

# BISeg: Band-Interval Selection for Explainable Hyperspectral Segmentation

Compact Spectral Windows via Center–Width Parameterization with Length and Count Penalties

Jaek Bae

Department of AI Convergence  
Gwangju Institute of Science and Technology (GIST)  
Gwangju, Republic of Korea  
jaeikb38@gm.gist.ac.kr

Yong-Gu Lee<sup>†</sup>

Department of AI Convergence  
Gwangju Institute of Science and Technology (GIST)  
Gwangju, Republic of Korea  
lygu@gist.ac.kr

## ABSTRACT

We introduce BISeg, a novel hyperspectral image (HSI) segmentation model that achieves both accuracy and interpretability through band-interval selection. Unlike conventional approaches that rely on post-hoc band attribution, BISeg intrinsically learns to identify compact spectral windows during training. Each interval is parameterized by its center wavelength and width, with a maximum of ten intervals activated. A dedicated regularization encourages the use of fewer and narrower intervals, ensuring explanations remain concise and physically meaningful. The proposed spectral–spatial backbone integrates these intervals with spatial features via cross-axis attention, and a query-based head outputs pixel-level mask. Experiments on our custom HSI dataset, covering nine material classes, demonstrate that BISeg achieves competitive segmentation accuracy while producing interpretable spectral intervals aligned with known material properties. This work demonstrates the potential of interval-based intrinsic interpretability to enhance human understanding of hyperspectral image segmentation. Furthermore, BISeg’s interval selection mechanism provides direct insight into which wavelength regions contribute most to material discrimination, enabling domain experts to validate model reasoning against known spectral signatures. We also visualize the learned intervals across multiple samples, revealing consistent selection patterns for similar materials and distinct shifts for spectrally ambiguous ones. Compared to attention-based or saliency-driven interpretability methods, BISeg’s intrinsic design avoids the instability and noise often observed in post-hoc explanations. In addition, the model maintains computational efficiency, as the compact interval representation substantially reduces spectral redundancy without degrading accuracy. Overall, BISeg bridges the gap between performance-driven and interpretation-oriented hyperspectral segmentation, suggesting a new direction for transparent AI in remote sensing and material analysis.

<sup>†</sup>Corresponding author: Yong-Gu Lee  
© 2025 Copyright held by the owner/author(s).

## KEYWORDS

Hyperspectral imaging, Segmentation, Explainable AI

## 1 Introduction

Hyperspectral imaging (HSI) captures hundreds of contiguous spectral bands, providing detailed material signatures that are invisible to conventional RGB cameras. This richness has enabled applications in remote sensing, agriculture, environmental monitoring, and material science. However, using the full spectral cube presents two major challenges:

1. Redundancy and noise: Many bands are highly correlated, and sensor noise often dominates at spectral extremes.
2. Interpretability: Deep learning models that process all bands equally do not reveal which spectral regions are truly responsible for their predictions.

Post-hoc explain-ability methods such as band attribution curves attempt to identify important wavelengths after training. Yet they remain limited: they do not influence the model’s internal reasoning and often highlight scattered, physically inconsistent bands. This gap prevents HSI models from being fully trustworthy in human-centered applications.

To address this, we propose BISeg, a segmentation framework that introduces band-interval selection as an intrinsic learning mechanism. Instead of analyzing importance retrospectively, BISeg directly learns spectral windows during training. Each window is defined by a center wavelength and width, with a maximum of ten intervals. A regularization scheme penalizes both interval length and count, thereby encouraging the model to use fewer, narrower, and more interpretable spectral regions.

BISeg integrates these learned intervals into a spectral–spatial backbone via cross-axis attention, and a query-based segmentation head produces pixel-level masks. Experiments on a custom HSI dataset demonstrate that BISeg not only achieves competitive segmentation accuracy but also yields compact spectral intervals aligned with material properties.

Our contributions are summarized as follows:

1. We introduce a band-interval selection module that learns center–width parameterized spectral windows within the network.
2. We design a regularization strategy that enforces sparsity and compactness of intervals.
3. We integrate the interval-masked spectral features into a spectral–spatial backbone for segmentation.
4. We validate BISeg on a custom HSI dataset, showing both high segmentation performance and interpretable interval patterns.

This work demonstrates the potential of interval-based intrinsic interpretability for hyperspectral segmentation, advancing toward more transparent and human-centered AI systems.

## 2 Related Work

### 2.1 Deep Learning for HSI Segmentation

Spectral–spatial attention modules enhance CNN backbones and remain deployment-friendly [5], while early 3D residual networks [4] and RL-based selection [6] chart complementary routes to compact, high-performing HSI models. These threads collectively point to the need for continuity-aware selection under explicit usage constraints. BISeg addresses this by learning a small number of contiguous spectral windows and regularizing their number and width, yielding accurate segmentation with faithful, compact spectral evidence.

### 2.2 Spectral Band Selection

Embedded selection integrates band choice into training; EHBS uses stochastic gates with  $l_0$ -style regularization to keep a compact subset without sacrificing segmentation accuracy [2]. Another strand frames band choice as sequential decision making via deep reinforcement learning to balance compactness and accuracy [6]. Because these methods typically select discrete individual bands, they may disrupt spectral continuity; BISeg instead learns center–width intervals with penalties on length and count, preserving contiguous, interpretable ranges.

### 2.3 Explainability in Hyperspectral Models

Explainability has been explored via attention and post-hoc attribution. EFBASN combines attention with feature alignment to highlight material-relevant wavelengths while maintaining strong accuracy [3], but attention remains soft and post-hoc explanations do not alter the decision path. BISeg makes the selected intervals part of the forward computation, ensuring that explanations correspond to wavelengths the model truly uses.

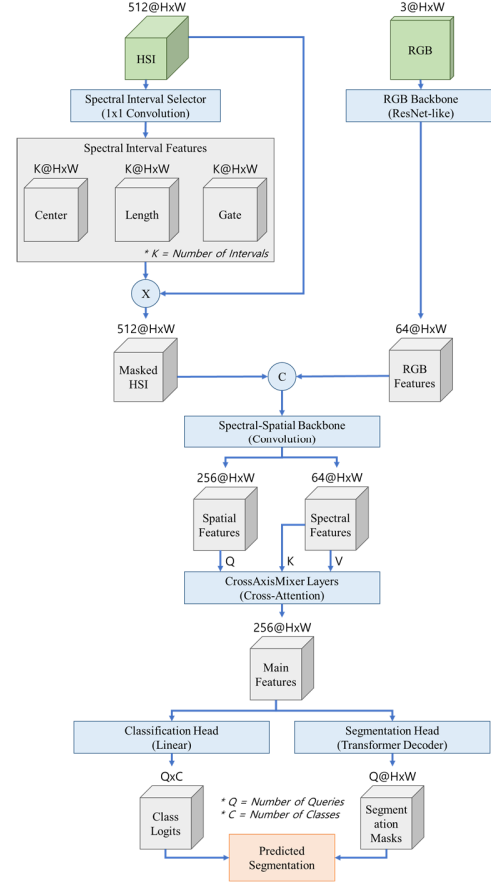
### 2.4 Additional Directions

Early 3D residual CNNs (SSRN) established strong spectral–spatial baselines; spectral–spatial attention further enhanced CNN backbones; and deep-RL framed band choice as sequential

decision making [4–6]. These threads motivate BISeg’s intrinsic, interval-level selection rather than isolated band scoring.

## 3 Method

### 3.1 Overall Architecture



**Figure 1: Overall architecture of BISeg. The Spectral Interval Selector learns up to  $K$  intervals (center, length, gate) to mask HSI inputs. Masked hyperspectral and RGB features are fused via a spectral–spatial backbone and cross-axis mixer, and task-specific heads output class logits and segmentation masks.**

Figure 1 shows the overall architecture of BISeg. The framework begins with a Spectral Interval Selector that learns up to  $K$  spectral windows, each parameterized by center, length, and gate. The masked HSI features are then fused with RGB features through a spectral–spatial backbone and cross-axis mixer. Finally, classification and segmentation heads output class logits and segmentation masks, producing interpretable results aligned with the selected spectral intervals.

### 3.2 Spectral Interval Selector

The hyperspectral input cube is first processed by a Spectral Interval Selector, implemented as  $1 \times 1$  convolutions. This module predicts three parameters for each interval: center, length, and gate. A maximum of  $K$  intervals is allowed ( $K \leq 10$ ). The gate determines whether an interval is active, while the center and length define its spectral window. The resulting interval masks are applied to the hyperspectral cube, producing a masked HSI that preserves only the most informative spectral ranges.

### 3.3 Multimodal Fusion

The masked hyperspectral features are then combined with RGB features extracted by an auxiliary backbone (ResNet-like). The concatenated representation is passed to the Spectral-Spatial Backbone, which applies convolutional layers to separately encode spectral features and spatial features.

### 3.4 Cross-Axis Mixer

To capture interactions across domains, CrossAxisMixer layers perform cross-attention between spectral and spatial features. This design ensures that spatial context can guide spectral selection and vice versa, producing a set of main fused features.

### 3.5 Segmentation and Classification Heads

The fused representation is fed into two parallel heads. The Segmentation Head (transformer decoder) predicts query-based segmentation masks, while the Classification Head (linear projection) outputs class logits. The final predicted segmentation is obtained by combining mask predictions with class assignments, resulting in an interpretable segmentation map aligned with the learned spectral intervals.

## 4 Experiments

### 4.1 Dataset

We evaluate BISeg on a custom hyperspectral dataset covering nine material classes: iron, glass, leather, plastic, rock, rubber, stainless steel, water, and wood. Each sample contains 512 spectral bands (420–1728 nm) with a spatial resolution of  $320 \times 320$  pixels. Pixel-level annotations are provided, with background pixels labeled as 255.

### 4.2 Implementation Details

All experiments are implemented in PyTorch 2.3.1 with CUDA 12.1 and trained on  $8 \times$  NVIDIA RTX A6000 GPUs. Mixed precision is enabled to optimize memory usage and training speed. We adopt the AdamW optimizer with an initial learning rate of  $1e-4$  and weight decay of  $1e-4$ . A cosine learning rate scheduler is applied throughout 500 epochs. Batch size is set to 8–10 depending on GPU memory availability. Interval regularization is applied with penalties on both length and count, with a maximum of  $K=10$  intervals.

### 4.3 Evaluation Metrics

We evaluate BISeg using three standard segmentation metrics:

1. **Overall Accuracy (OA)**: the percentage of correctly classified pixels across all classes.
2. **Mean Intersection-over-Union (mIoU)**: the average IoU across classes, measuring overlap between predicted and ground-truth masks.
3. **Mean F1-Score**: the harmonic mean of precision and recall, averaged across classes, reflecting balance between false positives and false negatives.

In addition, we analyze interpretability by reporting the average number of active intervals and average interval width, which indicate how compactly BISeg represents the spectrum.

### 4.4 Results

**Table 1: Quantitative segmentation results under different maximum interval settings ( $K=3, 5, 10, 20$ ). Best values are in bold in underline. And second-best values are in bold**

$K$ (max intervals)	OA (%)	mIoU (%)	Mean F1 (%)
3	80.27	66.60	69.08
5	<b>82.50</b>	<b>77.59</b>	<b>79.06</b>
10	<b><u>82.58</u></b>	<b><u>78.85</u></b>	<b><u>80.24</u></b>
20	81.08	59.24	61.68

**Table 2: Interval statistics showing the number of active intervals and average interval width under different  $K$  ( $K=3, 5, 10, 20$ ).**

$K$ (max intervals)	Avg. Active Intervals	Avg. Width (bands)	Avg. Gate
3	3	76.93	0.92
5	5	46.27	0.89
10	10	23.14	0.89
20	20	11.55	0.83

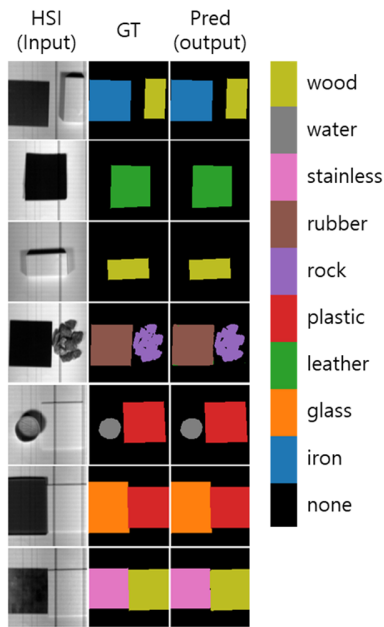
*4.4.1 Quantitative Results.* Table 1 summarizes OA, mIoU, and Mean F1 under different maximum interval settings ( $K=3, 5, 10, 20$ ). Increasing  $K$  improves segmentation accuracy, with  $K=10$  yielding the best performance (OA  $\approx 82.6\%$ , mIoU  $\approx 78.9\%$ , Mean F1  $\approx 80.2\%$ ). However,  $K=5$  offers almost equal results, achieving strong accuracy while maintaining fewer intervals for interpretability.

*4.4.2 Coverage-preserving compensation.* Table 2 indicates a nearly constant covered spectral support of  $\approx 231$  bands across  $K$ .

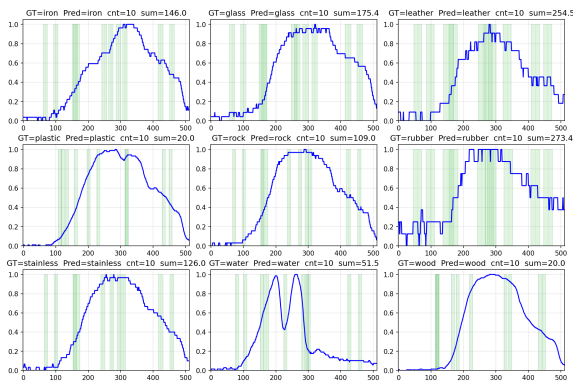
When  $K$  is smaller (fewer intervals), each interval widens to preserve coverage; as  $K$  increases, the model partitions the same support into more, narrower windows. The average gate magnitude remains high and stable ( $0.92 \rightarrow 0.89 \rightarrow 0.89$ ), showing

BISeg maintains activation strength while trading interval count for width instead of collapsing spectral coverage.

**4.4.3 Qualitative Visualization.** Figure 2 presents segmentation maps for the best-performing setting ( $K=10$ ). BISeg accurately delineates object boundaries across different materials. Figure 3 illustrates the learned spectral intervals, showing that selected ranges align with known absorption features (e.g., NIR absorption for water, visible absorption for iron oxide). The green bars intuitively show that the regions selected by the model include the characteristic ranges of the corresponding wavelengths.



**Figure 2: Qualitative segmentation results. From left to right: input HSI (center band), ground truth, and BISeg prediction ( $K=10$ ). BISeg accurately delineates both small irregular objects and large homogeneous regions.**



**Figure 3: Visualization of learned spectral intervals. The blue curve shows the reflectance spectrum, and green bars indicate the learned intervals selected by BISeg. The learned windows align with known absorption features for different materials.**

## 5 Conclusion

We presented BISeg, an intrinsically interpretable HSI segmentation model that learns contiguous spectral intervals via center-width parameterization with length/count penalties. On a custom dataset (9 classes, 12 objects, 78 images), BISeg attained strong results—up to mIoU 78.85% (OA 82.58%, Mean F1 80.24) at  $K=10$ , while  $K=5$  offered similar accuracy with fewer windows. Interval statistics reveal a coverage-preserving behavior: fewer gates widen intervals, more gates split the same support into narrower windows ( $\sim 231$  bands in total). Thus, BISeg balances accuracy and interpretability without post-hoc explanations, tying decisions to compact, physically meaningful spectral ranges.

## 6 Limitation & Discussion

### 6.1 Empirical behavior

BISeg maintains a nearly constant spectral support ( $\sim 231$  bands) across  $K \in \{3, 5, 10, 20\}$ : fewer intervals are wider; more intervals are narrower, with similar average gates ( $\sim 0.9$ ).

### 6.2 Limitations

- (i) Scale/diversity: small, controlled dataset; generalization to other sensors/scenes remains open.
- (ii) Baselines: evaluation relies on ablations (varying  $K$ ), not external baselines.
- (iii) Design bias: contiguous-interval parameterization may under-express very narrow or scattered cues unless  $K$ /widths adapt.
- (iv) Compute/tuning: end-to-end training benefited from multi-GPU; regularization strengths may need per-sensor tuning.

### 6.3 Future Work

Public benchmarks and cross-sensor validation; adding external baselines (attention/embedded selection); adaptive  $K$  per class/region and multi-scale intervals; interactive tools for inspecting windows and masks.

## ACKNOWLEDGMENTS

This work was supported by Center for Applied Research in Artificial Intelligence (CARAI) grant funded by Defense Acquisition Program Administration (DAPA) and Agency for Defense Development (ADD) (UD230017TD)

## REFERENCES

- [1] Danfeng Hong, Zhu Han, Jing Yao, Lianru Gao, Bing Zhang, Antonio Plaza, and Jocelyn Chanussot. 2022. SpectralFormer: Rethinking Hyperspectral Image Classification with Transformers. *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022), 1–15. DOI: <https://doi.org/10.1109/TGRS.2021.3130716>
- [2] Yaniv Zimmer, Ofir Lindenbaum, and Oren Glickman. 2024. Embedded Hyperspectral Band Selection with Adaptive Optimization for Image Semantic Segmentation. *arXiv preprint arXiv:2401.11420* (2024). DOI: <https://doi.org/10.48550/arXiv.2401.11420>
- [3] Jirui Liu, Jinhui Lan, Yiliang Zeng, Wei Luo, Zhixuan Zhuang, and Jinlin Zou. 2025. Explainability Feature Bands Adaptive Selection for Hyperspectral Image Classification. *Remote Sensing* 17, 9 (2025), Article 1620. DOI: <https://doi.org/10.3390/rs17091620>

- [4] Zilong Zhong, Jonathan Li, Zhiming Luo, and Michael Chapman. 2018. Spectral–Spatial Residual Network for Hyperspectral Image Classification: A 3-D Deep Learning Framework. *IEEE Transactions on Geoscience and Remote Sensing* 56, 2 (2018), 847–858. DOI: <https://doi.org/10.1109/TGRS.2017.2755542>
- [5] Hao Sun, Xiangtao Zheng, Xiaoqiang Lu, and Siyuan Wu. 2020. Spectral–Spatial Attention Network for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing* 58, 5 (2020), 3232–3245. DOI: <https://doi.org/10.1109/TGRS.2019.2951160>
- [6] Lichao Mou, Sudipan Saha, Yuansheng Hua, Francesca Bovolo, Lorenzo Bruzzone, and Xiao Xiang Zhu. 2021. Deep Reinforcement Learning for Band Selection in Hyperspectral Image Classification. *arXiv preprint arXiv:2103.08741* (2021). DOI: <https://doi.org/10.1109/TGRS.2021.3067096>