# Wavelet-Guided Spectral–Spatial Learning for Medical Hyperspectral Image Segmentation

Junya Ji[1], Nick Stone[2], Yanda Meng[3], and Xujiong Ye[1(✉)]

[1] Department of Computer Science, University of Exeter, Exeter, UK
[2] Department of Physics and Astronomy, University of Exeter, Exeter, UK
[3] Biomedical Sciences Division (BioMed), King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia
X.Ye2@exeter.ac.uk

**Abstract.** Medical Hyperspectral Imaging (MHSI) provides rich pixel-level biochemical information and has demonstrated strong potential in computational pathology. However, the high spectral dimensionality and strong inter-band redundancy make effective spectral–spatial representation learning challenging. Most existing methods encode spectra as flat channel-wise features or tokenized sequences, without explicitly modeling their intrinsic multi-scale structure. To address this limitation, we propose a wavelet-guided spectral–spatial segmentation framework for MHSI. By performing one-dimensional wavelet decomposition along the spectral dimension, we explicitly separate global biochemical trends and local spectral variations. Instead of replacing the original spectrum, the low-frequency component is incorporated as a structural prior via a per-pixel adaptive scalar gate, stabilizing deep spectral learning. Furthermore, we introduce Band-Aware Aggregation and Spectral-Enhanced Skip Connections to enhance spectral–spatial interaction in the decoding process. Extensive experiments on public medical hyperspectral datasets demonstrate that our method consistently improves segmentation performance compared with state-of-the-art methods. The source code is publicly available at https://github.com/Junya99/Wavelet-Guided-MHSI.

**Keywords:** Medical Hyperspectral Image Segmentation · Wavelet-Guided Spectral–Spatial Learning · Spectral Representation Reparameterization.

## 1 Introduction

Medical Hyperspectral Imaging (MHSI) has emerged as a powerful imaging modality for computational pathology and medical diagnosis. It produces a three-dimensional (3D) spatial–spectral cube in which each pixel is associated with a continuous spectral curve, reflecting underlying molecular composition and tissue physiology [11]. As shown in Fig. 1(a), pixels sampled from cancerous and normal tissue regions exhibit distinct spectral profiles, highlighting the discriminative potential of spectral information. However, the high dimensionality, strong inter-band redundancy, and multiscale spectral characteristics of MHSI data fundamentally complicate effective representation learning.

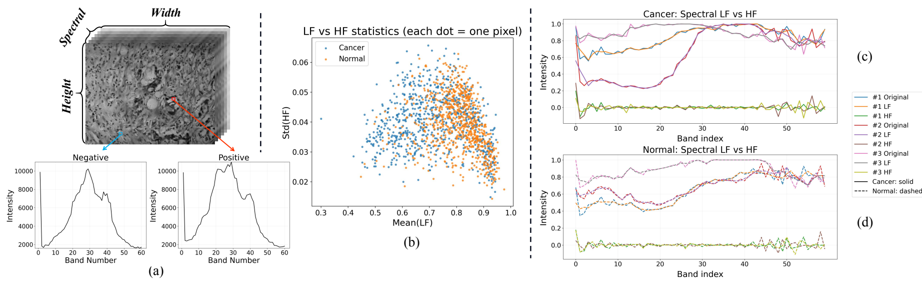


**Fig. 1.** (a) A microscopic hyperspectral image from the Multi-Dimensional Choledoch Dataset (MDC). Positive and negative regions exhibit distinct spectral profiles. (b) Pixel-wise statistics of low-frequency (LF) mean and high-frequency (HF) variation, which show that the HF spectral statistics exhibit substantial overlap between cancerous and normal tissue, whereas the LF mean distributions show clearer separation, indicating that LF components carry more discriminative tissue-level information. (c) Representative cancer spectra with corresponding LF and HF components. (d) Representative normal spectra with corresponding LF and HF components.

To address these challenges, a wide range of methods have been developed for medical hyperspectral image segmentation. CNN-based architectures [6,13] are widely adopted due to their strong local feature extraction, and are often combined with spectral dimensionality reduction (e.g., PCA [15] and low-rank decomposition [20]) to reduce redundancy and improve efficiency. Dual-stream designs [22] further model spatial structures and spectral signatures in parallel with feature fusion. However, their local receptive fields limit modeling of long-range spectral dependencies and global spectral trends. To mitigate this limitation, attention-based models have been introduced in hyperspectral analysis to model long-range spectral–spatial dependencies, including gated axial attention [18] and Transformer-based architectures [5,8,19], but the quadratic cost of self-attention can be prohibitive for high-dimensional MHSI. More recently, selective state-space models such as Mamba [10,14] provide linear complexity, yet their unidirectional aggregation and centralized spectral representations may still limit global and continuous spectral modeling. From a representation perspective, most existing approaches encode spectra as channel-wise features or tokenized sequences, treating the continuous signature as a flat sequence without explicit scale separation. However, in MHSI, spectral curves are intrinsically multi-scale, coupling global biochemical trends with local spectral variations [16].

The wavelet transform has been widely used in hyperspectral imaging for denoising, compression, and spectral analysis because of its multiscale frequency localization property. Early research on MHSI [2,12] has shown that wavelet decomposition along the spectral dimension can facilitate tissue-aware spectral characterization and segmentation. More recent efforts have further applied wavelet-based spectral compression to medical SWIR hyperspectral data [3], aiming to reduce spectral dimensionality while preserving diagnostically relevant spectral fidelity in clinical datasets. In parallel, recent work in remote sensing [1] has explored integrating wavelet-based multiscale representations with deep learning architectures for hyperspectral image analysis, demonstrating the technical feasibility of wavelet-guided deep models.

Despite these efforts, wavelet transforms are typically used as fixed preprocessing tools for denoising or dimensionality reduction in MHSI, rather than being integrated into the representation learning process. Moreover, multi-frequency fusion strategies developed in remote sensing may not align with the distinct spectral characteristics of clinical MHSI, where low-frequency (LF) components tend to capture stable global trends while high-frequency (HF) components are often less reliable. This motivates us to revisit spectral wavelet decomposition as a structural prior for deep spectral modeling.

In this paper, we propose a wavelet-guided spectral–spatial segmentation framework that explicitly reparameterizes spectral representations in the wavelet domain. Rather than treating wavelet decomposition itself as the main novelty, we inject LF components as adaptive structural priors through a per-pixel gated mechanism to guide deep spectral learning. To further strengthen spectral–spatial interaction, we incorporate Spectral-Enhanced Skip Connections and Band-Aware Aggregation modules to adaptively reweight band-group features in the skip pathways used for decoding. Together, these designs provide a structured alternative to conventional spectral encoding strategies for MHSI segmentation.

## 2 Methodology

### 2.1 Problem Definition and Network Overview

Mathematically, let $\mathbf{X} \in \mathbb{R}^{H\times W\times S}$ denote a MHSI, where $H \times W$ represents the spatial resolution and $S$ denotes the number of spectral bands. The task is to predict a dense pixel-wise label map $\hat{\mathbf{Y}} \in \{0,1\}^{H\times W}$. Given a training set $\mathcal{D} = \{(\mathbf{X}_i, \mathbf{Y}_i)\}_{i=1}^{N}$ with $N$ training samples, where $\mathbf{Y}_i$ is the ground-truth annotation corresponding to $\mathbf{X}_i$, our objective is to learn a mapping function that accurately assigns tissue labels to each spatial location by jointly modeling spatial structures and spectral signatures.

An overview of the proposed framework is illustrated in Fig. 2. We design a wavelet-guided spectral–spatial architecture to jointly model spatial morphology and spectral signatures in medical hyperspectral images. The spectral stream explicitly reparameterizes spectral representations in the wavelet domain and incorporates a gated LF structural prior to stabilize spectral modeling. The spatial stream processes grouped spectral bands using a 2D backbone, where Band-Aware Aggregation adaptively fuses band-group features across encoder stages. To preserve spectral context during reconstruction, Spectral-Enhanced Skip Connections inject spectrally informative cues into the decoding pathway.

### 2.2 Wavelet-based Spectral Stabilization Prior

As illustrated in Fig. 1(b), the high-frequency spectral statistics exhibit substantial overlap between cancerous and normal tissues, indicating limited discrim-

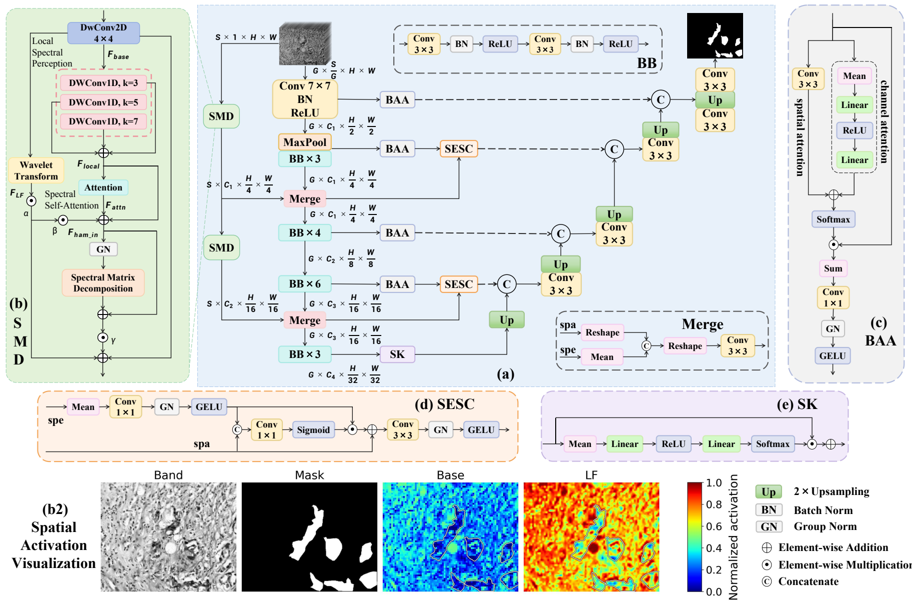


**Fig. 2.** Framework architecture of Wavelet-Guided Spectral–Spatial Segmentation. (a) Overall spectral–spatial architecture; (b) Spectral Modulation and Decomposition (SMD) module; (b2) Spatial activation visualization of $F_{\text{base}}$ and $F_{\text{LF}}$, where white contours indicate pathological regions; (c) Band-Aware Aggregation (BAA); (d) Spectral-Enhanced Skip Connections (SESC); (e) Selective Kernel (SK) block.

inative power. In contrast, low-frequency components capture smooth, tissue-specific spectral trends that remain consistent across different spatial locations. This observation motivates us to integrate wavelet-based LF decomposition into the network as a structural stabilization prior for deep spectral modeling, rather than treating it as an external preprocessing operation.

Guided by this insight, we design the Spectral Modulation and Decomposition (SMD) module, in which LF information is injected as a sigmoid-gated residual prior.

### 2.3 Spectral Modulation and Decomposition (SMD)

Given an input MHSI, we first apply a depthwise spectral convolution to obtain compact spectral features, and reshape them into per-pixel spectral sequences $F_{\text{base}}$. Next, we model local spectral continuity using a lightweight depthwise 1D convolution along the spectral dimension, generating locally enhanced features $F_{\text{local}}$. To capture spectral patterns with different receptive field sizes along the spectral dimension, we employ multi-scale depthwise convolutions with multiple kernel sizes ($k = 3, 5, 7$) and fuse them via a residual summation.

Based on $F_{\mathrm{local}}$, we further apply a lightweight spectral self-attention mechanism to aggregate long-range dependencies across distant wavelength bands, thus obtaining a globally contextualized spectral representation $F_{\mathrm{attn}}$.

Simultaneously, we extract a LF spectral trend prior by applying a one-level wavelet decomposition along the spectral dimension:

$$F_{\mathrm{LF}} = \mathcal{W}_{\mathrm{db2}}(F_{\mathrm{base}}), \tag{1}$$

where $\mathcal{W}_{\mathrm{db2}}(\cdot)$ denotes a one-level discrete wavelet transform implemented with fixed Daubechies-2 low-pass filters, and only the approximation component is retained. Subsequently, linear interpolation is performed to align the spectral length with $S$. As shown in Fig. 2(b2), $F_{\mathrm{LF}}$ produces more spatially coherent and less fragmented activations compared to $F_{\mathrm{base}}$, suggesting improved structural stability in spectral representation learning.

Rather than replacing the original spectrum or fusing all frequency scales, we inject $F_{\mathrm{LF}}$ as a stabilizing residual prior via a pixel-adaptive gate $\alpha \in (0,1)$,

$$\alpha = \sigma\left(g([\mu(F_{\mathrm{base}}),\ \mathrm{std}(F_{\mathrm{base}}),\ \mu(F_{\mathrm{LF}})])\right), \tag{2}$$

where $\mu(\cdot)$ and $\mathrm{std}(\cdot)$ are computed along the spectral dimension, $g(\cdot)$ is a lightweight MLP, and $\sigma$ represents the sigmoid function.

We incorporate the gated LF prior in two stages. First, it gently modulates the input to the spectral matrix decomposition (Hamburger [7]) as

$$F_{\mathrm{ham\_in}} = F_{\mathrm{local}} + F_{\mathrm{attn}} + \beta \cdot \alpha \cdot F_{\mathrm{LF}}, \tag{3}$$

where $\beta \in (0, 0.2)$ is a learnable scaling factor parameterized via a sigmoid function to keep the LF guidance subtle. The Hamburger output is further scaled by a learnable gate $\gamma$ before residual fusion. Second, the gated LF prior is added again as a residual term to provide a stable LF correction at the output (Fig. 2(b)).

### 2.4 Spectral-Aware Feature Aggregation and Decoding

**Band-Aware Aggregation (BAA).** In the spatial stream, spectral bands are evenly divided into $G$ groups (spectral agents [20]), each group is processed independently by the 2D encoder. Conventional dual-stream approaches aggregate spectral agents via simple averaging, implicitly assuming equal contributions from all band groups. However, different wavelength groups encode distinct biochemical properties and may exhibit spatially varying discriminative power.

To address this, we propose a BAA module to learn adaptive per-pixel importance weights across spectral groups. Given grouped features $F_i = \{F_i^{(g)}\}_{g=1}^{G}$, the group-wise weights are computed as $w_g = \mathrm{Softmax}_g\left(\mathrm{CA}(F_i^{(g)}) \oplus \mathrm{SA}(F_i^{(g)})\right)$, where $\mathrm{CA}(\cdot)$ and $\mathrm{SA}(\cdot)$ denote channel and spatial attention operations. The aggregated feature is

$$F_i^{\mathrm{BAA}} = \sum_{g=1}^{G} w_g \odot F_i^{(g)}. \tag{4}$$

This design enables adaptive band-group fusion and improves spectral discriminability in spatial representations (see Fig. 2(c)). For the final encoder stage with low spatial resolution, we adopt a simplified global group fusion strategy inspired by Selective Kernel (SK) Networks [9] (Fig. 2(e)), because spatially adaptive weighting becomes less critical at coarse feature scales.

**Spectral-Enhanced Skip Connections (SESC).** In conventional dual-stream architectures [10,20], spectral features are typically injected only at encoder stages and are not explicitly propagated to the decoder. To preserve spectral context during spatial reconstruction, we introduce Spectral-Enhanced Skip Connections.

Given spatial encoder feature $F_i^{\text{spa}}$ and spectral feature $F_i^{\text{spe}}$, the spectral tokens are first averaged along the band dimension and projected to the spatial feature space via a $1 \times 1$ convolution, yielding a projected feature $F_i^{\text{proj}}$. A spatially adaptive gating mask is then learned as

$$\delta = \sigma\Big(\text{Conv}\left([F_i^{\text{spa}}, F_i^{\text{proj}}]\right)\Big), \tag{5}$$

where $[\cdot,\cdot]$ denotes concatenation and $\sigma$ is the sigmoid function.

The enhanced skip feature is computed as

$$\tilde{F}_i = F_i^{\text{spa}} + \delta \cdot F_i^{\text{proj}}. \tag{6}$$

This mechanism selectively injects spectral information into the decoding pathway, enabling spatial features to leverage discriminative spectral cues when necessary (Fig. 2(d)).

## 3 Experimental Results

### 3.1 Experimental Setup

**MDC Dataset.** The Multi-Dimensional Choledoch Dataset (MDC) [21] contains 538 microscopic hyperspectral scenes with pixel-level annotations, acquired by AOTF-based wavelength switching and sCMOS imaging under a microscope, where each band records the transmitted intensity at a specific wavelength. Each scene comprises 60 spectral bands spanning the range of 550 nm to 1000 nm, and the spatial resolution of each band image is resized to 256 × 320.

**MOD Dataset.** The Multispectral Oral Disease Image Dataset (MOD) [4] contains 138 oral imaging scenes acquired in 16 spectral bands spanning 460–600 nm, using a multispectral reflectance imaging system where each band records reflected intensity at a specific wavelength. All images are resized to 256 × 256.

For the MDC dataset, following [17], we adopt a patient-centric evaluation protocol, partitioning the data into training, validation, and test sets in a 3:1:1 ratio. The split is performed at the patient-ID level, ensuring that all samples from the same patient are assigned to the same subset. For the MOD dataset, we follow a standard 8:1:1 split for training, validation, and testing, respectively.

**Table 1.** Performance comparison with state-of-the-art (SOTA) methods on MDC and MOD datasets. The best results are **highlighted**.

| Method | | MDC | | | MOD | | |
|---|---|---|---|---|---|---|---|
| | | **IoU↑** | **DSC↑** | **HD↓** | **IoU↑** | **DSC↑** | **HD↓** |
| **CNN** | PCA-UNet [15] | 54.63 | 68.83 | 80.60 | 47.08 | 62.76 | 54.21 |
| | 3D UNet [6] | 56.43 | 70.69 | 80.07 | 47.37 | 62.95 | 47.53 |
| | FSS [20] | 60.43 | 73.72 | 78.50 | 50.59 | 65.83 | 47.34 |
| **Mamba** | VMUNet [14] | 60.56 | 73.63 | 78.28 | 53.04 | 68.58 | 46.53 |
| | MDN [10] | 59.67 | 73.20 | 80.44 | 54.65 | 69.69 | 51.20 |
| **Attention/ Transformer** | Spec-Tr [19] | 61.14 | 74.55 | **76.37** | 53.50 | 67.99 | 43.42 |
| | GSA-Net [18] | 58.37 | 72.04 | 79.43 | 42.74 | 57.47 | 70.83 |
| **Ours** | | **64.34** | **76.76** | 76.44 | **55.33** | **70.05** | **43.37** |

We apply data augmentation including random flipping and rotation during training. The model is trained for 100 epochs with a batch size of 8, using the AdamW optimizer with an initial learning rate of $5 \times 10^{-4}$. The training loss combines Dice loss, BCE loss, and an auxiliary boundary loss, where the boundary target is automatically generated from the ground-truth mask to supervise the predicted edge map without additional annotations. Segmentation performance is primarily evaluated using Dice Similarity Coefficient (DSC, %) and Intersection over Union (IoU, %). Additionally, Hausdorff Distance (HD, px) is reported for boundary quality assessment. For fair comparison, all competing methods were re-implemented and retrained from scratch under the same data splits, preprocessing pipeline, data augmentation strategies, and experimental protocol. All experiments are implemented in PyTorch and conducted on NVIDIA GH200 GPUs.

### 3.2 Performance Evaluation

Table 1 presents the quantitative comparison with state-of-the-art methods on MDC and MOD datasets. Compared with models based on CNNs, Mamba, and Transformers, our approach achieves the best IoU (64.34%) and DSC (76.76%) on the MDC dataset, outperforming the strongest competing method (Spec-Tr) by 2.21% in DSC. This improvement is not simply due to increased model size, as our model has slightly fewer parameters than FSS (26.61M vs. 27.06M), and the internal 1D wavelet operation accounts for only 1.1% of the total forward time. These results demonstrate that our method consistently improves segmentation accuracy while maintaining competitive boundary performance. Similar trends are observed on the MOD dataset, indicating good cross-dataset generalization of the proposed framework.

Qualitative comparisons further support the quantitative results. As illustrated in Fig. 3, on both MDC and MOD samples, competing methods tend

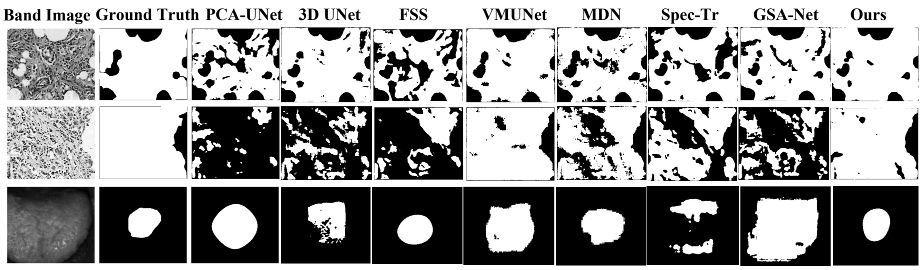


**Fig. 3.** Visual Comparisons on MDC (top two rows) and MOD (bottom row) datasets.

**Table 2.** Ablation study of wavelet-based spectral designs and various decoding modules on the MDC dataset. Our final model (Residual LF + Db2 + SK + BAA + SESC) corresponds to the last row. The best results are **highlighted**.

| Wavelet | | | | | | | | Decoding | | | Metrics | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Subband Mode | | | Integration | | Basis | | | | | | | | |
| HF | LF+HF | LF | Residual | Pure | Haar | Sym2 | Db2 | SK | BAA | SESC | IoU↑ | DSC↑ | HD↓ |
| ✓ | × | × | ✓ | × | × | × | ✓ | × | × | × | 62.52 | 75.43 | 77.54 |
| × | ✓ | × | ✓ | × | × | × | ✓ | × | × | × | 62.84 | 75.58 | 78.88 |
| × | × | ✓ | × | ✓ | × | × | ✓ | × | × | × | 61.00 | 74.18 | 79.38 |
| × | × | ✓ | ✓ | × | ✓ | × | × | × | × | × | 62.68 | 75.37 | 76.74 |
| × | × | ✓ | ✓ | × | × | ✓ | × | × | × | × | 62.16 | 75.38 | 76.47 |
| × | × | ✓ | ✓ | × | × | × | ✓ | × | × | × | 63.18 | 75.71 | 76.41 |
| × | × | ✓ | ✓ | × | × | × | ✓ | ✓ | × | × | 63.19 | 76.13 | **74.74** |
| × | × | ✓ | ✓ | × | × | × | ✓ | ✓ | ✓ | × | 63.76 | 76.19 | 77.92 |
| × | × | ✓ | ✓ | × | × | × | ✓ | ✓ | ✓ | ✓ | **64.34** | **76.76** | 76.44 |

to produce fragmented predictions, boundary leakage, or missed small regions. In contrast, our approach generates more continuous segmentation maps with clearer boundary delineation, aligning better with the ground truth. These results demonstrate the effectiveness of the proposed residual spectral modeling and structured decoding strategy for medical hyperspectral image segmentation.

### 3.3 Ablation Study

We conduct an ablation study to evaluate different wavelet-based spectral designs and decoding modules, as summarized in Table 2. Replacing the original spectrum with pure LF components results in a noticeable performance drop (74.18% DSC), indicating that LF trends alone are insufficient. Direct concatenation of LF and HF components provides only a minor improvement (75.58%), while the Haar basis performs worse than db2 (75.37% vs. 75.71%), demonstrating the advantage of smoother wavelet decomposition. Introducing LF as a gated residual prior yields consistent gains (75.71%), and progressively incorporating decoding modules further improves performance: adding SK increases DSC to 76.13%, and the full decoding pipeline (SK + BAA + SESC) achieves the best result of 76.76%. These results indicate that residual LF modeling and structured decoding modules contribute complementary improvements to spectral–spatial segmentation performance.

## 4 Conclusion

In this paper, we propose a wavelet-guided spectral–spatial framework for medical hyperspectral image segmentation. By redesigning the spectral stream with wavelet-domain reparameterization, we incorporate low-frequency components as adaptive structural priors to stabilize deep spectral modeling. Furthermore, Band-Aware Aggregation and Spectral-Enhanced Skip Connections enhance spectral–spatial interaction during decoding. Extensive experiments on the MDC and MOD datasets demonstrate consistent performance improvements over the strongest competing SOTA method, achieving gains of 2.21% in DSC and 3.20% in IoU on the MDC dataset. These results validate the effectiveness of structured multi-scale spectral modeling for clinical MHSI data.

**Acknowledgments.** This work is supported by the China Scholarship Council (CSC) under Grant 202408410119.

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

## References

1. Ahmad, M., Usama, M., Mazzara, M., Distefano, S.: Wavemamba: Spatial-spectral wavelet mamba for hyperspectral image classification. IEEE Geoscience and Remote Sensing Letters **22**, 1–5 (2024)
2. Akbari, H., Kosugi, Y., Kojima, K., Tanaka, N.: Wavelet-based compression and segmentation of hyperspectral images in surgery. In: International Workshop on Medical Imaging and Virtual Reality. pp. 142–149. Springer (2008)
3. Biswas, H., Tang, R., Mollah, S., Berezin, M.Y.: Wavelet-based compression method for scale-preserving swir hyperspectral data. medRxiv (2025)
4. Chand, S., Namasivayam, K., Dave, J., Preejith, S., Jayachandran, S., Sivaprakasam, M.: In-vivo non-contact multispectral oral disease image dataset with segmentation. Scientific Data **11**(1), 1298 (2024)
5. Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A.L., Zhou, Y.: Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)
6. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3d u-net: learning dense volumetric segmentation from sparse annotation. In: International conference on medical image computing and computer-assisted intervention. pp. 424–432. Springer (2016)
7. Geng, Z., Guo, M.H., Chen, H., Li, X., Wei, K., Lin, Z.: Is attention better than matrix decomposition? arXiv preprint arXiv:2109.04553 (2021)

8. Hong, D., Han, Z., Yao, J., Gao, L., Zhang, B., Plaza, A., Chanussot, J.: Spectralformer: Rethinking hyperspectral image classification with transformers. IEEE Transactions on Geoscience and Remote Sensing **60**, 1–15 (2021)
9. Li, X., Wang, W., Hu, X., Yang, J.: Selective kernel networks. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 510–519 (2019)
10. Lin, S., Yun, B., Shen, W., Li, Q., Yang, A., Wang, Y.: Mdn: mamba-driven dual-stream network for medical hyperspectral image segmentation. In: ICASSP 2025-2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1–5. IEEE (2025)
11. Lu, G., Fei, B.: Medical hyperspectral imaging: a review. Journal of biomedical optics **19**(1), 010901–010901 (2014)
12. Rajpoot, K.M., Rajpoot, N.M.: Wavelet based segmentation of hyperspectral colon tissue imagery. In: 7th International Multi Topic Conference, 2003. INMIC 2003. pp. 38–43. IEEE (2003)
13. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical image computing and computer-assisted intervention. pp. 234–241. Springer (2015)
14. Ruan, J., Li, J., Xiang, S.: Vm-unet: Vision mamba unet for medical image segmentation. ACM Transactions on Multimedia Computing, Communications and Applications (2024)
15. Wang, J., Wang, Y., Tao, X., Li, Q., Sun, L., Chen, J., Zhou, M., Hu, M., Zhou, X.: Pca-u-net based breast cancer nest segmentation from microarray hyperspectral images. Fundamental Research **1**(5), 631–640 (2021)
16. Wołk, K., Wołk, A.: Hyperspectral imaging system applications in healthcare. Electronics (2079-9292) **14**(23) (2025)
17. Xie, X., Jin, T., Yun, B., Li, Q., Wang, Y.: Exploring hyperspectral histopathology image segmentation from a deformable perspective. In: Proceedings of the 31st ACM International Conference on Multimedia. pp. 242–251 (2023)
18. Xue, J., Chen, X., Kim, S.H.: Adapting gated axial attention for microscopic hyperspectral cholangiocarcinoma image segmentation. Electronics **14**(20), 3979 (2025)
19. Yun, B., Lei, B., Chen, J., Wang, H., Qiu, S., Shen, W., Li, Q., Wang, Y.: Spectr: Spectral transformer for microscopic hyperspectral pathology image segmentation. IEEE Transactions on Circuits and Systems for Video Technology **34**(6), 4610–4624 (2023)
20. Yun, B., Li, Q., Mitrofanova, L., Zhou, C., Wang, Y.: Factor space and spectrum for medical hyperspectral image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 152–162. Springer (2023)
21. Zhang, Q., Li, Q., Yu, G., Sun, L., Zhou, M., Chu, J.: A multidimensional choledoch database and benchmarks for cholangiocarcinoma diagnosis. IEEE access **7**, 149414–149421 (2019)
22. Zhang, Q., Pei, G., Wang, Y.: Omni-fusion of spatial and spectral for hyperspectral image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 471–481. Springer (2025)