

# Deep Learning-Based Spectral Band Selection for Spectral Imaging Tasks

EMMANUEL MARTINEZ<sup>1</sup>, KEBIN CONTRERAS<sup>1</sup>, AND JORGE BACCA<sup>1,\*</sup>

<sup>1</sup>Universidad Industrial de Santander, Bucaramanga, Colombia

\*jbacquin@uis.edu.co

Compiled June 20, 2025

Spectral Images (SI) are acquired at multiple wavelengths across the electromagnetic spectrum, providing information that enhances performance in tasks such as material segmentation and classification by resolving ambiguities inherent in RGB images. SI devices are designed to capture a large number of spectral bands, which increases both cost and acquisition time, thereby limiting their practical deployment. However, not all spectral bands contribute equally to task-specific performance. To address this issue, a Deep Spectral Band Selection (DSBS) framework is proposed for spectral imaging tasks. Unlike previous methods that emphasize the preservation of non-task-specific information, DSBS identifies the most informative bands for a given task by jointly training a fully differentiable band selector and a neural network within an end-to-end learning framework. The selection process is guided by a proposed bin function and a custom  $\ell_p$ -norm regularization term to achieve the desired number of spectral bands. Experimental results in material segmentation and classification tasks indicate that DSBS outperforms state-of-the-art machine and deep learning methods.

<http://dx.doi.org/10.1364/ao.XX.XXXXXX>

## 1. INTRODUCTION

Spectral imaging is a technique for acquiring and processing spectral information from a scene, resulting in a three-dimensional data cube,  $\mathcal{F} \in \mathbb{R}^{M \times N \times L}$ , referred to as a Spectral Image (SI), where  $(M, N)$  denotes the spatial dimensions and  $L$  represents the number of spectral bands. The spectral content in an SI reveals the material composition within a scene, enabling a wide range of real-world applications, including food quality assessment [1], oil and mineral exploration [2], and precision agriculture [3]. For food analysis, for instance, SI enables non-destructive estimation of physicochemical properties, in contrast to conventional techniques that require sample destruction [4–6].

Recent advances in SI systems have focused on capturing hundreds or even thousands of narrow spectral bands to enhance spatial and spectral resolution, enabling identification of targets based on spectral variations [7, 8]. However, these systems often involve high acquisition costs and prolonged capture times, which may restrict their use in practical applications [9]. Furthermore, not all acquired bands are necessary for tasks such as classification or material segmentation only performance specific task. As an alternative, Band Selection (BS) methods aim to identify the most informative bands for a given task, supporting the design of customized and lower-cost spectral imaging systems [10, 11]. For example, in SI systems based on color filter arrays [12], by selecting a reduced number of bands or filters can lead to improved spatial resolution, while in systems em-

ploying band-by-band illumination [13], the sensing time can be decreased. Additionally, the cost of filter-based systems based on BS is generally lower than that of standard SI devices, which often rely on complex optical elements, making filter-based approaches more economically viable for SI acquisition.

BS methods leverage prior information such as correlation, covariance, similarity, low-rank structures, or entropy, to identify the most relevant spectral bands [14–16]. Alternative approaches formulate the BS problem using clustering, graph theory, or ranking strategies [17–23]. Recently, deep learning-based methods have introduced attention mechanisms, convolutional networks, reinforcement learning, and statistical learning to guide the BS process [24–26]. These techniques focus on selecting bands that best preserve information and are evaluated through classification tasks. However, most existing methods fail to account for the specific requirements of computational tasks such as material classification or segmentation, leading to suboptimal performance. Selecting bands without considering task-specific objectives can introduce redundancy and irrelevance, as the selected bands may not correspond to the discriminative features required for accurate modeling. Incorporating task-based criteria into the BS process is therefore essential for improving the effectiveness of spectral imaging tasks.

To address this limitation, a Deep Spectral Band Selection (DSBS) method is proposed, which integrates spectral BS with spectral imaging tasks within an End-To-End (E2E) framework.

DSBS introduces a fully differentiable BS optical encoder that identifies the most informative spectral bands through a discretization function and a custom  $\ell_p$  norm regularization, enabling the selection of an optimized number of bands tailored to the specific task. By embedding this BS encoder within spectral imaging networks, such as those used for classification and material segmentation, the E2E approach dynamically adapts BS to enhance task performance. This integration also reduces hardware complexity and acquisition time, offering a practical approach for applications such as food quality assessment, environmental monitoring, and remote sensing. A review of existing literature indicates that BS methods have not been previously assessed for material segmentation tasks. The present work provides the first evaluation of BS techniques in this context, extending their applicability to a broader range of spectral imaging applications.

## 2. RELATED WORK

BS methods in spectral imaging are commonly categorized into two main approaches: task-independent selection, where bands are chosen before the spectral imaging task, and task-based selection, where the BS process is optimized jointly with model training. Traditional methods follow a two-stage pipeline, in which bands are first selected and subsequently evaluated using classification models. In contrast, E2E approaches embed the BS process within model optimization, thereby aligning selected bands with task-specific objectives and simplifying the workflow.

### A. Task-independent Band Selection Methods

Tasks-independent BS methods perform selection as a standalone preprocessing step. Once the band subset is identified, the selected bands are evaluated in spectral imaging tasks such as classification, using established machine learning and deep learning algorithms such as Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Fully Connected Networks (FC), and CNN, among others [13]. In addition, the Correntropy-based Sparse Spectrum Clustering (CSSC) [27] uses a correntropy measure to create an affinity matrix, capture non-linear relationships, and select bands to reduce noise. Fast and Latent Low-Rank Subspace Clustering (FLLRSC) [28] clusters bands based on latent low-rank subspaces, identifying bands that capture essential spectral features. Laplacian Regularized Low-Rank Subspace Clustering (LLRSC) [29] combines low-rank representation and Laplacian regularization to preserve global and local correlations. The Optimal Clustering Framework [19] and the Adaptive Subspace Partition Strategy (ASPS) [30] use dynamic programming and noise-based subspace partitioning, respectively, to minimize redundancy while retaining spectral information. The Dropout Concrete Autoencoder (Dropout CAE) [26] uses a concrete selector layer within an autoencoder to classify and retain only the essential bands for reconstruction. In contrast, BS-Nets [24] use a Band Attention Module (BAM) to weight bands for reconstruction, but its effectiveness is highly dataset-dependent, limiting its broader applicability.

### B. Task-based Band Selection Methods

To address the limitations of task-independent strategies, recent research has focused on integrated approaches that jointly optimize BS and spectral imaging tasks within a unified E2E framework. This integration enables the identification of task-relevant bands, improving model performance while also reduc-

ing the complexity and cost of SI system design. For instance, Aya et al. [11] proposed an E2E framework incorporating a Bernoulli distributed binarization vector. This vector facilitates stochastic BS through a learnable mask, which is optimized via backpropagation to minimize classification loss under constraints on the number of selected bands. Recently, Karen et al. [10] introduced binary and ranking-based regularization terms to guide the selection process and ensure an optimal subset of bands, thereby enhancing classification accuracy. Building on these developments, the present work formulates a method that employs discrete functions for a realistic BS process and generalizes the ranking-based regularization using an  $\ell_p$  norm formulation. This advancement provides selection mechanism that extends beyond classification, enabling its application to tasks such as material segmentation and object detection. Nevertheless, most existing BS methods continue to be evaluated primarily in classification contexts, which limits their broader applicability. Expanding BS techniques to additional tasks has the potential to enhance performance and reduce acquisition costs across various domains, including environmental monitoring, food quality assessment, and remote sensing [31, 32].

## 3. BAND SELECTION OPTICAL ENCODER

An SI can be acquired using the filter-based method, which involves an optical system composed of two lenses, a filter wheel with multiple spectral responses, and a 2-Dimensional (2D) sensor. Light from the scene is first captured by an objective lens, which collects and directs it toward the filter wheel. Each filter, characterised by a specific spectral transfer function, attenuates wavelengths outside its designated range, isolating a narrow-band portion of the spectrum. A second lens then refocuses the filtered light onto the 2D sensor, generating a monochromatic image corresponding to a single spectral band of the scene.

Mathematically, the scene can be represented by the spectral radiance function  $f(x, y, \lambda)$ , where  $(x, y)$  denotes the spatial coordinates and  $\lambda$  denotes the wavelength. The spectral transfer function of a given filter can be modeled using the Johnson's SU distribution [33] as

$$\mathcal{H}(\lambda|\lambda_l, \sigma, \beta, \delta) = \frac{\delta}{\sigma\sqrt{2\pi}} \cdot \frac{1}{\sqrt{1 + \left(\frac{\lambda - \lambda_l}{\sigma}\right)^2}} \times \exp\left(-\frac{1}{2} \left(\delta \sinh^{-1}\left(\frac{\lambda - \lambda_l}{\sigma}\right) - \beta\right)^2\right), \quad (1)$$

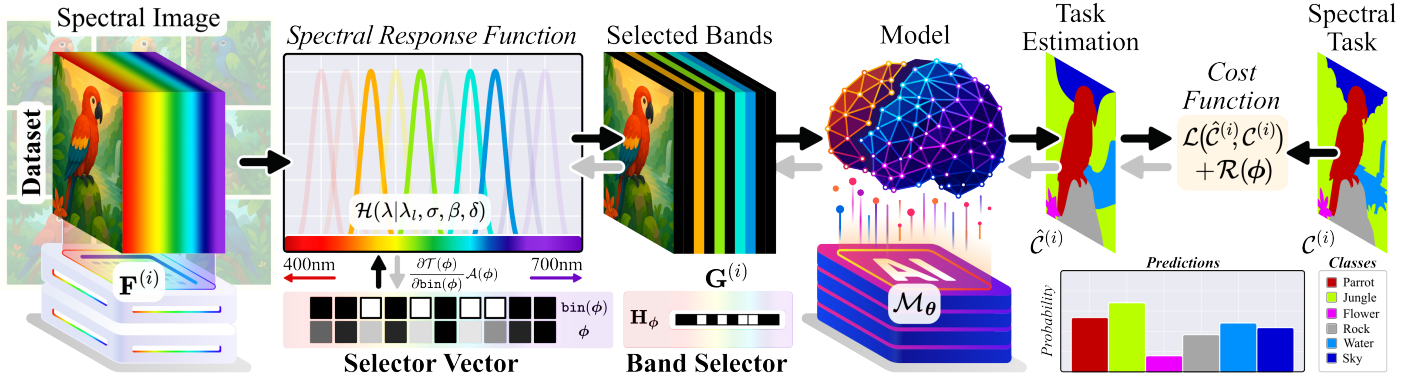
where  $\lambda_l$  is the central wavelength,  $\sigma$  is the bandwidth,  $\beta$  controls the asymmetry, and  $\delta$  governs the spectral energy concentration. The intensity measured by the sensor at the corresponding filter position is given by

$$\tilde{f}_l(x, y) = \int \mathcal{H}(\lambda|\lambda_l, \sigma, \beta, \delta) \cdot f(x, y, \lambda) d\lambda, \quad (2)$$

where  $l$  indexes the spectral band associated with the central wavelength  $\lambda_l$ . Assuming a discretized formulation of Eq. (2), the observed image at band  $l$  can be expressed as

$$\tilde{\mathbf{f}}_l = \mathbf{h}_l^T \mathbf{F}, \quad (3)$$

where  $\mathbf{h} \in \mathbb{R}^{L \times 1}$  is the discrete transfer function for the  $l$ -th filter, and  $\mathbf{F} \in \mathbb{R}^{L \times MN}$  is the discretized SI matrix, with spatial resolution  $(M, N)$  and  $L$  spectral bands. As the filter wheel



**Fig. 1. Overview of the proposed DSBS framework.** A spectral image  $\mathbf{F}^{(i)}$  is processed by a band selection encoder parameterized by a binary selector  $\phi$ . Selected bands  $\mathbf{G}^{(i)}$  are modeled via  $\mathbf{H}_\phi$  and used as input to a model  $\mathcal{M}_\theta$  to estimate the task output  $\hat{\mathcal{C}}^{(i)}$ . The model is trained end-to-end by minimizing a task-specific loss  $\mathcal{L}$  regularized by  $\mathcal{R}(\phi)$ .

rotates to acquire each spectral band sequentially, the full SI can be described mathematically as

$$\tilde{\mathbf{F}} = \mathbf{H}\mathbf{F}, \quad (4)$$

where  $\mathbf{H} = [\mathbf{h}_1, \dots, \mathbf{h}_L]^T$  represents the stacked spectral transfer functions for the entire filter set. However, as the number of filters  $L$  increases, the system becomes more complex and acquisition times longer. To mitigate this, the BS process aims to identify a reduced subset of  $B \ll L$  filters for acquisition. Mathematically, this process is equivalent to selecting a subset of rows from  $\mathbf{H}$ . Consequently, by defining a selector vector  $\phi \in \{0, 1\}^L$ , with  $\|\phi\|_1 = B$ , the BS can be modeled as

$$\mathbf{G} = \text{diag}(\phi)\mathbf{H}\mathbf{F} := \mathbf{H}_\phi(\mathbf{F}), \quad (5)$$

where  $\mathbf{G} \in \mathbb{R}^{L \times MN}$  denotes the reduced SI composed of the  $B$  selected bands, with the remaining bands set to zero, and  $\mathbf{H}_\phi$  represents the BS optical encoder parameterized by the binary vector  $\phi$ . In this context, inference is performed using  $\mathbf{G}$  instead of the full spectrum  $\mathbf{F}$ , while maintaining the task performance by optimally learning the binary selector vector  $\phi$ .

#### 4. END-TO-END SPECTRAL BAND SELECTION FOR SPECTRAL IMAGE TASKS

The proposed method performs a joint optimization of the BS encoder  $\mathbf{H}_\phi(\cdot)$  and a Deep Neural Network (DNN)  $\mathcal{M}_\theta(\cdot)$  with learnable parameters  $\theta$ , to achieve task-adaptive spectral band selection within an end-to-end learning framework. Given a dataset  $\mathcal{D} = \{\mathbf{F}^{(i)}, \mathcal{C}^{(i)}\}_{i=1}^K$  consisting of  $K$  paired samples, where each  $\mathbf{F}^{(i)}$  represents a SI and  $\mathcal{C}^{(i)} \subset \Omega$  denotes the corresponding ground truth labels within the output space  $\Omega$  associated with the SI task, the joint optimization problem can be formulated as

$$(\phi^*, \theta^*) \in \arg \min_{\phi, \theta} \mathbb{E}_{\mathbf{F}, \mathcal{C} \sim \mathcal{D}} [\mathcal{L}(\mathcal{M}_\theta(\mathbf{H}_\phi(\mathbf{F})), \mathcal{C}) + \mathcal{R}(\phi)], \quad (6)$$

where  $\phi^*$  and  $\theta^*$  are the optimized parameters for the band selector and DNN, respectively. The function  $\mathcal{L}$  corresponds to the task-specific loss (See following subsections), and  $\mathcal{R}(\phi)$  is a regularization term that enforces convergence toward a pre-defined number of selected bands. At initialization, the selector vector is set to ones, i.e.,  $\phi \in \{1\}^L$ . During training, continuous values in  $\mathbb{R}$  are allowed for gradient-based optimization. However, a binarization function is subsequently applied to enforce a

discrete selection in  $\{0, 1\}^L$ , ensuring that the intensity structure of spectral signatures is preserved in the reduced representation. The binarization function is then formulated as

$$\text{bin}(\phi)_l = \begin{cases} 1 & \text{if } \phi_l \geq 0 \\ 0 & \text{if } \phi_l < 0 \end{cases}, \quad (7)$$

where  $\text{bin}(\cdot)$  is non-differentiable and thus incompatible with gradient-based optimization. To address this, a custom Straight-Through Estimator (STE) [34] is designed using a surrogate function  $\mathcal{A}(\phi)$ , which approximates the gradient as

$$\frac{\partial \mathcal{T}(\phi)}{\partial \phi} = \frac{\partial \mathcal{T}(\phi)}{\partial \text{bin}(\phi)} \frac{\partial \text{bin}(\phi)}{\partial \phi} \approx \frac{\partial \mathcal{T}(\phi)}{\partial \text{bin}(\phi)} \mathcal{A}(\phi), \quad (8)$$

where  $\mathcal{T} \in \{\mathcal{L}, \mathcal{R}\}$  is associated with Eq. (6) and denotes any loss function dependent on  $\phi$ . The surrogate function  $\mathcal{A}(\cdot)$  is defined element-wise as a customized variant of the  $\text{hardtanh}$  function, designed to smooth the gradients of the band selector and enable stable optimization. Specifically, it is formulated as

$$\mathcal{A}(\phi)_l = \begin{cases} 1 & \text{if } \phi_l \geq 0.5 \\ \phi_l & \text{if } 0.5 > \phi_l > -0.5 \\ 0 & \text{if } \phi_l \leq -0.5 \end{cases}, \quad (9)$$

where the gradient values are constrained to the range  $\{0, 1\}$  to remain consistent with the behavior of the binarization function  $\text{bin}(\cdot)$ . Finally, the regularization function  $\mathcal{R}(\cdot)$  imposes a constraint on the number of selected spectral bands  $B$ , generalizing the formulation presented in [10]. It is defined as

$$\mathcal{R}(\phi) = \rho |B - \|\text{bin}(\phi)\|_1|^p, \quad (10)$$

where  $\rho$  is a regularization parameter,  $\|\text{bin}(\phi)\|_1$  represents the number of bands selected the binarization vector  $\phi$ , and  $p$  controls the penalty strength. The subsequent subsections describe the application of this framework to spectral imaging tasks such as classification and material segmentation, along with their respective loss functions  $\mathcal{L}$ .

##### A. Classification

For SI classification, the fidelity term in Eq. (6) is defined as the cross-entropy loss

$$\mathcal{L}(\mathcal{M}_\theta(\mathbf{H}_\phi(\mathbf{F})), c) = - \sum_{j=1}^C \mathbb{1}_{c=j} \log([\mathcal{M}_\theta(\mathbf{H}_\phi(\mathbf{F}))]_j), \quad (11)$$



where  $\{\mathbf{F}, c\}$  denotes a sample-label pair,  $c^{(i)} \in \Omega = \{0, \dots, C - 1\}$  is the ground-truth class label, and  $C$  is the maximum number of classes. The indicator function  $\mathbb{1}_v$  equals 1 if  $c = j$  and 0 otherwise, and  $[\mathcal{M}_\theta(\cdot)]_j$  represents the  $j^{\text{th}}$  element of the DNN output.

## B. Material Segmentation

For material segmentation, each spectral signature in the SI  $\mathbf{F}$  is assigned a material class, resulting in a label vector  $\mathbf{c} \in \mathbb{R}^{MN}$ . In this task, the predicted segmentation map is computed as

$$\hat{\mathbf{c}} = \mathcal{M}_\theta(\mathbf{H}_\phi(\mathbf{F})) \in \mathbb{R}^{MN}, \quad (12)$$

where  $\hat{\mathbf{C}}$  denotes the estimated class scores for each spatial location. Following the approach in [31], the fidelity term in Equation Eq. (6) is defined using the focal loss function

$$\mathcal{L}(\mathcal{M}_\theta(\mathbf{H}_\phi(\mathbf{F})), \mathbf{c}) = - \sum_{j=1}^C \mathbb{1}_{c=j} \alpha (1 - e^{\hat{c}_j})^\gamma \log(\hat{c}_j), \quad (13)$$

where  $\alpha \in \mathbb{R}$  is a weighting factor, and  $\gamma \in \mathbb{R}$  is the focusing parameter, both of which control the contribution of well-classified versus hard-to-classify samples during training.

## 5. SIMULATIONS AND RESULTS

This section presents the experimental validation of the proposed method DSBS, which is designed to identify optimal spectral bands for spectral imaging tasks, improving performance while reducing hardware and operational costs. The experimental setup includes a comparative analysis of DSBS against state-of-the-art machine and deep learning based algorithms in classification and material segmentation tasks. All experiments were conducted on a single NVIDIA RTX 3090 GPU and repeated five times, which reported results representing the average performance across these runs. The code is available at <https://github.com/Enmartz/DSBS>.

### A. Classification Task

**Datasets.** The evaluated datasets are Indian Pines, Pavia University, and Salinas. These datasets differ in spatial resolution, spectral range, and number of land cover classes, making them standard benchmarks for assessing SI classification algorithms, as summarized in Tab. 1. Indian Pines and Salinas, both acquired by the AVIRIS sensor, provide broad spectral coverage from 400 to 2500 nm, and are commonly used in agricultural and environmental monitoring tasks. In contrast, the Pavia University dataset, acquired by the ROSIS sensor, focuses on urban land cover analysis with a narrower spectral range of 430 to 860 nm. The number of land cover classes also varies, with Indian Pines and Salinas containing 16 classes each, whereas Pavia University includes 9 classes. Additionally, all the datasets were partitioned into 5% of the spectral signatures for training, 45% for validation and 50% for testing. These variations present diverse challenges for SI classification, requiring effective feature extraction and task-specific learning strategies.

**Models.** A Convolutional Neural Network (CNN), inspired in [10], is employed as the classification model. This architecture consists of 6 one-dimensional convolutional layers followed by 4 fully connected (FC) layers, each with ReLU activation functions. Each input corresponds to a spectral signature  $\mathbf{f} \in \mathbb{R}^L$ , representing a column for each SI  $\mathbf{F}$ . The model outputs a one-hot encoded label vector  $\mathbf{c}$ , performing classification according

**Table 1.** Overview of the spectral image datasets used for land cover classification.

Dataset	Sensor	Location	Size	Bands	Range	Classes
Indian Pines	AVIRIS	USA	145 × 145	200	400–2500	16
Pavia Univ.	ROSIS	Italy	610 × 340	103	430–860	9
Salinas	AVIRIS	USA	512 × 217	204	400–2500	16

to Eq. (11). The training configuration includes a learning rate of  $10^{-3}$ , a batch size of 32, and 8000 epochs, enabling convergence of both the BS optical encoder until the desired number of bands and classification model. Additionally, an  $\ell_2$ -norm regularization with a weight of  $10^{-4}$  is applied to the FC layers. A preprocessing step is applied to each dataset, consisting of global min-max normalization for each spectral band.

After training the DSBS framework and identifying the selected bands, an additional evaluation is conducted using only the selected bands in the form  $\mathbf{G}$ . This evaluation is carried out using four SOTA ML and DL models: Support Vector Machine (SVM), K-Nearest Neighbors (KNN), FC, and CNN (with the same architecture using during E2E DSBS training) [35]. All models are trained from scratch on the selected bands to assess their classification performance and to evaluate the generalization capacity of the BS.

**Metrics.** Performance is evaluated using Overall Accuracy (OA), Average Accuracy (AA), and the Kappa coefficient. Given a confusion matrix  $\mathbf{N}^{C \times C}$ , where each entry  $n_{ij}$  denotes the number of samples whose ground-truth class is  $i$  and that were predicted as class  $j$ , OA quantifies the proportion of correctly classified samples over the total number of samples, providing a global measure of classification accuracy as

$$OA = \frac{1}{N} \sum_{i=1}^C n_{ii}, \quad \text{with } N = \sum_{i=1}^C \sum_{j=1}^C n_{ij} \quad (14)$$

AA computes the average of per-class accuracies, offering a balanced view of model performance across all classes, particularly in the presence of imbalanced datasets as

$$AA = \frac{1}{C} \sum_{i=1}^C \frac{n_{ii}}{n_i}, \quad \text{where } n_i = \sum_{j=1}^C n_{ij} \quad (15)$$

where  $n_i$  is the total number of samples with in class  $i$ . Lastly, the Kappa coefficient ( $\kappa$ ) adjusts the observed agreement (OA) for the agreement expected by chance (PE), providing a more robust measure of prediction reliability as

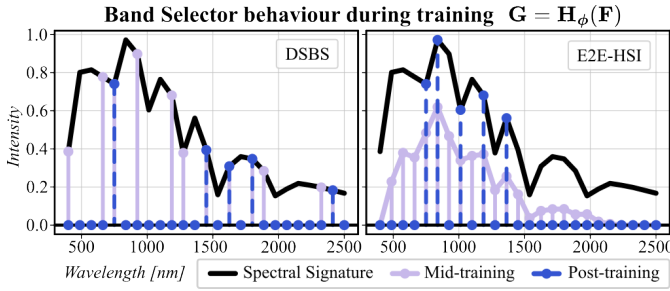
$$\kappa = \frac{OA - PE}{1 - PE}, \quad \text{with } PE = \frac{1}{N^2} \sum_{i=1}^C n_i \cdot n_i^p \quad (16)$$

where  $n_i^p = \sum_{j=1}^C n_{ji}$  is the number samples predicted as class  $i$ , and  $N$  is the total number of samples.

### A.1. Ablation Study of the Band Selector

This experiment evaluates the impact of the two main components of the proposed DSBS framework: the BS optical encoder incorporating the  $\text{bin}(\phi)$  function, and the general regularization term  $\mathcal{R}(\phi)$  that controls the number of selected bands. To quantify the improvements, DSBS is compared against the E2E-HSI method [10] using OA, AA, and  $\kappa$  classification metrics.

Table 2 presents the results on Indian Pines dataset using 10 selected bands. The evaluation includes a variation on the regularization exponent  $p \in \{0.5, 1, 2\}$ , associated with the regularization term. DSBS consistently outperforms E2E-HSI across all metrics, achieving an improvement of 6 – 8%. The main



**Fig. 2. Evolution of the Band Selector.** Behavior mid-training (purple) and post-training (blue), visualized using a spectral signature (black), for E2E-HSI [10] and DSBS. The spectral signature was scaled for visualization. During training, global min-max normalization and standardization were applied to the full dataset for DSBS and E2E-HSI, respectively.

advantage is attributed to the application of the binarization function  $\text{bin}$ , which is not present in E2E-HSI. Moreover, tuning the regularization strength to  $\rho = 0.5$  results in an additional improvement of 0.25 – 0.5% compared to other values. Consequently, the value is fixed for all subsequent experiments.

To further illustrate the benefits of the proposed method, Fig. 2 visualises input of the classification model  $\mathbf{G} = \mathbf{H}_\phi(\mathbf{F})$ , over the course of training using a representative spectral signature  $\mathbf{F}$ . Compared to E2E-HSI, the proposed approach preserves both the structure and intensity of the spectral signature, while E2E-HSI introduces notable distortions during training. This improvement is primarily attributed to the inclusion of the binarization function, as corroborated by the quantitative results in Tab. 2. These findings highlight the superiority of the proposed DSBS approach over E2E-HSI.

### A.2. Comparison with Task-Independent and Task-Based Band Selection Methods

To evaluate the performance of the proposed DSBS framework, it is compared against several State-Of-The-Art (SOTA) BS methods. These include deep learning-based approaches, such as BS-Nets-FC [24], BS-Net-Conv [24], Dropout CAE [26], and E2E-HSI [10], as well as traditional methods such as RL [25], RDFBSS [37], MLBS [11], NC-OC-IE [19], NC-OC-MVPCA [19], and TSC-OC-FDPC [38]. The evaluation is performed on the 3 selected datasets using 10 selected spectral bands and four classifiers: CNN, FC, KNN, and SVM. Table 3 reports the classification performance of the proposed DSBS relative to these methods.

In the Indian Pines dataset, DSBS achieves the highest performance in all three metrics (OA, AA, and  $\kappa$ ) for most classifiers.

**Table 2. Classification results on the Indian Pines dataset using 10 selected bands under different ablation settings.** The effects of the binarization function ( $\text{bin}$ ) and regularization exponent  $p$  are evaluated. Best results are in **bold green**, and second-best in blue underline

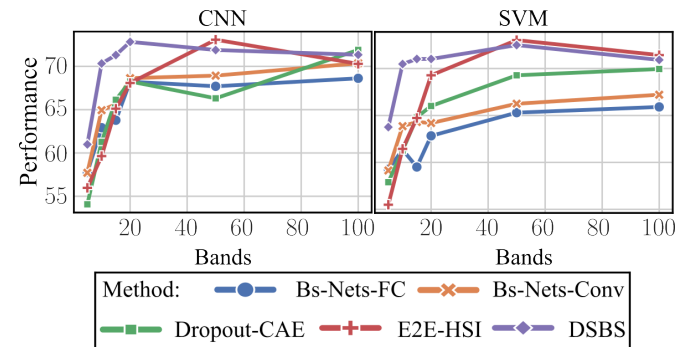
Method	Ablation		OA ( $\uparrow$ )	AA ( $\uparrow$ )	$\kappa \times 100$ ( $\uparrow$ )
	Bin	Reg			
E2E-HSI [36]	$\times$	2	65.39	53.20	60.33
	$\checkmark$	2	72.56	65.19	68.70
DSBS	$\checkmark$	1	<u>73.96</u>	<u>66.12</u>	<u>70.28</u>
	$\checkmark$	0.5	<b>74.18</b>	<b>66.23</b>	<b>70.58</b>

For example, DSBS reaches 74.18% OA with CNN and 73.43% with SVM, outperforming the second-best E2E-HSI and NC-OC-IE. In terms of  $\kappa$ , DSBS also leads with values of 70.58 CNN, 68.84 FC, and 69.72 SVM. These results demonstrate the robustness of the model under limited band configurations. For Pavia University, DSBS again outperforms all baselines in most scenarios. It obtains the highest OA in every classifier: 90.51% (CNN), 90.42% FC, 86.47% KNN, and 90.66% (SVM). Furthermore, DSBS achieves the best AA in machine learning methods with 83.80% KNN and 88.25% SVM. The  $\kappa$  scores are also highest for CNN and FC, further supporting the method's consistent performance across classification backbones. In the Salinas dataset, DSBS maintains top performance across most of the metrics and classifiers. It achieves 90.79 – 91.74% OA, 93.21 – 94.75 AA, and  $\kappa \times 100$  values up to 90.79, surpassing strong baselines such as RDFBSS and NC-OC-IE.

Figure 3 shows the classification maps generated for each dataset. The results obtained with DSBS exhibit improved spatial consistency and fewer classification artefacts compared to conventional BS methods. These visual results align with the observed improvements in OA, AA, and  $\kappa \times 100$ , demonstrating the model's capacity to preserve relevant spectral-spatial structures using a reduced number of bands.

### A.3. Number of selected bands

To analyse the impact of the number of selected bands on classification accuracy, the Indian Pines dataset is used. The evaluation considers deep learning-based BS methods, including BS-Nets-FC [24], BS-Nets-Conv [24], Dropout CAE [26], E2E-HSI [10], and the proposed DSBS. These methods are assessed using two classifiers (CNN and SVM) with varying band configurations  $B = \{5, 10, 15, 20, 50, 100\}$ . Figure 4 shows the OA results obtained on the Indian Pines dataset. DSBS consistently outperforms all other methods across all selected bands, with the most significant improvements observed in low-band regimes  $B < 20$ , demonstrating that our method provides superior performance, especially when a reduced number of spectral bands is available.



**Fig. 4. OA performance of band selection methods on the Indian Pines dataset using varying numbers of selected bands.** Results are shown for CNN and SVM classifiers with  $B \in \{5, 10, 15, 20, 50, 100\}$ .

## B. Material Segmentation Task

To extend the applicability of the proposed DSBS framework beyond classification, the method was evaluated on a material segmentation task. In this context, the goal is to segment objects such as concrete, vegetation, metal, and glass using a reduced number of spectral bands.

**Table 3. Classification performance using 10 selected bands across classifiers and datasets.** Results are reported for OA, AA, and  $\kappa \times 100$  using CNN, FC, KNN, and SVM. Best results are in **bold green**, and second-best in blue underline

Dataset	Method	OA ( $\uparrow$ )				AA ( $\uparrow$ )				$\kappa \times 100$ ( $\uparrow$ )			
		CNN	FC	KNN	SVM	CNN	FC	KNN	SVM	CNN	FC	KNN	SVM
Indian Pines	TRC-OC-FDPC [19]	66.00	70.01	62.63	<u>71.50</u>	57.59	<u>64.85</u>	51.06	66.73	61.16	65.88	57.18	<u>67.55</u>
	RDFBSS [37]	62.32	64.73	57.59	62.52	54.44	59.99	47.50	59.39	56.80	59.71	51.22	57.25
	NC-OC-IE [19]	70.38	<u>70.96</u>	<u>64.90</u>	70.62	64.99	<b>65.35</b>	<u>52.36</u>	<u>67.75</u>	66.12	<u>66.82</u>	<u>59.69</u>	66.55
	NC-OC-MVPCA [19]	<u>71.83</u>	70.16	64.40	69.62	<u>65.08</u>	63.57	51.89	67.03	<u>67.83</u>	65.83	59.11	65.43
	BS-Nets-FC [24]	66.65	64.70	60.04	64.64	60.28	56.67	47.21	59.75	61.82	59.58	54.06	59.67
	BS-Nets-Conv [24]	70.24	67.97	55.03	66.69	60.54	58.71	44.37	58.09	66.13	63.46	48.42	62.10
	Dropout CAE [26]	64.88	65.46	58.70	64.41	59.07	59.23	47.28	59.89	59.88	60.47	52.59	59.49
	E2E-HSI [10]	65.39	66.29	60.38	65.45	53.20	54.92	46.74	58.27	60.33	61.33	54.60	60.58
	DSBS	<b>74.18</b>	<b>72.80</b>	<b>66.40</b>	<b>73.43</b>	<b>66.23</b>	64.46	<b>53.38</b>	<b>68.30</b>	<b>70.58</b>	<b>68.84</b>	<b>61.46</b>	<b>69.72</b>
Pavia University	TRC-OC-FDPC [19]	79.86	81.46	78.72	82.33	78.20	78.99	75.17	77.92	73.12	75.08	71.11	75.71
	NC-OC-MVPCA [19]	80.85	81.52	79.14	82.55	79.32	79.76	75.62	78.06	74.39	75.30	71.68	76.01
	NC-OC-IE [19]	79.76	81.21	78.98	82.44	77.78	78.75	74.28	79.26	79.25	74.67	71.47	75.84
	RDFBSS [37]	85.25	85.88	79.42	87.35	82.20	83.36	74.22	84.11	80.34	81.23	72.09	83.08
	BS-Nets-FC [24]	84.77	86.46	83.05	87.77	82.45	84.99	79.39	84.63	79.73	82.09	77.08	83.51
	BS-Nets-Conv [24]	82.91	83.29	79.91	85.46	80.07	80.45	74.73	81.59	77.18	77.65	72.72	80.38
	Dropout CAE [26]	84.45	86.12	83.34	88.07	82.12	83.33	79.45	84.55	79.66	81.67	77.43	83.98
	E2E-HSI [10]	<u>89.76</u>	<u>90.16</u>	<u>86.11</u>	<u>89.95</u>	<b>88.19</b>	<b>89.03</b>	<u>83.33</u>	<u>87.24</u>	<u>86.33</u>	<u>86.98</u>	<u>81.22</u>	<u>86.53</u>
	DSBS	<b>90.51</b>	<b>90.42</b>	<b>86.47</b>	<b>90.66</b>	<u>87.95</u>	<u>88.02</u>	<b>83.80</b>	<b>88.25</b>	<b>87.35</b>	<b>87.22</b>	<b>81.71</b>	<b>87.54</b>
Salinas	TRC-OC-FDPC [19]	88.83	89.07	87.21	90.66	92.76	92.70	92.26	94.29	87.52	87.81	85.76	89.57
	NC-OC-MVPCA [19]	88.89	89.54	87.72	<u>91.65</u>	93.50	93.08	92.76	<u>95.27</u>	81.63	88.33	86.32	<u>90.69</u>
	NC-OC-IE [19]	<u>89.00</u>	<u>89.85</u>	87.88	91.59	93.01	93.91	92.75	95.17	<u>87.74</u>	<u>88.69</u>	86.28	90.62
	RDFBSS [37]	88.64	<b>90.35</b>	<u>88.22</u>	90.99	<u>93.68</u>	<b>94.12</b>	<u>93.23</u>	94.63	87.36	<b>89.23</b>	<u>86.88</u>	87.95
	BS-Nets-FC [24]	88.85	89.30	87.57	91.00	93.46	93.01	92.81	95.09	87.60	87.50	86.16	89.96
	BS-Nets-Conv [24]	88.06	88.94	86.19	90.34	92.42	92.90	91.26	94.34	86.66	87.67	84.61	89.23
	Dropout CAE [26]	88.25	89.19	87.03	90.59	93.19	93.44	92.58	94.77	86.90	87.95	85.55	89.50
	E2E-HSI [10]	88.89	89.85	87.33	90.62	93.43	<u>94.02</u>	92.75	94.73	87.62	88.67	85.89	89.53
	DSBS	<b>90.79</b>	88.31	<b>88.62</b>	<b>91.74</b>	<b>94.75</b>	93.21	<b>93.50</b>	<b>95.49</b>	<b>89.72</b>	87.02	<b>87.30</b>	<b>90.79</b>

**Dataset.** The LIB-HSI dataset [31] was used for this task. It was captured using the Specim IQ handheld hyperspectral camera (Specim, Spectral Imaging Ltd.), covering the 400 to 1000 nm spectral range with a spatial resolution of  $512 \times 512$  pixels and 204 spectral bands. The dataset comprises 513 spectral images of building facades in a light industrial environment. Each pixel is labeled with one of 44 material classes. To minimize the effects of illumination variability, all images were acquired under shaded or overcast conditions. For computational efficiency, the images were downsampled to  $256 \times 256$  pixels while preserving the discrete class annotations. The dataset was divided into 393 paired samples for training, 45 for validation, and 75 for testing.

**Models.** Following the protocol described in [31], a Fully Convolutional Network (FCN) based on the ResNet-50 backbone was employed for segmentation. The input layer of the network was modified to match the specific number of spectral bands. The network processes full spectral images  $\mathbf{F} \in \mathbb{R}^{L \times MN}$  and produces a segmentation map  $\mathbf{c} \in \mathbb{R}^{C \times MN}$ , as defined in Eq. (13). To ensure consistency with the original training procedure, a reduce-on-plateau learning rate scheduler was used, with a patience threshold of 15 epochs. The training setup included an initial learning rate of  $10^{-3}$  and a batch size of 8. Finally, the

parameter to control the number of selected bands is  $\rho = 0.05$ .

**Metrics.** Material segmentation was evaluated using four metrics: Focal Loss, Cross-entropy (CE), Intersection over Union (IoU), and the Dice Coefficient. These metrics capture different aspects of segmentation quality, including class imbalance and region-level overlap between predicted and ground-truth labels.

IoU quantifies the ratio between the predicted and ground-truth regions as

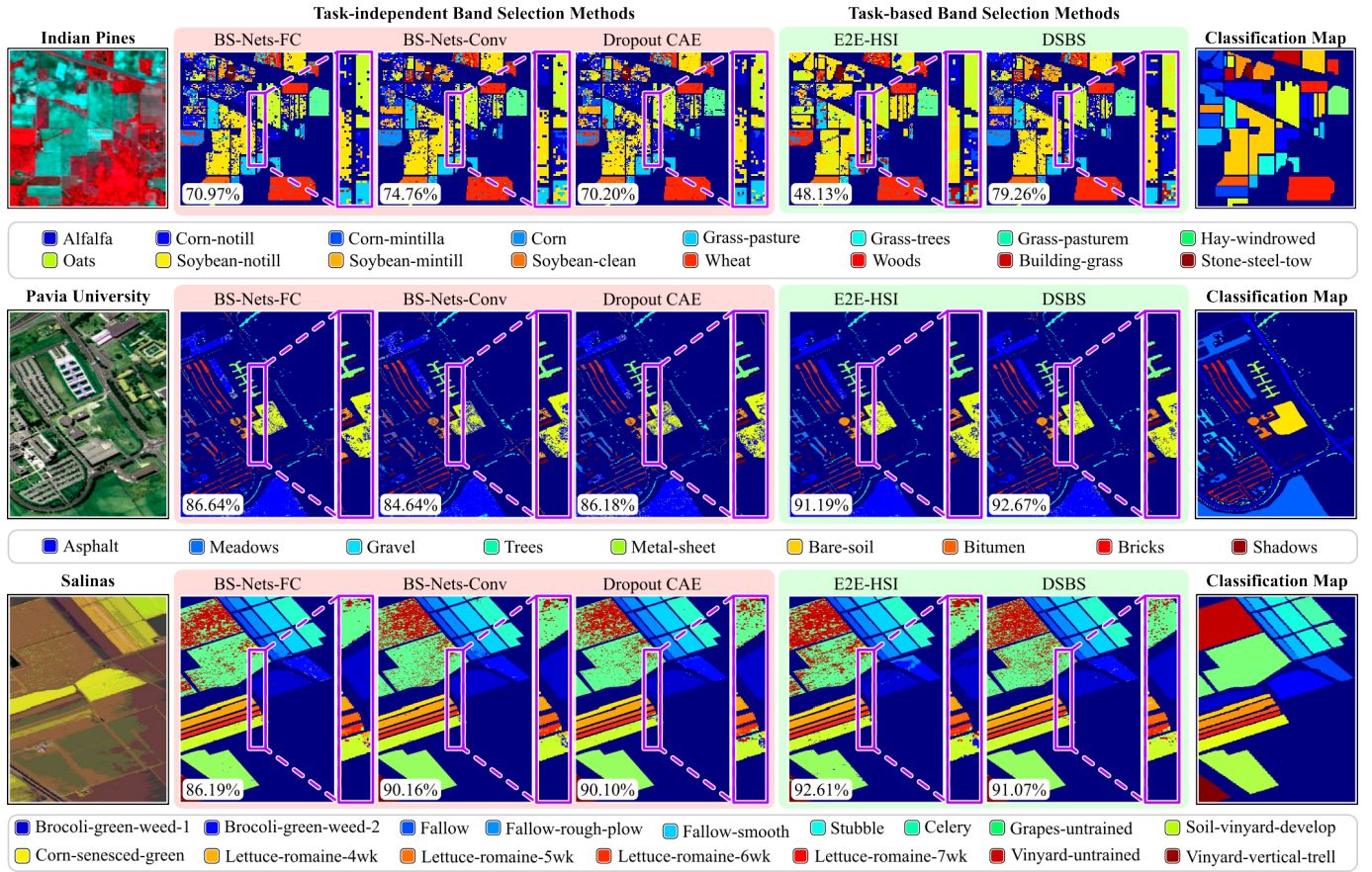
$$\text{IoU} = \frac{|P \cap G|}{|P \cup G|}, \quad (17)$$

where  $P$  is the set of pixels predicted as belonging to a class, and  $G$  is the corresponding ground-truth set. The Dice Coefficient, closely related to IoU, measures the overlap between prediction and ground-truth. It is defined as

$$\text{Dice Coefficient} = \frac{2 \times |P \cap G|}{|P| + |G|}, \quad (18)$$

where  $|P|$  and  $|G|$  are the total number of predicted and ground-truth pixels for the target class, respectively. The Dice score is particularly relevant in scenarios with imbalanced class distributions.





**Fig. 3. Visual Classification Results with OA metric on the Indian Pines, Pavia University, and Salinas datasets.** From left to right: false-color RGB image, classification maps from task-independent methods (BS-Nets-FC, BS-Nets-Conv, Dropout CAE), task-based methods (E2E-HSI, DSBS), and the ground-truth map. Estimated maps correspond to CNN-based classification. Each reported result corresponds to the best OA performance for each method across 5 runs.

### B.1. Segmentation Performance Varying Number of Bands.

The DSBS framework is evaluated in the material segmentation setting, where it can adapt to different tasks by integrating the BS optical encoder. As no prior BS methods have addressed segmentation, comparisons are made against the baseline from [31], which uses full-spectrum bands with 204 bands and RGB images with 3 channels. Additional experiments are conducted using the proposed method with  $B = \{1, 3, 10\}$  spectral bands.

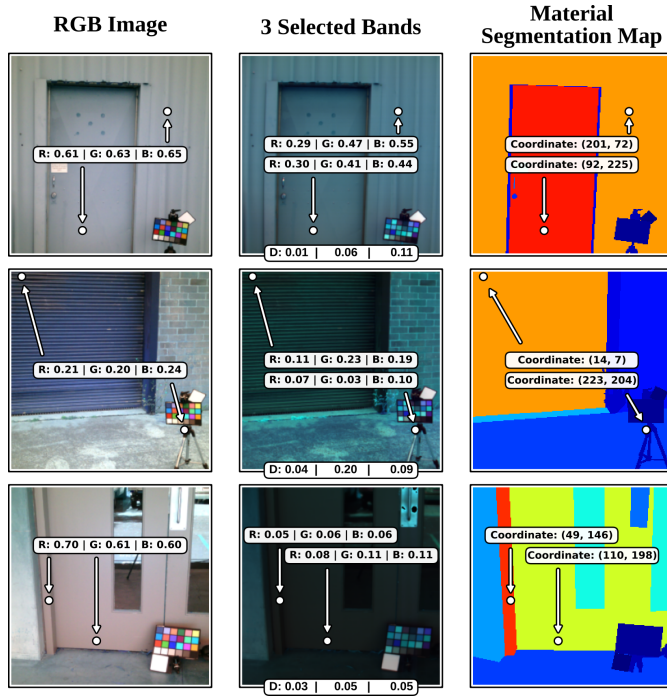
As shown in Table 4, the DSBS method was evaluated with

**Table 4. Performance of the band selector optical encoder on material segmentation using different numbers of selected bands.** Metrics include Focal loss, Cross-Entropy (CE), Dice score, and Intersection over Union (IoU). The selector was configured with  $\phi \in \{0, 1\}^L$  and  $\rho = 0.05$ . Best results are in **green bold**, and second-best results in **blue underline**.

Method	Bands	Focal ( $\downarrow$ )	CE ( $\downarrow$ )	Dice ( $\uparrow$ )	IoU ( $\uparrow$ )
RGB	3	0.1932	0.7482	73.57	58.66
DSBS	1	0.2405	0.8374	73.81	58.88
	3	0.1688	0.7231	75.51	60.95
	5	<u>0.1560</u>	<u>0.6976</u>	<u>76.02</u>	<b>61.66</b>
	10	<b>0.1542</b>	<b>0.6926</b>	<b>76.04</b>	<u>61.64</u>
HSI	204	0.1638	0.7126	75.14	60.58

selected spectral bands 10, 5, 3 and 1. Using 10 bands, the segmentation performance approaches that of the full spectrum baseline (204 bands), with an IoU of 61.64% compared to 60.58%, a difference of only 1.06%. This result indicates that the selection of a small number of bands for the segmentation task can be compared to the input of the full spectrum. With 5 bands, the IoU reaches 61.66%. Therefore, these findings demonstrate that the DSBS method can preserve segmentation quality while reducing spectral dimensionality with a DICE of 76.04%, which is beneficial for systems with limited computational resources.

When reducing the number of selected bands to 3, the DSBS method achieves a Dice score of 75.51% and an IoU of 60.95%, both of which exceed the RGB baseline (Dice = 73.57%, IoU = 58.66%). The improvement in IoU is approximately 2.29%, and in Dice 1.94%. Therefore, learnt band selection provides better feature representation than fixed RGB channels. The three selected bands correspond to central wavelengths of 400 nm, 747 nm, and 918 nm. To provide a visual representation of these new bands, Figure 5 presents both the original RGB image and a false RGB image constructed using the newly selected bands, highlighting two distinct points. As shown, two pixels in the original RGB image, which share the same RGB values but belong to different classes, are clearly separated in the new image. This demonstrates a significant visual improvement achieved by the proposed method. In the case of using only one band, DSBS achieves an IoU of 58.88%, nearly matching the RGB



**Fig. 5. Comparison of RGB images, DSBS selected spectral bands, and material segmentation maps for three scenes.** Each row shows one scene with sampled points. Each row shows one scene with sampled points. R, G, B indicate red, green, and blue channel intensities; D represents the spectral difference between paired points. RGB values (left), falsr RGB images constructed using the three selected bands (center), and material segmentation map with coordinates (right).

baseline with a deviation of just 0.22%. This result confirms that even minimal spectral content, when optimally selected, can encode task-relevant information.

### B.2. Adapted Band Selection Methods for Material Segmentation

This section evaluates the SOTA band selection methods using 10 selected bands, such as BS-Nets-FC [24], BS-Net-Conv [24], Dropout CAE [26], and E2E-HSI [10] on the material segmentation task for 10 spectral bands. As these methods were originally developed for classification, they were adapted, and the selected bands were used for training the segmentation model from scratch. Unique exception is the E2E-HSI method, which

**Table 5. Classification performance across various classification methods using 10 selected bands, evaluated across different band selection methods for Material Segmentation.** The best results are highlighted in **green bold**, and the second-best results are in blue underline.

Method	Focal ( $\downarrow$ )	CE ( $\downarrow$ )	Dice ( $\uparrow$ )	IoU ( $\uparrow$ )
BS-Nets-FC	0.1646	0.7115	74.27	59.47
BS-Nets-Conv	0.1913	0.7672	73.63	58.58
Dropout CAE	0.1611	0.7048	75.37	60.84
E2E-HSI	<u>0.1555</u>	<u>0.6986</u>	<u>75.74</u>	<u>61.21</u>
DSBS	<b>0.1542</b>	<b>0.6926</b>	<b>76.04</b>	<b>61.64</b>

was trained with the E2E framework. Quantitative results in Tab. 6 show that the proposed DSBS method achieves superior performance across all segmentation metrics, confirming its adaptability and effectiveness beyond classification tasks.

Figure 6 illustrates qualitative segmentation results for three representative samples from the LIB-HSI dataset, comparing task-independent and task-based band selection methods. The DSBS method yields the most accurate segmentation maps across all samples, closely matching the ground-truth boundaries and material distributions. In particular, it captures fine structural details and material transitions with greater consistency than competing methods. Task-independent approaches such as BS-Nets-FC, BS-Nets-Conv, and Dropout CAE exhibit fragmented or oversmoothed outputs, while E2E-HSI fails to resolve certain class boundaries. The reported IoU values further support this observation, with DSBS consistently achieving the highest scores, reaching up to 82.37%. These results highlight the advantages of task-driven band selection in improving segmentation quality while reducing the number of required spectral bands.

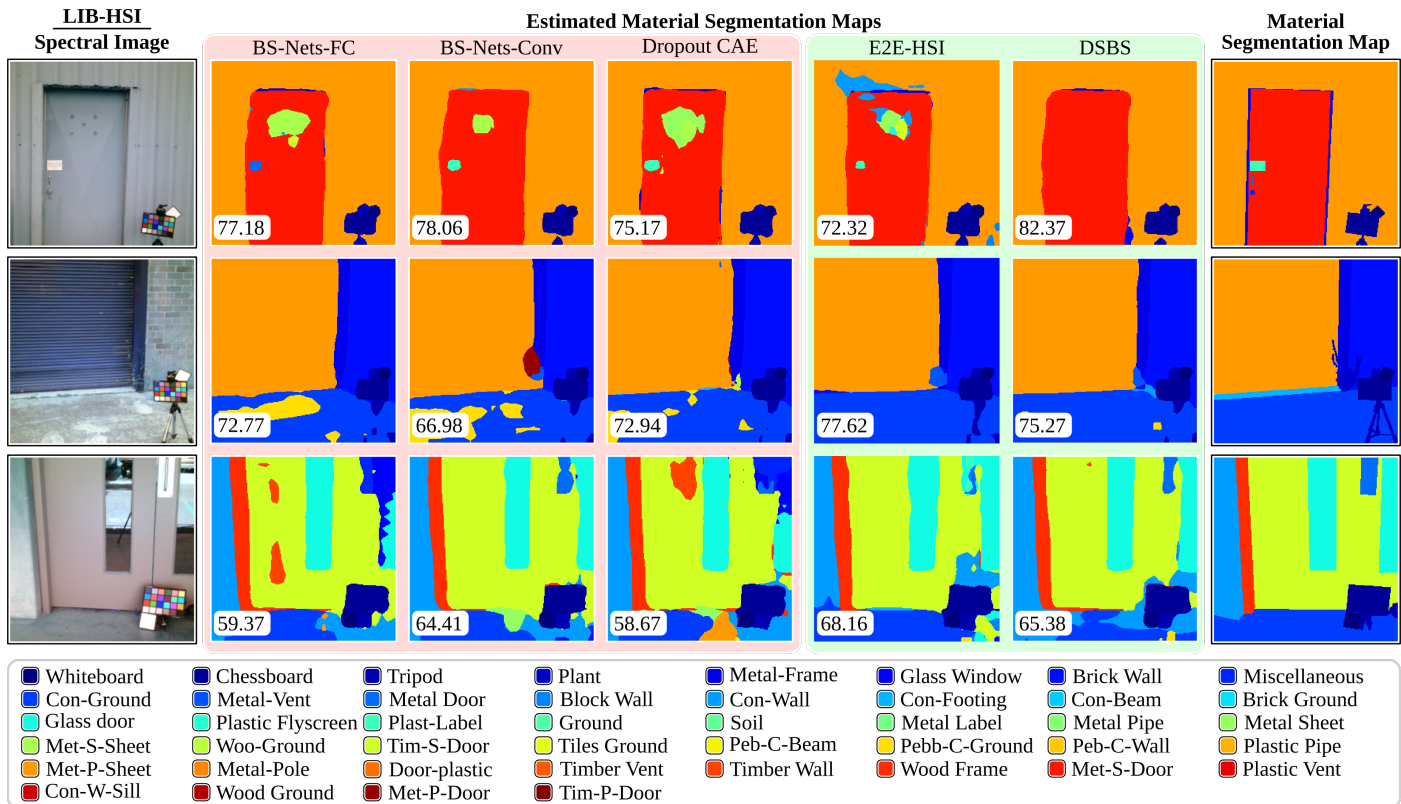
## 6. CONCLUSIONS AND FUTURE WORK

This work introduces DSBS, an E2E framework that integrates spectral BS with SI tasks. The DSBS framework provides an efficient solution for spectral imaging applications by incorporating a fully differentiable BS optical encoder, designed to identify and utilize the most relevant spectral bands for task-specific tasks, such as material segmentation and classification. By employing a discretization function and custom norm regularization, DSBS optimizes the selection of bands for each specific task, reducing the need for full-spectrum data. This results in decreased hardware costs and acquisition times, while achieving performance that is either comparable to or exceeds traditional full-spectrum imaging methods using only a fraction of the spectral bands. Future work can expand DSBS to other tasks, such as object detection and anomaly detection. Moreover, incorporating interpretability analyses into DSBS may offer deeper insights into how specific spectral bands contribute to various tasks, improving transparency and making the framework more suitable for decision-critical applications.

## REFERENCES

- Y. Liu, H. Pu, and D.-W. Sun, "Hyperspectral imaging technique for evaluating food quality and safety during various processes: A review of recent applications," *Trends food science & technology* **69**, 25–35 (2017).
- S. Peyghambari and Y. Zhang, "Hyperspectral remote sensing in lithological mapping, mineral exploration, and environmental geology: an updated review," *J. Appl. Remote. Sens.* **15**, 031501–031501 (2021).
- B. Lu, P. D. Dao, J. Liu, *et al.*, "Recent advances of hyperspectral imaging technology and applications in agriculture," *Remote. Sens.* **12**, 2659 (2020).
- N. Caporaso, M. B. Whitworth, and I. D. Fisk, "Near-infrared spectroscopy and hyperspectral imaging for non-destructive quality assessment of cereal grains," *Appl. spectroscopy reviews* **53**, 667–687 (2018).
- J. Y. Barnaby, T. D. Huggins, H. Lee, *et al.*, "Vis/nir hyperspectral imaging distinguishes sub-population, production environment, and physicochemical grain properties in rice," *Sci. reports* **10**, 9284 (2020).
- K. Sánchez, J. Bacca, L. Arévalo-Sánchez, *et al.*, "Classification of cocoa beans based on their level of fermentation using spectral information," *Tecnologías* **24**, 172–188 (2021).
- P. Ghamisi, J. Plaza, Y. Chen, *et al.*, "Advanced spectral classifiers for hyperspectral images: A review," *IEEE Geosci. Remote. Sens. Mag.* **5**, 8–32 (2017).





**Fig. 6.** Visual material segmentation results with IoU metric across three samples from the LIB-HSI dataset. From left to right: false-color RGB representation of the spectral image, estimated segmentation maps produced by task-independent methods (BS-Nets-FC, BS-Nets-Conv, Dropout CAE), task-based methods (E2E-HSI, DSBS), and the ground-truth material map. Reported values indicate segmentation performance for each method. Each reported result corresponds to the best OA performance for each method across 5 runs.

- H. Arguello, S. Pinilla, Y. Peng, *et al.*, "Shift-variant color-coded diffractive spectral imaging system," *Optica* **8**, 1424–1434 (2021).
- J. Zhang, R. Su, Q. Fu, *et al.*, "A survey on computational spectral reconstruction methods from rgb to hyperspectral imaging," *Sci. reports* **12**, 11905 (2022).
- K. Fonseca, J. Bacca, H. Garcia, and H. Arguello, "Deep jointly optical spectral band selection and classification learning," *Appl. Opt.* **63**, 5505–5514 (2024).
- C. O. Ayna, R. Mdrafi, Q. Du, and A. C. Gurbuz, "Learning-based optimization of hyperspectral band selection for classification," *Remote. Sens.* **15**, 4460 (2023).
- R. Jacome, J. Bacca, and H. Arguello, "D2uf: Deep coded aperture design and unrolling algorithm for compressive spectral image fusion," *IEEE J. Sel. Top. Signal Process.* **17**, 502–512 (2022).
- J. Bacca, E. Martinez, and H. Arguello, "Computational spectral imaging: a temporary overview," *JOSA A* **40**, C115–C125 (2023).
- J. Shi, X. Zhang, X. Liu, *et al.*, "Multicriteria semi-supervised hyperspectral band selection based on evolutionary multitask optimization," *Knowledge-Based Syst.* **240**, 107934 (2022).
- X. Sun, H. Zhang, F. Xu, *et al.*, "Constrained-target band selection with subspace partition for hyperspectral target detection," *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **14**, 9147–9161 (2021).
- T. Li, Y. Cai, Z. Cai, *et al.*, "Nonlocal band attention network for hyperspectral image band selection," *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **14**, 3462–3474 (2021).
- H. Su and Q. Du, "Hyperspectral band clustering and band selection for urban land cover classification," *Geocarto Int.* **27**, 395–411 (2012).
- Y. Yuan, J. Lin, and Q. Wang, "Dual-clustering-based hyperspectral band selection by contextual analysis," *IEEE Trans. on Geosci. Remote. Sens.* **54**, 1431–1445 (2015).
- Q. Wang, F. Zhang, and X. Li, "Optimal clustering framework for hyperspectral band selection," *IEEE Trans. on Geosci. Remote. Sens.* **56**, 5910–5922 (2018).
- S. Ning, W. Wang, and H. Liu, "Global-local consistency constrained deep embedded clustering for hyperspectral band selection," *IEEE access* (2023).
- C. Tang, X. Liu, E. Zhu, *et al.*, "Hyperspectral band selection via spatial-spectral weighted region-wise multiple graph fusion-based spectral clustering," in *IJCAI*, (2021), pp. 3038–3044.
- W. Sun, G. Yang, J. Peng, *et al.*, "A multiscale spectral features graph fusion method for hyperspectral band selection," *IEEE Trans. on Geosci. Remote. Sens.* **60**, 1–12 (2021).
- Y. Zhou, Q. Yao, S. Huo, and X. Li, "Hyperspectral band selection with iterative graph auto-encoder," *IEEE Trans. on Geosci. Remote. Sens.* (2023).
- Y. Cai, X. Liu, and Z. Cai, "Bs-nets: An end-to-end framework for band selection of hyperspectral image," *IEEE Trans. on Geosci. Remote. Sens.* **58**, 1969–1984 (2019).
- L. Mou, S. Saha, Y. Hua, *et al.*, "Deep reinforcement learning for band selection in hyperspectral image classification," *IEEE Trans. on Geosci. Remote. Sens.* **60**, 1–14 (2021).
- L. Xu, M. Ahishali, and M. Gabbouj, "Dropout concrete autoencoder for band selection on hsi scenes," *arXiv preprint arXiv:2401.16522* (2024).
- W. Sun, J. Peng, G. Yang, and Q. Du, "Correntropy-based sparse spectral clustering for hyperspectral band selection," *IEEE Geosci. Remote. Sens. Lett.* **17**, 484–488 (2019).
- W. Sun, J. Peng, G. Yang, and Q. Du, "Fast and latent low-rank subspace clustering for hyperspectral band selection," *IEEE Trans. on Geosci. Remote. Sens.* **58**, 3906–3915 (2020).
- H. Zhai, H. Zhang, L. Zhang, and P. Li, "Laplacian-regularized low-rank

- subspace clustering for hyperspectral image band selection," *IEEE Trans. on Geosci. Remote. Sens.* **57**, 1723–1740 (2018).
30. Q. Wang, Q. Li, and X. Li, "Hyperspectral band selection via adaptive subspace partition strategy," *IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens.* **12**, 4940–4950 (2019).
31. N. Habili, E. Kwan, W. Li, *et al.*, "A hyperspectral and rgb dataset for building façade segmentation," in *European Conference on Computer Vision*, (Springer, 2022), pp. 258–267.
32. L. Yan, M. Zhao, X. Wang, *et al.*, "Object detection in hyperspectral images," *IEEE Signal Process. Lett.* **28**, 508–512 (2021).
33. N. L. Johnson, "Systems of frequency curves generated by methods of translation," *Biometrika* **36**, 149–176 (1949).
34. Y. Bengio, N. Léonard, and A. Courville, "Estimating or propagating gradients through stochastic neurons for conditional computation," *arXiv preprint arXiv:1308.3432* (2013).
35. S. Dargan, M. Kumar, M. R. Ayyagari, and G. Kumar, "A survey of deep learning and its applications: a new paradigm to machine learning," *Arch. Comput. Methods Eng.* **27**, 1071–1092 (2020).
36. K. Fonseca, H. Garcia, F. da Silva, *et al.*, "Joint deep learning optical band selection and classification method for spectral data," in *Computational Optical Sensing and Imaging*, (Optica Publishing Group, 2023), pp. CTh2A–4.
37. C.-I. Chang, Y.-M. Kuo, and P. F. Hu, "Unsupervised rate distortion function-based band subset selection for hyperspectral image classification," *IEEE Trans. on Geosci. Remote. Sens.* **61**, 1–18 (2023).
38. S. Huang, H. Zhang, J. Xue, and A. Pizurica, "Heterogeneous regularization-based tensor subspace clustering for hyperspectral band selection," *IEEE Trans. on Neural Networks Learn. Syst.* **34**, 9259–9273 (2022).