAM2_{Base}$ baseline. However, mixed training significantly improves results. While $SAM2_{Tuned}$ shows inconsistent benefits from mixed training (improved on Heipor, decreased on Hib), SAMSA maintains balanced performance,

Table 3: Model performance Macro D@0.5 using cross and mixed training

Training→	**None**	**Heipor**		**Hib**		**Mixed**	
HSI Channels	-	100		128		238	
Num Classes	-	20		4		24	
Test↓	$SAM2_{Base}$ (0-shot)	$SAM2_{Tuned}$	SAMSA	$SAM2_{Tuned}$	SAMSA	$SAM2_{Tuned}$	SAMSA
Heipor	0.600	0.806	0.811	0.445	0.433	0.807	0.810
Hib	0.523	0.454	0.497	0.771	0.810	0.695	0.765

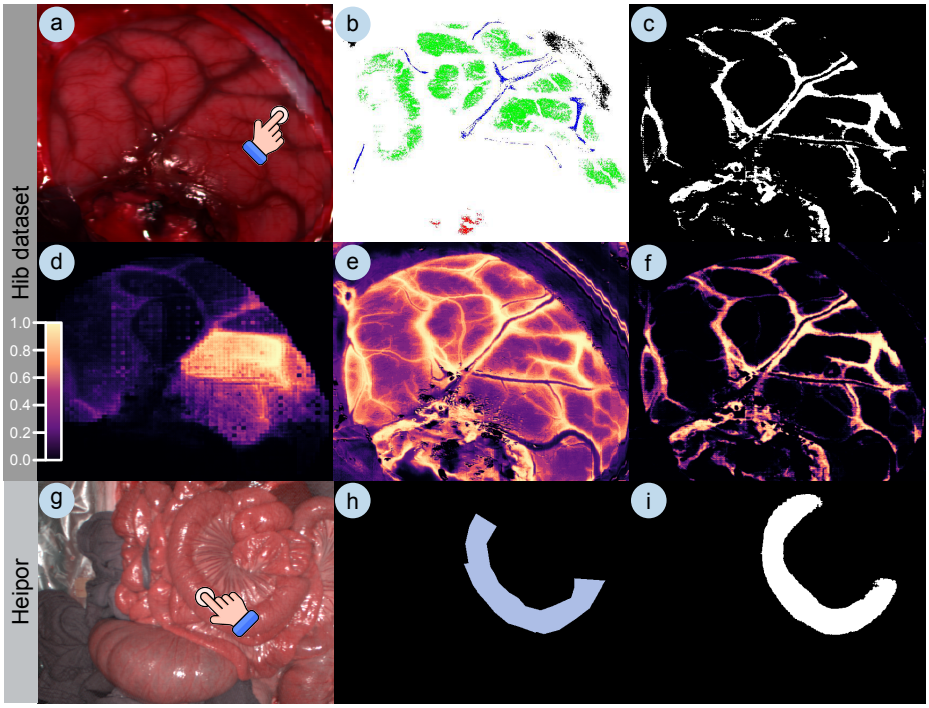


Fig. 3: Comparison of results. (a) RGB image with vascular click. (b) Corresponding label image, where Tumor is red, Vascular structures are blue, Healthy tissue is green, Background non-tissue structures are black, and unlabeled regions are white. (c) SAMSA prediction of (a). (d-f) Probability maps from SAM2, SA, and SAMSA. (g) RGB image with a small bowel click. (h) Corresponding label image, where Small Bowel is gray and Background is black. (i) SAMSA prediction of (g).

outperforming SAM2$_{Tuned}$ on both datasets (Hib +0.07, Heipor +0.003). This confirms SAMSA's ability to generalize across heterogeneous HSI datasets with varying spectral properties and clinical domains.

In fig. 3 we present qualitative results on the Hib dataset. When clicking on vascular tissue (a), SAM2$_{Tuned}$ (d) struggles to effectively segment the vascular class without spectral information. The SA map (e) clearly identifies vascular structures but introduces noise around the tumor region. In contrast, SAMSA (f) produces a well-localized probability map for vascular tissue. For the Heipor dataset, clicking on small bowel tissue (g) demonstrates SAMSA's ability to precisely delineate class boundaries compared to the ground truth (h). We believe the proposed method can be incorporated in two key ways: 1) to assist efficient annotation during dataset creation and 2) for intraoperative use as a real-time decision support tool, enabling surgeons to interactively explore and localize suspicious regions during procedures.

4 Conclusion

SAMSA is a unique method for generalizing across different HSI datasets, enabling effective segmentation in scenarios with limited training data and diverse imaging conditions. The proposed framework's ability to combine spectral and RGB information provides significant advantages, particularly in detecting challenging medical structures and maintaining performance across different datasets. Our approach shows promise in handling unseen classes and adapting to heterogeneous HSI datasets under low data regimes, opening new possibilities for flexible and robust hyperspectral interactive medical image analysis.

Acknowledgments. This work was supported by the Engineering and Physical Sciences Research Council (EP/T51780X/1) and Intel R&D UK. Dr Giannarou is supported by the Royal Society (URF\R\201014). This paper is independent research partially funded by the National Institute for Health Research (NIHR) Imperial Biomedical Research Centre (BRC), the Cancer Research UK (CRUK) Imperial Centre and the Imperial Brain Tumour Research Centre (BTR)

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Anichini, G., Leiloglou, M., Hu, Z., O'Neill, K., Daniel Elson: Hyperspectral and multispectral imaging in neurosurgery: a systematic literature review and meta-analysis. European Journal of Surgical Oncology p. 108293 (2024). https://doi.org/10.1016/j.ejso.2024.108293
2. Boardman, J.: Spectral angle mapping: a rapid measure of spectral similarity. AVIRIS. Delivered by Ingenta (1993)
3. Clancy, N.T., Jones, G., Maier-Hein, L., Elson, D.S., Stoyanov, D.: Surgical spectral imaging. Medical Image Analysis **63**, 101699 (July 2020). https://doi.org/10.1016/j.media.2020.101699, epub 2020 Apr 13
4. Czempiel, T., Roddan, A., Leiloglou, M., Hu, Z., O'Neill, K., Anichini, G., Stoyanov, D., Elson, D.: Rgb to hyperspectral: Spectral reconstruction for enhanced surgical imaging. Healthcare Technology Letters **11**(6), 307–317 (2024). https://doi.org/10.1049/htl2.12098
5. Dice, L.R.: Measures of the amount of ecologic association between species. Ecology **26**(3), 297–302 (1945). https://doi.org/10.2307/1932409
6. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Pearson, New York, NY, fourth edition, global edition edn. (2018)
7. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press (2016), `http://www.deeplearningbook.org`, book in preparation for MIT Press
8. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R.: Segment anything. arXiv:2304.02643 (2023)
9. Leon, R., Fabelo, H., Ortega, S., Cruz-Guerrero, I., Campos-Delgado, D., Szolna, A., Piñeiro, J., Espino, C., O'Shanahan, A., Hernandez, M., Carrera, D., Bisshopp, S., Sosa, C., Balea-Fernandez, F., Morera, J., Clavo, B., Callico, G.: Hyperspectral imaging benchmark based on machine learning for intraoperative brain tumour detection. NPJ Precision Oncology **7**(1), 119 (November 2023). https://doi.org/10.1038/s41698-023-00475-9

10. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature Communications **15**(1), 654 (Jan 2024). https://doi.org/10.1038/s41467-024-44824-z
11. Martinez, B., Leon, R., Fabelo, H., Ortega, S., Piñeiro, J.F., Szolna, A., Hernandez, M., Espino, C., J. O'Shanahan, A., Carrera, D., et al.: Most relevant spectral bands identification for brain cancer detection using hyperspectral imaging. Sensors **19**(24), 5481 (2019). https://doi.org/10.3390/s19245481
12. Meneses, P.R.: Spectral correlation mapper (scm): An improvement on the spectral angle mapper (sam). In: Proceedings of the AVIRIS 2000 Workshop. Jet Propulsion Laboratory, California Institute of Technology (2000), jPL Publication 00-18
13. Murali, A., Mascagni, P., Mutter, D., Padoy, N.: Cyclesam: One-shot surgical scene segmentation using cycle-consistent feature matching to prompt sam (2024), https://arxiv.org/abs/2407.06795
14. Ravi, N., Gabeur, V., Hu, Y.T., Hu, R., Ryali, C., Ma, T., Khedr, H., Rädle, R., Rolland, C., Gustafson, L., Mintun, E., Pan, J., Alwala, K.V., Carion, N., Wu, C.Y., Girshick, R., Dollár, P., Feichtenhofer, C.: Sam 2: Segment anything in images and videos. arXiv preprint arXiv:2408.00714 (2024), https://arxiv.org/abs/2408.00714
15. Roddan, A., Czempiel, T., Elson, D.S., Giannarou, S.: Calibration-jitter: Augmentation of hyperspectral data for improved surgical scene segmentation. Healthcare Technology Letters **11**(6), 345–354 (2024). https://doi.org/10.1049/htl2.12102
16. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. pp. 234–241. Springer International Publishing, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
17. Seidlitz, S., Sellner, J., Odenthal, J., Özdemir, B., Studier-Fischer, A., Knödler, S., Ayala, L., Adler, T.J., Kenngott, H.G., Tizabi, M., Wagner, M., Nickel, F., Müller-Stich, B.P., Maier-Hein, L.: Robust deep learning-based semantic organ segmentation in hyperspectral images. Medical Image Analysis **80**, 102488 (2022). https://doi.org/10.1016/j.media.2022.102488
18. Shapey, J., Xie, Y., Nabavi, E., Bradford, R., Saeed, S.R., Ourselin, S., Vercauteren, T.: Intraoperative multispectral and hyperspectral label-free imaging: A systematic review of in vivo clinical studies. Journal of Biophotonics **12**(9), e201800455 (Sep 2019). https://doi.org/10.1002/jbio.201800455, epub 2019 Apr 29
19. Studier-Fischer, A., Seidlitz, S., Sellner, J., Bressan, M., Özdemir, B., Ayala, L., Odenthal, J., Knoedler, S., Kowalewski, K.F., Haney, C.M., Salg, G., Dietrich, M., Kenngott, H., Gockel, I., Hackert, T., Müller-Stich, B.P., Maier-Hein, L., Nickel, F.: Heiporspectral - the heidelberg porcine hyperspectral imaging dataset of 20 physiological organs. Scientific Data **10**(1), 414 (June 2023). https://doi.org/10.1038/s41597-023-02315-8
20. Wang, G., Li, W., Zuluaga, M.A., et al.: Interactive medical image segmentation using deep learning with image-specific fine tuning. IEEE Transactions on Medical Imaging **37**(7), 1562–1573 (2018). https://doi.org/10.1109/TMI.2018.2791721
21. Wang, G., Zuluaga, M.A., Li, W., Pratt, R., Patel, P.A., Aertsen, M., Doel, T., David, A.L., Deprest, J., Ourselin, S., Vercauteren, T.: Deepigeos: A deep interactive geodesic framework for medical image segmentation. IEEE Transactions on Pattern Analysis and Machine Intelligence **41**(7), 1559–1572 (Jul 2019). https://doi.org/10.1109/tpami.2018.2840695

22. Wu, J., Wang, Z., Hong, M., Ji, W., Fu, H., Xu, Y., Xu, M., Jin, Y.: Medical sam adapter: Adapting segment anything model for medical image segmentation. Medical Image Analysis **102**, 103547 (2025). https://doi.org/https://doi.org/10.1016/j.media.2025.103547
23. Zhao, F., Xie, X.: An overview on interactive medical image segmentation. The Annals of the BMVA **2013**(7) (August 2013), http://www.bmva.org/annals/2013/2013-0007.pdf