

# LiDAR-Guided Cross-Attention Fusion for Hyperspectral Band Selection and Image Classification

Judy X Yang, *Student Member, IEEE*, Jun Zhou, *Senior Member, IEEE*, Jing Wang, *Senior Member, IEEE*, Hui Tian, *Senior Member, IEEE*, and Alan Wee-Chung Liew, *Senior Member, IEEE*

**Abstract**—The fusion of hyperspectral and LiDAR data has been an active research topic. Existing fusion methods have ignored the high-dimensionality and redundancy challenges in hyperspectral images, despite that band selection methods have been intensively studied for hyperspectral image (HSI) processing. This paper addresses this significant gap by introducing a cross-attention mechanism from the transformer architecture for the selection of HSI bands guided by LiDAR data. LiDAR provides high-resolution vertical structural information, which can be useful in distinguishing different types of land cover that may have similar spectral signatures but different structural profiles. In our approach, the LiDAR data are used as the “query” to search and identify the “key” from the HSI to choose the most pertinent bands for LiDAR. This method ensures that the selected HSI bands drastically reduce redundancy and computational requirements while working optimally with the LiDAR data. Extensive experiments have been undertaken on three paired HSI and LiDAR data sets: Houston 2013, Trento and MUUFL. The results highlight the superiority of the cross-attention mechanism, underlining the enhanced classification accuracy of the identified HSI bands when fused with the LiDAR features. The results also show that the use of fewer bands combined with LiDAR surpasses the performance of state-of-the-art fusion models.

**Index Terms**—Remote sensing, multimodal data fusion, hyperspectral images, LiDAR data, band selection, cross-attention.

## I. INTRODUCTION

Remote sensing has witnessed remarkable advances in recent decades and generated significant impacts in agriculture, environmental monitoring, and urban planning [1]. An active research topic in remote sensing is multimodal data fusion, which aims to explore heterogeneous features extracted from various sources and achieves results that single source data cannot achieve on its own [2].

Among data captured by various sensors, hyperspectral image (HSI) and light detection and range (LiDAR) data have unique advantages. HSI sensors capture the Earth’s surface in a wide range of electromagnetic spectrum, offering detailed spectral information for material identification. LiDAR data, on the other hand, are capable of providing 3D profiles to measure the distance between the sensor and the surface of the Earth, offering precise topographic and structural information [3].

Judy X Yang, Jun Zhou, Hui Tian and Alan Wee-Chung Liew are with the School of Information and Communication Technology, Griffith University, Australia (corresponding author: Jun Zhou, jun.zhou@griffith.edu.au).

Jing Wang is with the Queensland Department of Agriculture and Fisheries, Australia.

Fusion of hyperspectral and LiDAR data can leverage the complementary advantages of both data modalities and lead to improved image classification accuracy [4].

The rich spectral detail provided by the HSI results in high dimensionality and redundancy in HSI data, which leads to challenges such as the Hughes phenomenon, a large computational load, and inefficiency in storage and transmission. Hyperspectral band selection is one way to mitigate these challenges, which attempts to identify and preserve a subset of spectral bands from the original HSI that contribute the most pertinent information while discarding redundant or irrelevant bands. Hyperspectral band selection has been widely studied, whose methods fall mainly into five categories: ranking-based, search-based, clustering-based, deep learning-based, and their hybrid [5].

In parallel, significant research efforts have been dedicated to the fusion of hyperspectral and LiDAR data to improve the classification of land cover. Methods on this topic can be broadly categorised into traditional [6] and deep learning-based approaches [7], [8]. Traditional fusion methods typically involve pixel-based or feature-based approaches, using statistical and signal processing techniques such as PCA or wavelet transform to combine and extract valuable features from data sets [9]–[11]. On the other hand, deep learning-based approaches have introduced deep neural networks [12], [13] and vision transformers for joint hyperspectral and LiDAR classification, which address limitations in heterogeneous feature representation and information fusion [14], [15].

Fig. 1 shows the correlation coefficients between LiDAR and each hyperspectral band of three datasets: Houston 2013, Trento, and MUUFL. [16]–[18]. We can see that a certain level of correlation between LiDAR and hyperspectral bands exists. This correlation fluctuates among different spectral bands. It is generally believed that LiDAR data contain distance information, and HSI contain chemical components information, which seems to be independent and complementary. However, two factors could contribute to such correlation phenomenon. First, the characteristics of the ground cover structures may cause the correlation between LiDAR and HSI bands. For example, tree canopy and ground soil in a same area have difference heights, which can be captured by LiDAR. Meanwhile, these two classes have different chemical components, and thus have different spectral responses. This leads to a certain level of correlation between LiDAR and HSI. Since different spectral bands are sensitive to different materials, we can see that the

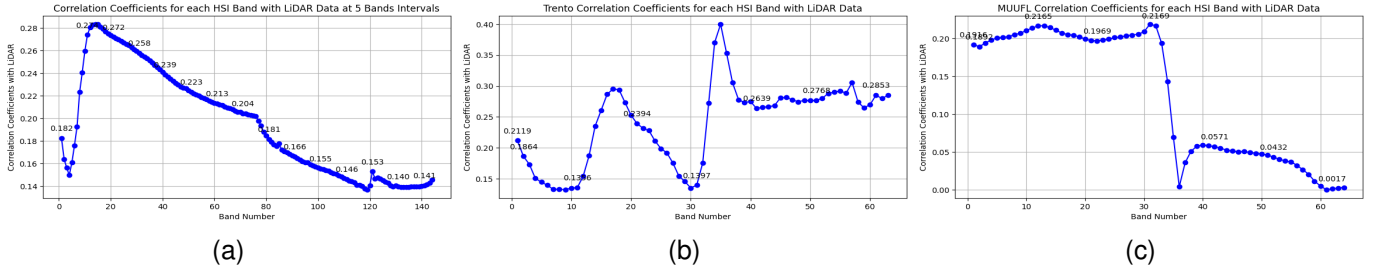


Fig. 1. Correlation coefficients between LiDAR and each hyperspectral band of Three Datasets. (a) Houston2013 HSI and LiDAR. (b) Trento HSI and LiDAR. (c) MUUFLE HSI and LiDAR. The horizontal axis shows the band index.

correlation varies across different bands and different datasets. Second, LiDAR also relies on the specific wavelength of the laser to sense the objects. The relationship of wavelength in LiDAR and HSI bands could have contribution to the correlation phenomenon

Additionally, HSI sensors produce 2D raster-based images that have a low spatial resolution compared to LiDAR data. LiDAR sensors produce 3D. Although there is no direct correspondence one-to-one between the two datasets [13], [19], it can be seen that some hyperspectral bands have correlations with the LiDAR data more than others. In addition to containing complementary information, some hyperspectral bands, after adding the LiDAR information into the raster-based HSI process, helped to distinguish spectrally similar materials, such as trees and grass [19]. When selecting important hyperspectral bands with LiDAR, we also need to consider the overlap of this information. The importance level of a hyperspectral band decreases when it is highly correlated with the already available LiDAR. In statistics, Pearson's correlation coefficient measures the linear relationship between two sets of data. A higher coefficient indicates a stronger linear correlation. Therefore, when the coefficient is high, it suggests that the data points of the two data sets (in this case, LiDAR and certain hyperspectral bands) have a strong linear relationship, implying that they contain similar information [20], [21]

The integration of LiDAR data into raster-based HSI analysis increases the distinction of spectrally similar materials, exemplified by the differentiation of trees and grass [19]. Therefore, it is imperative to consider the informational overlap when selecting pertinent HSI bands to pair with LiDAR data. The utility of an HSI band may diminish if its correlation with existing LiDAR data is substantial. Pearson's correlation coefficient, as a statistical measure, gauges the linearity of the relationship between two datasets. A pronounced coefficient suggests a robust linear association, signifying considerable overlap in the information conveyed by the respective datasets. Thus, the high correlation coefficients between LiDAR and certain HSI bands underscore the presence of analogous information within these paired datasets.

Building on the hyperspectral band selection and HSI and LiDAR fusion techniques, we identify the following two research gaps in the existing literature. First, there is a lack of well-defined methods for selecting the most pertinent bands in HSI when LiDAR data are available, limiting the potential synergies that can be exploited by integrating LiDAR and HSI

data for enhanced data classification and analysis. Although traditional band selection methods can select important bands from hyperspectral data, they are designed for hyperspectral data only. When paired with LiDAR data, the selected band of these methods is not guaranteed to be optimal.

Second, although it is widely accepted that the Hughes phenomenon hampers the performance of hyperspectral image classification, to the best of our knowledge, almost all hyperspectral and LiDAR fusion methods still use all hyperspectral bands. It is important to investigate whether the fusion of fewer bands with LiDAR would lead to better classification performance.

We strive to bridge these gaps in current research on band selection and data fusion. Inspired by the attention mechanism of transformer architectures in domains such as natural language processing and image processing, we propose a novel paradigm for HSI band selection, guided by the attributes of LiDAR data. We aim to show that leveraging LiDAR's spatial properties with HSI's spectral properties can lead to a new solution in HSI and LiDAR data fusion.

The contributions of this paper are summarised as follows:

- 1) We identify the existing gaps in HSI and LiDAR data fusion and band selection research. Then we leverage the latest success of transformer architecture and propose a novel cross-attention mechanism that uses LiDAR data to guide HSI band selection.
- 2) Our model is tested on three HSI and LiDAR data fusion datasets and consistently outperforms existing band selection techniques in HSI classification. The experiments demonstrate the effectiveness of our model.
- 3) Our approach demonstrates that selecting significant bands and combining them with LiDAR data can achieve better performance than current fusion models, suggesting new prospects for remote sensing data fusion.

The paper is structured as follows. Section II reviews related works in the literature. Our method is described in Section III, followed by the experiments and results presented in Section IV. Finally, Section V concludes this article and highlights directions for future research.

## II. RELATED WORK

HSI and LiDAR data fusion has shown its effectiveness in improving the performance of various remote sensing applications, such as land cover classification, target detection,

and environmental monitoring. In this section, we review the literature with a focus on band selection strategies and data fusion methodologies relevant to our study.

### A. Hyperspectral Band Selection

The core objective of hyperspectral band selection is to identify crucial and distinguishing bands in HSI, which aims to reduce data dimensionality and computational cost while preserving vital information in the images. We briefly review five predominant categories of hyperspectral band selection methods: ranking-based, searching-based, clustering-based, deep learning-based, and hybrid methods [5], [22].

Ranking-based methods prioritise band indexes using different statistical data analyses, for example, covariance [23], correlation [24], and structural similarity [25]. These approaches rank the bands based on the uniqueness or the representative of the bands on a given dataset. Searching-based methods perform band selection using a distance measurement function with a searching strategy. Distance measurements used in hyperspectral imaging are crucial for various tasks. Spectral Angle Mapping (SAM), as mentioned in Liu et al. (2013) [26], quantifies uncertainty or randomness in spectral information, useful for assessing data complexity. The Maximum Ellipsoid Volume, as discussed in [27], is used to maximise the data scatter in a reduced-dimensional space. The Structural Similarity Index (SSI) [25] is another important measure for assessing the similarity between images or spectra. The categorisation of search strategies in hyperspectral band selection, as described by Sun et al. (2019) [5], includes incremental, updated and eliminating strategies, each with a specific methodology for selecting the optimal bands. Incremental strategy iteratively adds bands, updated strategy refines choices as new data become available, while eliminating strategy removes less informative bands systematically.

Clustering-based methods group the original bands into clusters, selecting representative bands from each for the final subset. Initially using hierarchical clustering with Ward linkage [26], [28], subsequent methods used information measurements such as MI [29] or Kullback-Leibler divergence [26]. Commonly used clustering approaches consist of k-means [30], affinity propagation [31], graph clustering [32], optimal band clustering [33], fuzzy clustering [34], and spectral-spatial subspace clustering [35].

Deep learning-based methods adopt deep neural networks for band selection. Among them, BSNets [36] adopt the Band Attention Module (BAM) to model complex relationships between spectral bands, then use a Reconstruction Network (RecNet) to validate the chosen bands by reconstructing the original HSI from the BAM-selected bands. Deep reinforcement learning [37] frames unsupervised band selection as a Markov decision process and automatically learns a policy to select an optimal subset of bands. Constrained measurement learning network [38] is composed of a constrained measurement learning network for band selection, followed by a classification network. These two networks are jointly trained to minimise the classification loss along band selection. Deep learning with spatial features [22] integrates spectral-spatial features into a

deep learning framework for the selection of hyperspectral bands. These methods provide a variety of ways to select informative and non-redundant bands from hyperspectral data using deep learning.

Hybrid band-selection methods employ multiple strategies for optimal band selection. Commonly, they combine clustering and ranking approaches [39]. Zhao et al. presented an unsupervised Spectral-Spatial genetic algorithm-based method [40], which integrated search-based, deep learning-based, rank-based, and clustering-based approaches. The spectral separability index algorithm integrates clustering, classification and searching for the best band combination [40]–[42]. These hybrids aim to achieve higher classification accuracy and band significance by leveraging the strengths of multiple selection techniques.

Although current band-selection studies predominantly focus on HSI data, our objective is to identify the most informative hyperspectral bands in the presence of LiDAR data, and we aim to compare our method with existing band-selection approaches using the same paired HSI and LiDAR data sets.

### B. HSI and LiDAR Fusion Models

HSI and LiDAR data fusion has witnessed substantial progress. Although traditional methods have their merits [43], deep learning techniques have proven superior in handling the complexity of these data sets [13]. In particular, convolutional neural network (CNN)-based architectures have been instrumental in HSI and LiDAR data fusion, offering robust spatial feature extraction capabilities [8], [44]. Huang et al. [44] introduced a framework that uses two coupled convolutional neural networks (CNN) to fuse hyperspectral and LiDAR data, extracting spectral-spatial features and elevation information, respectively. EndNet [45], an Encoder-Decoder fusion network, addresses the limitations of single-modal remote sensing data and uses a reconstruction strategy to activate neurons in both modalities. The collaborative contrastive learning (CCL) method [46] uses a pretraining and fine-tuning strategy to extract features from HSI and LiDAR data separately and produce coordinated representation and matching of features between the two-modal data without label samples. Integrating single-band and multi-band LiDAR fusion with hyperspectral data, a deep learning fusion framework [47] was introduced to merge HSI and LiDAR for tree species mapping.

With the integration of deep learning and attention-based mechanisms [48], the effectiveness of data fusion becomes more effective. The FusAtNet model [49] fuses self-attention and cross-attention module, highlighting the harmonious integration of the HSI and LiDAR data. Falahatnejad et al. [8] used hybrid 3D/2D CNNs and attention module to extract spatial-spectral characteristics of HSI and representative elevation characteristics of LiDAR. The extracted features were then integrated and classified using a softmax classifier. The collaborative attention network [50] addresses limitations in the representation of heterogeneous features. Both multiscale and single-branch networks are used to extract land cover features enhanced by self-attention and cross-attention module.

Deep learning and attention-based fusion techniques take advantage of the strengths of both types of data, providing

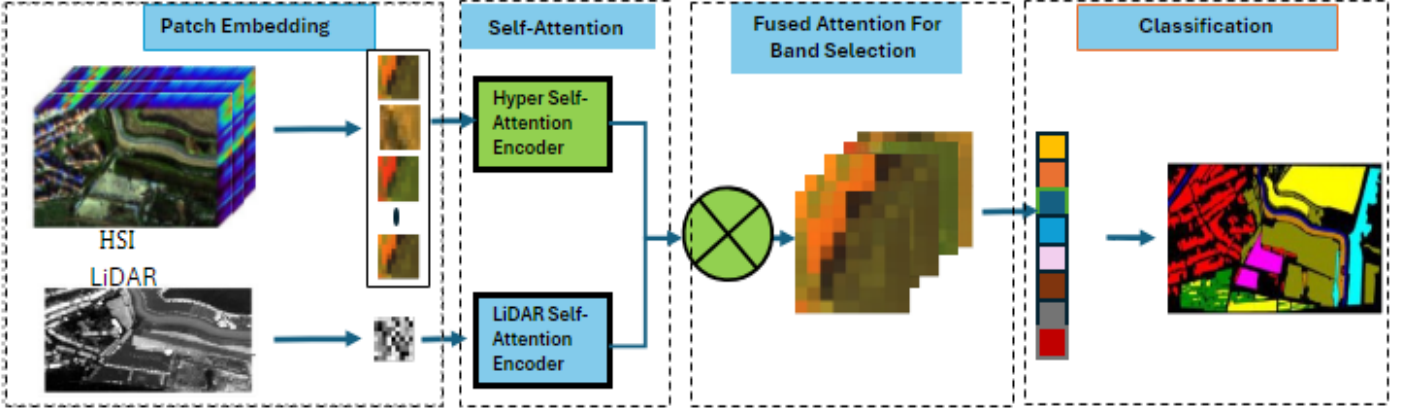


Fig. 2. The overall framework of the proposed LiDAR-guided cross-attention fusion for band selection.

comprehensive insight that is crucial for various applications. However, there is a noticeable gap as current approaches tend to use full HSI bands and LiDAR data without performing any band selection. This lack of band selection can lead to unnecessary computational complexity and potential information redundancy, highlighting the need for further research and development in this area.

### III. METHODOLOGY

This section introduces the proposed integrated LiDAR-guided band selection framework. This framework consists mainly of two modules: the self-attention module for feature extraction from hyperspectral and LiDAR data and the LiDAR-guided cross-attention module. We first introduce the overview framework, followed by the details of the attention module.

#### A. Overview of the Proposed Fusion Framework

Our method leverages the transformer, originally prominent in natural language processing (NLP) for sequence-to-sequence tasks such as machine translation [51], for HSI and LiDAR data fusion. Transformers have revolutionised data processing by using self-attention mechanisms to model long-range dependency in the data. Beyond NLP, the potential of transformers has been validated in computer vision and hyperspectral image classification, with innovations such as the Vision Transformer (ViT) [52] achieving or even surpassing traditional CNN benchmarks in various visual tasks.

Drawing from these advances, we propose an attention-based mechanism adapted from transformer architecture to model the relationship between HSI and LiDAR and obtain the significance of each HSI band in the context of LiDAR fusion. Fig. 2 shows the model work flow. The model consists mainly of four modules, in Figure 3, i.e. the input and patch embedding module, the self-attention module, the cross-attention module, and the Multi-Layer Perceptron (MLP) and classification module.

The input to our model is a hyperspectral image  $\mathbf{X}_H \in \mathbb{R}^{W \times H \times B}$  and the corresponding LiDAR data  $\mathbf{X}_L \in$

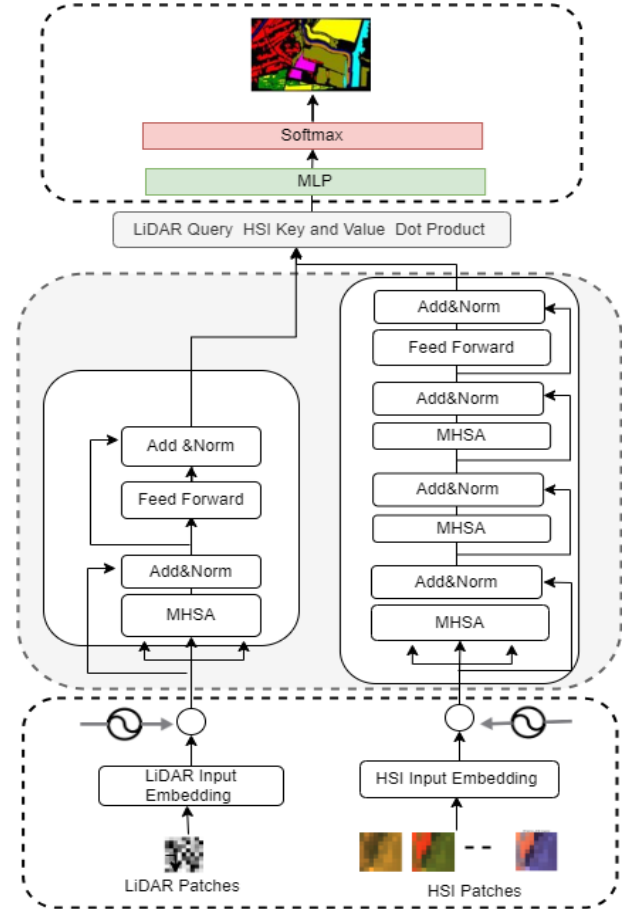


Fig. 3. The detailed illustration of the Proposed LiDAR-guided cross-attention fusion for band selection method. The network consists of four components: A) input and patch embedding module; B) self-attention module; C) cross-attention module; and D) MLP.

$\mathbb{R}^{W \times H \times C}$ , where  $W$  and  $H$  represent the width and height of the dataset, and  $B$  and  $C$  are the number of HSI bands and the number of LiDAR channels, respectively. A pair of input sample patches extracted from hyperspectral and LiDAR data is denoted as  $\mathbf{X}_h \in \mathbb{R}^{P \times P \times B}$  and  $\mathbf{X}_l \in \mathbb{R}^{P \times P \times C}$ , where  $P$

represents the size of the sample patch. To formulate the raw input data into sequences for the transformer-like architecture, we flatten the spatial dimensions of the HSI and LiDAR data so that they are represented as vectors, with each flattened band representing a token. Thus, the results are  $\hat{\mathbf{X}}_{h0} \in \mathbb{R}^{M \times B}$  and  $\hat{\mathbf{X}}_{l0} \in \mathbb{R}^{M \times C}$ , respectively, where  $M = P^2$ .

Although grouping neighbouring bands into one token is often adopted in the hyperspectral image classification [53], [54], we use a band-wise token. The reason is that we can directly assess the importance and relationships of individual bands from hyperspectral and LiDAR data, aligning with our goal of selecting LiDAR-guided hyperspectral bands. Like classical transformers, the inputs of both hyperspectral and LiDAR data go through an embedding layer and are added with positional encoding to keep the band-order information. After embedding, the characteristic of the hyperspectral data becomes  $(d, B)$ , where  $d$  denotes the embedding size.

The main part of the proposed method comprises two modules: the self-attention module and the cross-attention module. The hyperspectral and LiDAR features go through two separate branches of stacked self-attention module. These stacked self-attention modules extract feature representations of each hyperspectral band and LiDAR channel individually. Finally, the two streams interact with each other at the end of the network with a cross-attention module to fuse and weigh the importance of each hyperspectral band with respect to the LiDAR data. The motivation of this design is aligned with our goal of selecting hyperspectral bands using LiDAR as a query. Finally, at the end of the network, the fused and weighted features are sent through the fully connected layer and the softmax function to output the classification results.

### B. Self attention and cross attention modules

Multihead attention was first proposed in language models to process long sequences [51]. We adopted it here to learn the representation of each band while considering the context of information within each data modality. Specifically, each hyperspectral band representation  $r_i, i = 1, 2, \dots, B$  is mapped to three spaces, namely  $q_i, k_i,$  and  $v_i$ . The three spaces correspond to the query, key, and value space. Then the attention can be achieved by:

$$Attention(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = softmax\left(\frac{\mathbf{Q} \cdot \mathbf{K}^T}{\sqrt{d}}\right)\mathbf{V} \quad (1)$$

Equation (1) denotes a process of self-attention for a single head. In practical applications, it is beneficial to map the representations to different Query, Key, and Values spaces and perform attention in multiple different spaces. Finally, the results of multiple attentions are concatenated together. A typical transformer encoder layer often has a feed-forward network attached after multi-head attention to learn better features.

The stacked self-attention layers can learn highly distinctive features for hyperspectral and LiDAR, respectively, but cannot model the relationship and significance of each hyperspectral band with respect to LiDAR. To achieve this goal, we further

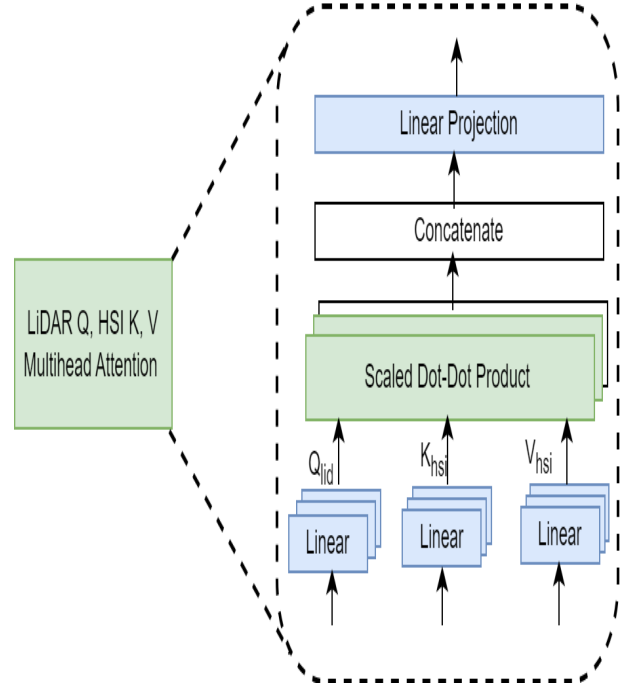


Fig. 4. Cross attention from LiDAR and HSI: LiDAR queries undergo linear transformation, interact with hyperspectral keys/values through scaled dot-product attention, concatenate across multiple heads, followed by a final linear projection for enriched data fusion.

design a cross-attention module that facilitates the LiDAR data to focus on specific hyperspectral bands, as shown in Fig. 4.

After going through the stacked self-attention module, the hyperspectral and LiDAR feature representation can be denoted by  $\hat{\mathbf{X}}_h$  and  $\hat{\mathbf{X}}_l$  with sizes  $(d, B)$  and  $(d, C)$ . To model the relationship of LiDAR to each hyperspectral band, we choose a different querying strategy from the self-attention module. Specifically, we only map LiDAR data to the query space and hyperspectral bands to the key space. Then  $Q_L$  is used to query the hyperspectral band's  $K_H$ . In this manner, the LiDAR and each hyperspectral band can interact directly with each other to model the importance of each hyperspectral band to the LiDAR, such that:

$$\mathbf{Q}_L = \mathbf{W}_q \times \hat{\mathbf{X}}_l, \quad (2)$$

$$\mathbf{K}_H = \mathbf{W}_k \times \hat{\mathbf{X}}_h, \quad (3)$$

$$\mathbf{V}_H = \mathbf{W}_v \times \hat{\mathbf{X}}_h. \quad (4)$$

where  $\mathbf{W}_q, \mathbf{W}_k$  and  $\mathbf{W}_v$  are the projection matrices for the three spaces, respectively. Cross-attention weights can be obtained by:

$$AttWeights_{singlehead} = softmax\left(\frac{\mathbf{Q}_L \cdot \mathbf{K}_H^T}{\sqrt{d}}\right) \quad (5)$$

For example, when LiDAR has one channel, hyperspectral has 144 bands, and the embedding dimension is 256,  $Q_L$  and  $K_H^T$  are in the size of  $1 \times 256$  and  $256 \times 144$ , respectively. Thus, the final result has the size of  $1 \times 144$ . After applying the softmax function, we can obtain a normalised 144-dimensional weight vector, each element of which represents the relevance



of the individual bands of hyperspectral in regards to LiDAR. The final attention weights are averaged over all heads when adopting multi-head in the cross-attention module:

$$AttWeights_{\text{multihead}} = \sum_{i=1}^N \frac{1}{N} \text{softmax} \left( \frac{\mathbf{Q}_L^i \cdot (\mathbf{K}_H^i)^T}{\sqrt{d_k}} \right) \quad (6)$$

where  $N$  is the number of heads.

By applying the attention weights on the hyperspectral band, we can obtain the output of the cross-attention module. The cross-attention module performs band weighting and data fusion at the same time. Training this whole network leads to the generation of the optimal weights of hyperspectral bands for LiDAR data on the fusion task. Thus, attention weights can be used to select the most useful hyperspectral bands for fusion with LiDAR, i.e., the bands with the highest weights.

Note that the learned attention weights depend on each input pair of hyperspectral and LiDAR samples. This means that the best hyperspectral band subsets for LiDAR fusion might be different for different ground objects. This also makes sense intuitively. On the contrary, we can also obtain a consistent set of selected bands for all classes. A simple solution is to input all training samples into the network, obtain attention weights for each sample, and then average over them.

#### IV. EXPERIMENTS

We have undertaken extensive experiments to assess the performance of the proposed cross-attention band selection method. The experiments were run on three paired HSI and LiDAR datasets, Houston 2013, Trento, and MUUFL.

**Houston 2013 Dataset:** This dataset contains hyperspectral and LiDAR data for the 2013 IEEE GRSS Data Fusion [55]. It includes a hyperspectral image (HSI) with 144 spectral bands (380-1050nm) and a LiDAR-derived digital surface model (DSM), both at 2.5 m resolution. The HSI is calibrated to the radiance of the sensor, while the DSM indicates elevation in meters above sea level. The data set contains 15 types of land cover, making it ideal for evaluating band selection and classification methods, despite urban complexity and challenges related to noise in the HSI. Details on the training and test samples from the experiment are given in TABLE I. Fig. 5 shows the false colour images of the Houston 2013 data set.

**Trento Data Set:** The Trento data set [17] was acquired in Trento, Italy, in 2015. The hyperspectral image comprises 48 spectral bands (400-950nm) with one-metre spatial resolution, while the LiDAR data presents the DSM. The data set covers approximately 100 hectares and includes six land cover classes. Table II outlines the number of training and test samples used in the current experiment for each class of interest. Fig. 6 shows the false colour images of the Trento data set.

**MUUFL Gulfport scene:** The MUUFL Gulfport scene [18] contains hyperspectral and LiDAR data collected over Gulfport, Mississippi, in 2010 and 2011. The data set includes four sub-images with different spatial resolutions and elevations and ground-truth information. TABLE III gives the distribution of the sample in this experiment. Fig. 7 shows the false-colour images of the MUUFL data set.

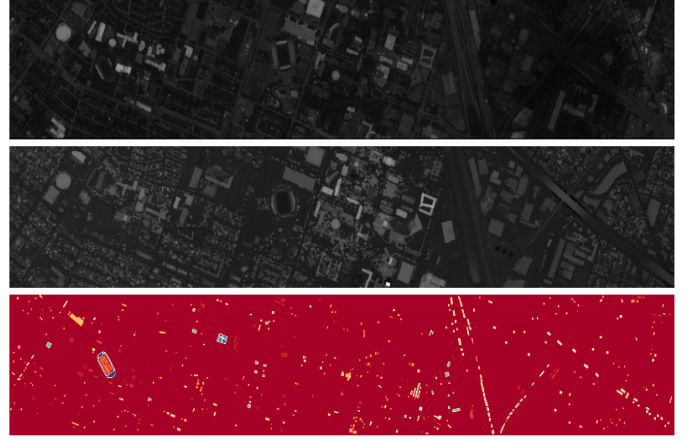


Fig. 5. Visualisation of the Houston 2013 data set. Top: HSI; Middle: LiDAR data; Bottom: ground truth.

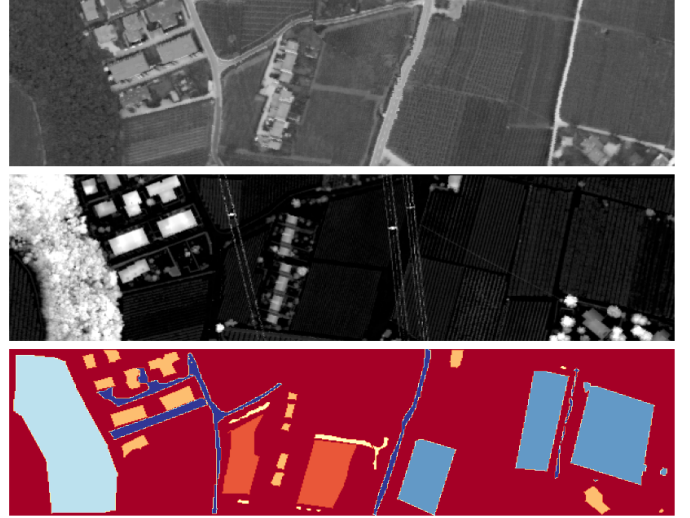


Fig. 6. Visualisation of the Trento data set. Top: HSI; Middle: LiDAR data; Bottom: ground truth.

##### A. Experimental Setup

We undertook two sets of experiments to analyse the performance of the proposed LiDAR-guided cross-attention band selection method. The first set of experiments compares our method with traditional hyperspectral band-selection methods, which do not consider the LiDAR data in the band-selection. The results are presented in terms of classification performance after fusion of the selected hyperspectral bands with the LiDAR data. The second set of experiments compares the results of our method with several hyperspectral and LiDAR fusion models without band selection.

In the experiment, both Support Vector Machine (SVM) and Convolutional Neural Networks (CNN) were used for the classification task. The standard Libsvm toolbox was adopted for the SVM. For the CNN, a conventional 6-layer network was selected, with kernel size  $3 \times 3$ , dropout 0.4, a dense layer, and a ReLU activation function. A softmax function was used for classification purposes. The batch size was set to 32, the learning rate was 0.0001, and the epoch was 50. The

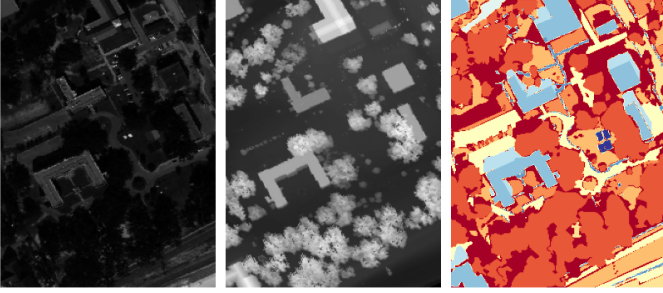


Fig. 7. Visualisation of the MUUFL data set. Left: HSI; centre: LiDAR data; right: ground truth.

TABLE I  
HOUSTON 2013 DATASET

No.	Class Name	Training	Test	Samples
1	Healthy grass	198	1053	1251
2	Stressed grass	190	1064	1254
3	Synthetic grass	192	505	697
4	Tree	188	1056	1244
5	Soil	186	1056	1242
6	Water	182	143	325
7	Residential	196	1072	1268
8	Commercial	191	1053	1244
9	Road	193	1059	1252
10	Highway	191	1036	1227
11	Railway	181	1054	1235
12	Parking lot 1	192	1041	1233
13	Parking lot 2	184	285	469
14	Tennis court	181	247	428
15	Running track	187	473	660
	<b>Total</b>	<b>2832</b>	<b>12197</b>	<b>15029</b>

classification performance was evaluated using three commonly used criteria, i.e., overall accuracy (OA), average accuracy (AA), and kappa coefficient (Kappa).

### B. Implementation Details

The experiments were carried out on the PyTorch platform, and the model was trained on the GPU-enabled machine of the Colab Premium version to accelerate the computation.

On the Houston and Trento dataset, the size of the HSI sample patch was  $9 \times 9$ . Following the patch embedding method of the visual transformer, the sample size  $9 \times 9 \times 144$  is divided into 144 pieces of  $9 \times 9$  patches, and the kernel size is set to  $1 \times 1$ , with the number of patches equal to the number of HSI bands, ensuring a one-to-one correspondence between the patches and the spectral bands for effective band selection. LiDAR data were processed similarly with the patch size set to  $9 \times 9$ , and the patch number was 1. For the MUUFL dataset, the patch size was reduced to  $3 \times 3$ . A linear projection layer transformed each patch into a flattened embedded representation. The embedding size was 256, ensuring that each spectral band was treated as a distinct entity. Consequently, both the HSI and LiDAR embedding were added with positional embedding of size 256 to retain spatial information.

Each self-attention module consisted of 3 layers with 8 heads, where the dimensionality of each head was 128 and the multilayer perceptron (MLP) dimension within the transformer

TABLE II  
TRENTO DATA SET

No.	Class Name	Training	Test	Samples
1	Apple trees	129	3905	4034
2	Buildings	125	2778	2903
3	Ground	105	374	479
4	Wood	154	8969	9123
5	Vineyard	184	10,317	10501
6	Roads	122	3052	3174
	<b>Total</b>	<b>819</b>	<b>29,395</b>	<b>30,214</b>

TABLE III  
MUUFL DATA SET

No.	Class Name	Training	Test	Samples
1	Trees	150	23,246	23,396
2	Mostly grass	150	4,270	4,420
3	Mixed ground surface	150	6,882	7,032
4	Dirt and sand	150	1,826	1,976
5	Road	150	6,687	6,837
6	Water	150	466	616
7	Building shadow	150	2,233	2,383
8	Building	150	6,240	6,396
9	Sidewalk	150	1,385	1,535
10	Yellow curb	150	183	333
11	Cloth panels	150	269	419
	<b>Total</b>	<b>1650</b>	<b>53,687</b>	<b>55,337</b>

blocks was 256. A cross-attention mechanism then integrated the information from both modalities, allowing each LiDAR patch to attend to all HSI patches, using the attention scores for band selection. Adam optimiser with a learning rate of 0.0001 was used to train the model. The model was trained with a batch size of 32.

Training was terminated after 50 epochs with an early stopping criterion to prevent overfitting. A cross-entropy loss function was employed to calculate the difference between the predicted class probabilities and the ground-truth labels. We report the results with and without augmented training data. In the former case, the augmented samples were generated through transformations including 45- and 90-degree rotations and both vertical and horizontal flips.

After training, the network cross-attention module assigned significance to each hyperspectral band based on its relevance to the corresponding LiDAR data. The cross-attention module outputs the attention weights, showing the relevance between the HSI and LiDAR data. The weights were then used to predict important band selection. Then, this subset of selected bands is used for subsequent data fusion tasks.

### C. Ablation Study

In this study, we investigate the incremental performance enhancements offered by our proposed cross-attention band selection model, particularly focusing on the accuracy of classification. This investigation is achieved by sequentially omitting the cross-attention component and instead employing alternatives such as hyperspectral only self-attention and a mixed self-attention mechanism that incorporates both LiDAR and hyperspectral imagery (HSI) within the network architecture. A series of ablation experiments were conducted using

TABLE IV  
 ABLATION ANALYSIS OF THE PROPOSED LiDAR GUIDED BAND SELECTION MODEL WITH A COMBINATION OF TWO SELF ATTENTION MODULES ON HOUSTON 2013 DATASET .THE BEST RESULTS OF SVM ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Band Number	Metrics	Self-Attention A Trained On HSI+LiDAR	Self-Attention B Trained on HSI	Self-Attention+Fusion OurMethod(Raw Data)	Self-Attention+Fusion OurMethod(Augmented Data)
30	OA	0.9125	0.9073	0.9201	0.9242
	AA	0.9248	0.9202	0.9301	0.9373
	kappa	0.9050	0.8994	0.9133	0.9177
25	OA	0.9125	0.9033	0.9193	0.9266
	AA	0.9248	0.9171	0.9284	0.9392
	kappa	0.9050	0.8950	0.9124	0.9203
20	OA	0.9118	0.9101	0.9198	0.9253
	AA	0.9235	0.9223	0.9283	0.9378
	kappa	0.9042	0.9024	0.9124	0.9189
15	OA	0.9134	0.9064	0.9019	0.9251
	AA	0.9246	0.9242	0.9187	0.9378
	kappa	0.9060	0.8984	0.8935	0.9186
10	OA	0.8742	0.9097	0.8938	0.9317
	AA	0.8909	0.9209	0.9107	0.9427
	Kappa	0.8634	0.9019	0.8848	0.9259
5	OA	0.8766	0.8762	0.9079	0.9202
	AA	0.8929	0.8879	0.9210	0.9335
	Kappa	0.8660	0.8557	0.9000	0.9134

the SVM classifier in the Houston 2013 data set to rigorously assess the contribution of the cross-attention module to the HSI and HSI + LiDAR classification tasks.

More specifically, the ablation study explores two different configurations of self-attention: Self-attention A leverages a composite importance score that amalgamates information from both HSI and LiDAR data for band selection. Conversely, Self-attention B uses only hyperspectral image data, where the selection of bands is guided by the importance scores derived from hyperspectral image self-attention post-model training. These methodologies were integrated into our model’s framework, and were evaluated across both raw and augmented data scenarios.

Subsequent SVM classification tests were conducted to compare the efficacy of different band selection strategies, incorporating the varied band indices generated by the four examined methods.

The results of these experiments are systematically presented in Table IV. We can see that the augmented data-enhanced cross-attention module secures the top performance ranking. This is closely followed by the model employing the cross-attention module trained on raw data. In contrast, the models utilizing Self-attention A and Self-attention B configurations exhibit comparatively diminished performance, underscoring the superior effectiveness of the cross-attention mechanism in our band selection model.

This ablation study conclusively affirms the pivotal role of the cross-attention module in optimising band selection for HSI and HSI + LiDAR fusion applications, highlighting its significant impact on enhancing classification accuracy.

#### D. Comparison with HSI band selection methods

In this set of experiments, we fused bands selected by several recent band selection approaches and concatenated them with the corresponding LiDAR data before calling the classifiers. The compared band selection methods included low-rank representation-based band selection (*LRR – BS*) [39],

adaptive spatial spectral patch selection (ASPS) [56], orthogonal projection band selection (OPBS) [27], optimal clustering framework (TRC-OC-FDPC) [33] and band selection network (BSNET) [36]. We tested the performance of each approach with 1 to 50 selected bands and all bands used based on both SVM and 1DCNN.

1) *Results on the Houston 2013 data set:* We show the experimental results in Fig. 8a for OA and in Tables V and VI for OA, AA and Kappa. Please note that due to space limitations, the curves on AA and kappa are not reported here, but can be accessed on GitHub<sup>1</sup>. Our methodology has consistently outperformed other band selection techniques in evaluations with SVM and CNN classifiers, demonstrating its robustness and effectiveness. When using models trained on raw data, our approach not only maintains stable performance but also achieves superior accuracy and Kappa scores, particularly when selecting from 15 to 40 bands. In contexts where CNN classifiers are employed, although BSNET shows promising initial performance, our method, applied to raw data, exceeds it in efficiency with an increasing number of bands, specifically at 5, 15, 20, 25, 30, and 35 bands. The performance enhancement becomes even more pronounced with models trained on augmented data. In such scenarios, both the SVM and CNN classifiers exhibit exceptional performance, significantly outperforming competing methods in nearly all band selection scenarios. This empirical evidence strongly supports our hypothesis that band selection, when combined with LiDAR data, significantly increases the effectiveness of the classification process after multimodal fusion.

2) *Results on the Trento data set:* Fig. 9 and Tables VII and VIII show OA, AA, and Kappa with various bands selected for the SVM and the CNN classifiers. According to the results, our method with the raw training dataset outperforms other methodologies using the SVM classifier when the band numbers

<sup>1</sup><https://github.com/Judyxyang/LiDAR-Guided-Band-Selection/tree/main/img>



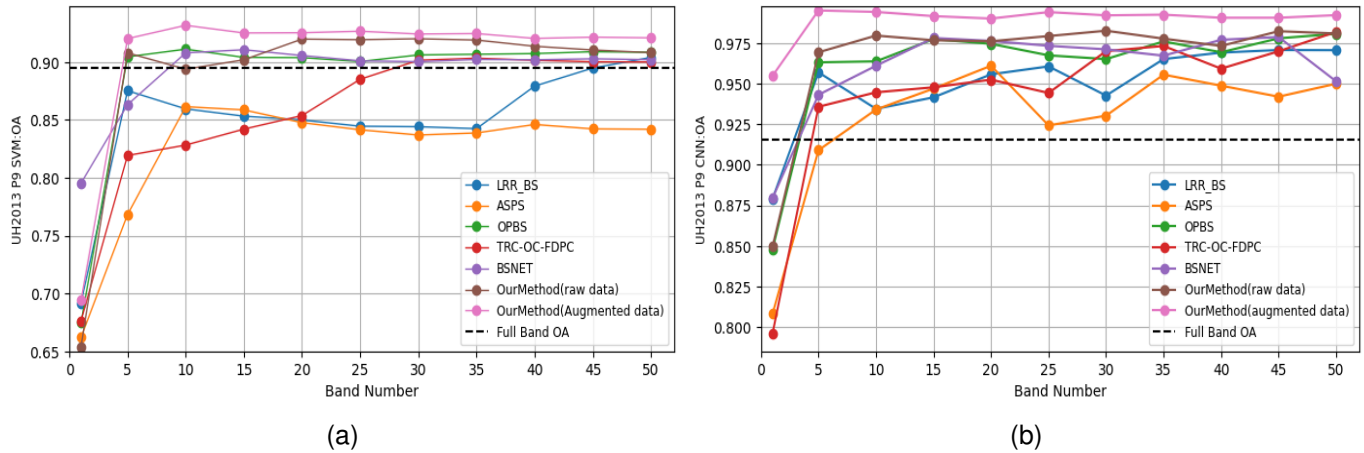


Fig. 8. Classification results on the Houston 2013 data set after fusing selected bands with LiDAR data. (a) OA with SVM; (b) OA with CNN.

TABLE V

CLASSIFICATION PERFORMANCE ON THE HOUSTON 2013 DATA SET (HSI+LiDAR) WITH THE SVM CLASSIFIER. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	LRR_BS [39]	ASPS [56]	OPBS [27]	TRC-OC-FDPC [33]	BSNET [36]	Our Method (Raw Data)	Our Method (Augmented Data)
OA (30 bands)	0.8441	0.8369	0.9061	0.9015	0.9003	0.9201	0.9242
AA (30 bands)	0.8684	0.8584	0.9199	0.9152	0.9152	0.9301	0.9373
Kappa (30 bands)	0.8308	0.8232	0.8981	0.8931	0.8918	0.9133	0.9177
OA (25 bands)	0.8445	0.8414	0.9004	0.8851	0.9010	0.9193	0.9266
AA (25 bands)	0.8695	0.8624	0.9155	0.9002	0.9157	0.9284	0.9392
Kappa (25 bands)	0.8313	0.8281	0.8919	0.8752	0.8925	0.9124	0.9203
OA (20 bands)	0.8499	0.8476	0.9038	0.8537	0.9056	0.9198	0.9253
AA (20 bands)	0.8736	0.7675	0.9173	0.8726	0.9193	0.9283	0.9378
Kappa (20 bands)	0.8371	0.8348	0.8955	0.8412	0.8976	0.9124	0.9189
OA (15 bands)	0.8532	0.8586	0.9040	0.8420	0.9106	0.9019	0.9251
AA (15 bands)	0.8770	0.8763	0.9175	0.8628	0.9231	0.9167	0.9378
Kappa (15 bands)	0.8408	0.8466	0.8958	0.8286	0.9030	0.8935	0.9186
OA (10 bands)	0.8595	0.8614	0.9111	0.8282	0.9079	0.8938	0.9317
AA (10 bands)	0.8822	0.8776	0.9233	0.8515	0.9203	0.9107	0.9427
Kappa (10 bands)	0.8476	0.8496	0.9035	0.8138	0.9000	0.8848	0.9259
OA (5 bands)	0.8752	0.7681	0.9045	0.8192	0.8631	0.9079	0.9202
AA (5 bands)	0.8934	0.7994	0.9171	0.8424	0.8819	0.9210	0.9335
Kappa (5 bands)	0.8646	0.7492	0.8963	0.8041	0.8515	0.9000	0.9134
OA (1 band)	0.6918	0.6619	0.6745	0.6762	0.7952	0.6539	0.6939
AA (1 band)	0.7361	0.7014	0.7214	0.7221	0.8237	0.6983	0.7242
Kappa (1 band)	0.6677	0.6355	0.6495	0.6510	0.7785	0.6269	0.6693

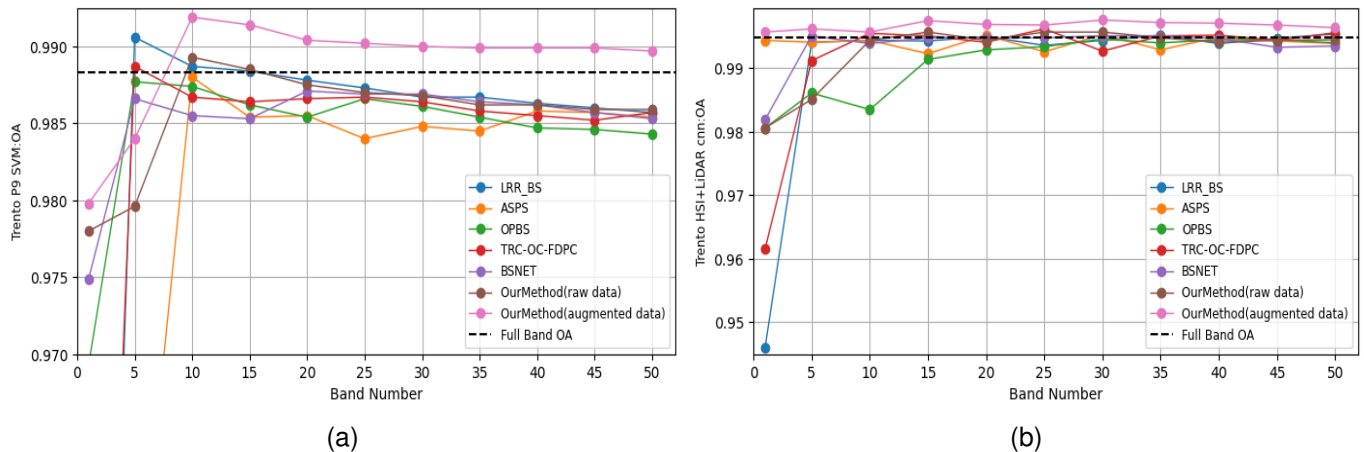


Fig. 9. Classification results on the Trento data set after fusing selected bands with LiDAR data. (a) OA with SVM; (b) OA with CNN

TABLE VI

CLASSIFICATION PERFORMANCE ON THE HOUSTON 2013 DATA SET (HSI+LiDAR) WITH THE CNN CLASSIFIER. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	LRR_BS [39]	ASPS [56]	OPBS [27]	TRC-OC-FDPC [33]	BSNET [36]	Our Method (Raw Data)	Our Method (Augmented Data)
OA (30 bands)	0.9427	0.9303	0.9652	0.9701	0.9712	0.9827	0.9922
AA (30 bands)	0.9521	0.9384	0.9701	0.9704	0.9687	0.9815	0.9940
Kappa (30 bands)	0.9379	0.9244	0.9622	0.9676	0.9688	0.9814	0.9915
OA (25 bands)	0.9607	0.9243	0.9675	0.9443	0.9734	0.9794	0.9942
AA (25 bands)	0.9647	0.9292	0.9696	0.9515	0.9763	0.9794	0.9955
Kappa (25 bands)	0.9574	0.9179	0.9647	0.9395	0.9711	0.9777	0.9938
OA (20 bands)	0.9558	0.9611	0.9746	0.9525	0.9764	0.9762	0.9902
AA (20 bands)	0.9612	0.9655	0.9732	0.9548	0.9755	0.9786	0.9904
Kappa (20 bands)	0.9520	0.9577	0.9724	0.9484	0.9744	0.9742	0.9894
OA (15 bands)	0.9416	0.9470	0.9775	0.9478	0.9782	0.9768	0.9917
AA (15 bands)	0.9513	0.9530	0.9690	0.9537	0.9789	0.9773	0.9935
Kappa (15 bands)	0.9366	0.9425	0.9756	0.9433	0.9763	0.9748	0.9910
OA (10 bands)	0.9345	0.9341	0.9638	0.9447	0.9611	0.9797	0.9943
AA (10 bands)	0.9440	0.9418	0.9670	0.9489	0.9646	0.9795	0.9956
Kappa (10 bands)	0.9289	0.9284	0.9607	0.9399	0.9577	0.9780	0.9939
OA (5 bands)	0.9570	0.9090	0.9632	0.9358	0.9432	0.9694	0.9952
AA (5 bands)	0.9583	0.9157	0.9614	0.9450	0.9512	0.9735	0.9963
Kappa (5 bands)	0.9533	0.9012	0.9600	0.9303	0.9383	0.9668	0.9948
OA (1 band)	0.8791	0.8081	0.8477	0.7959	0.8794	0.8496	0.9549
AA (1 band)	0.8905	0.8352	0.8670	0.8229	0.8965	0.8655	0.9639
Kappa (1 band)	0.8688	0.7919	0.8349	0.7790	0.8692	0.8371	0.9510

TABLE VII

CLASSIFICATION PERFORMANCE ON THE TRENTO DATA SET (HSI+LiDAR) WITH THE SVM CLASSIFIER. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	LRR_BS [39]	ASPS [56]	OPBS [27]	TRC-OC-FDPC [33]	BSNET [36]	Our Method (Raw Data)	Our Method (Augmented Data)
OA (30 bands)	0.9867	0.9848	0.9861	0.9864	0.9869	0.9868	0.9900
AA (30 bands)	0.9786	0.9768	0.9791	0.9793	0.9798	0.9793	0.9862
Kappa(30 bands)	0.9821	0.9795	0.9813	0.9818	0.9824	0.9823	0.9866
OA (25 bands)	0.9873	0.9840	0.9866	0.9867	0.9869	0.9870	0.9902
AA (25 bands)	0.9798	0.9758	0.9802	0.9790	0.9798	0.9795	0.9866
Kappa(25 bands)	0.9819	0.9785	0.9820	0.9822	0.9824	0.9826	0.9869
OA (20 bands)	0.9878	0.9855	0.9854	0.9866	0.9871	0.9875	0.9904
AA (20 bands)	0.9819	0.9783	0.9797	0.9779	0.9802	0.9804	0.9870
Kappa(20 bands)	0.9836	0.9806	0.9804	0.9821	0.9828	0.9833	0.9871
OA (15 bands)	0.9884	0.9854	0.9862	0.9864	0.9853	0.9885	0.9914
AA (15 bands)	0.9817	0.9784	0.9806	0.9783	0.9798	0.9817	0.9881
Kappa(15 bands)	0.9844	0.9804	0.9815	0.9817	0.9803	0.9845	0.9885
OA (10 bands)	0.9887	0.9813	0.9874	0.9804	0.9855	0.9893	0.9919
AA (10 bands)	0.9805	0.9726	0.9816	0.9716	0.9815	0.9849	0.9886
Kappa(10 bands)	0.9848	0.9748	0.9831	0.9737	0.9805	0.9856	0.9888
OA (5 bands)	0.9806	0.9512	0.9877	0.9887	0.9866	0.9796	0.9840
AA (5 bands)	0.9845	0.9316	0.9812	0.9816	0.9817	0.9713	0.9778
Kappa (5 bands)	0.9874	0.9339	0.9834	0.9848	0.9820	0.9726	0.9786
OA (1 band)	0.9018	0.9030	0.9693	0.9150	0.9749	0.9780	0.9798
AA (1 band)	0.7753	0.8946	0.9565	0.9266	0.9631	0.9657	0.9698
Kappa (1 band)	0.8677	0.8712	0.9587	0.8873	0.9663	0.9704	0.9729

are 1, 5 and 10 in most criteria. The results demonstrate a remarkable capability to discern pertinent features from a constrained band set, a crucial attribute in contexts with limited data availability or computational resources. When trained with augmented data, our method performs the best in all cases and presents its exceptional classification capability. This implies that the attention-based mechanism can work effectively with fused hyperspectral bands and LiDAR data. In particular, CNN achieved close to perfect classification results, validating the strength of a deep learning-based approach that integrates attention mechanisms with deep neural networks.

3) *Results on the MUUFL data set* : The results on the MUUFL dataset are shown in Fig. 10, Table IX and Table X.

The evaluation of the SVM classifier based on the MUUFL data set demonstrates that our method with limited training data exhibits close performance to the top performing methods. Meanwhile, when trained on the augmented data, our method consistently exceeds alternative techniques such as ASPS, OPBS, and BSNET across various band counts and with both SVM and CNN classifiers.

The experiments described above show that the use of LiDAR to guide band selection is a key factor in our method. LiDAR provides additional spatial and depth information that can enhance the feature selection process, leading to more accurate and reliable classifications. The implementation of a cross-

TABLE VIII

CLASSIFICATION PERFORMANCE ON THE TRENTO DATA SET (HSI+LiDAR) WITH THE CNN. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	LRR_BS [39]	ASPS [56]	OPBS [27]	TRC-OC-FDPC [33]	BSNET [36]	Our Method (Raw Data)	Our Method (Augmented Data)
OA (30 bands)	0.9944	0.9953	0.9946	0.9927	0.9951	0.9957	0.9976
AA (30 bands)	0.9915	0.9930	0.9911	0.9872	0.9925	0.9927	0.9962
Kappa (30 bands)	0.9926	0.9938	0.9928	0.9902	0.9934	0.9943	0.9969
OA (25 bands)	0.9936	0.9926	0.9934	0.9962	0.9948	0.9957	0.9968
AA (25 bands 25)	0.9906	0.9904	0.9898	0.9908	0.9904	0.9929	0.9948
Kappa (25 bands)	0.9915	0.9902	0.9911	0.9949	0.9931	0.9943	0.9958
OA (20 bands)	0.9950	0.9952	0.9929	0.9941	0.9948	0.9944	0.9969
AA (20 bands)	0.9921	0.9925	0.9884	0.9914	0.9923	0.9914	0.9950
Kappa (20 bands)	0.9933	0.9936	0.9905	0.9922	0.9930	0.9925	0.9959
OA (15 bands)	0.9943	0.9923	0.9914	0.9951	0.9948	0.9957	0.9975
AA (15 bands)	0.9922	0.9899	0.9867	0.9846	0.9922	0.9931	0.9960
Kappa (15 bands)	0.9923	0.9898	0.9884	0.9935	0.9931	0.9943	0.9967
OA (10 bands)	0.9944	0.9944	0.9835	0.9955	0.9939	0.9944	0.9957
AA (10 bands)	0.9900	0.9918	0.9754	0.9901	0.9907	0.9915	0.9934
Kappa (10 bands)	0.9925	0.9926	0.9778	0.9940	0.9919	0.9927	0.9943
OA (5 bands)	0.9941	0.9941	0.9861	0.9912	0.9951	0.9851	0.9962
AA (5 bands)	0.9914	0.9922	0.9862	0.9909	0.9922	0.9736	0.9964
Kappa (5 bands)	0.9921	0.9922	0.9813	0.9885	0.9934	0.9800	0.9950
OA (1 band)	0.9461	0.9944	0.9806	0.9616	0.9819	0.9806	0.9957
AA (1 band)	0.9007	0.9911	0.9710	0.9416	0.9694	0.9659	0.9946
Kappa (1 band)	0.9278	0.9929	0.9740	0.9486	0.9757	0.9740	0.9943

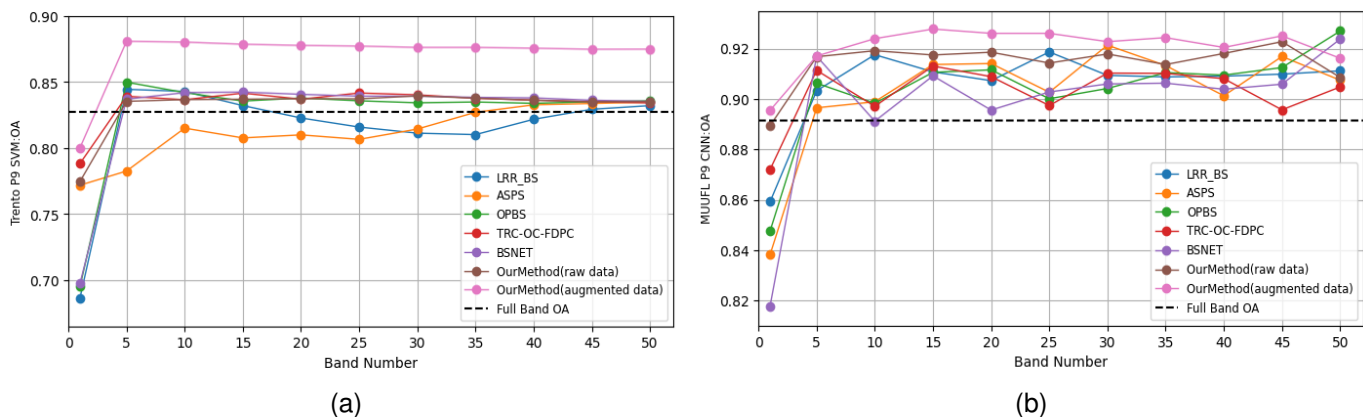


Fig. 10. Classification results on the MUUFL data set after fusing selected bands with LiDAR data. (a) OA with SVM; (b) OA with CNN.

attention mechanism contributes to the method's effectiveness. This mechanism can better capture the relationships between different types of data (HSI and LiDAR in this case), leading to more informative feature selection and, consequently, better classification performance.

The overall MUUFLE classification performance is lower than those of the other two datasets. We believe the reason could be that the MUUFLE dataset is more challenging than the other two datasets. This is also observed in other papers on HSI and LiDAR fusion.

#### E. Comparison with HSI and LiDAR Data Fusion Methods

The fusion models selected for this comparative study include the CNN-HSI model proposed by Yu et al. (2017) [57] and further developed by Mohla et al. (2020) [49], the CoupledCNN model from Hang et al. (2020) [58], FuseAtNet as detailed in Mohla et al. (2020) [49], and a general multimodal deep learning (MDL-RS) framework including the early-fusion,

middle-fusion and late-fusion models explored in Hong et al. (2020) [13]. Additional comparisons are made with EndNet (2020) [45] and the collaborative convolutional learning model (CCL) presented by Jia et al. (2023) [46].

These state-of-the-art fusion models typically operate on full-band HSI hyperspectral image data combined with LiDAR. All experimental comparisons adhere to the same test settings as those delineated for the CCL configuration [46]. Instead, our method uses only 10 HSI bands and fuses them with the LiDAR data. For our model, band selection integrating 1DCNN, and comparative models, analysis was carried out using a uniform patch size of 9 [46].

**1) Results on the Houston 2013 data set:** In the comprehensive analysis of the Houston 2013 dataset, our method has demonstrated exceptional performance, surpassing other fusion methods with a large margin. As highlighted in Table XI, our method has significantly outperformed alternative approaches that use all hyperspectral bands. In particular, our method

TABLE IX

CLASSIFICATION PERFORMANCE ON THE MUUFL DATA SET (HSI+LIDAR) WITH THE SVM CLASSIFIER. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	LRR_BS [39]	ASPS [56]	OPBS [27]	TRC-OC-FDPC [33]	BSNET [36]	Our Method (Raw Data)	Our Method (Augmented Data)
OA (30 bands)	0.8112	0.8143	0.8342	0.8402	0.8392	0.8394	0.8761
AA (30 bands)	0.8522	0.8534	0.8649	0.8752	0.8732	0.8745	0.8993
Kappa (30 bands)	0.7603	0.7639	0.7885	0.7960	0.7948	0.7949	0.8398
OA (25 bands)	0.8158	0.8065	0.8357	0.8416	0.8392	0.8372	0.8771
AA ((25 bands)	0.8539	0.8485	0.8664	0.8766	0.8721	0.8717	0.8994
Kappa ((25 bands5)	0.7660	0.7543	0.7905	0.7977	0.7945	0.7922	0.8411
OA (20 bands)	0.8227	0.8099	0.8378	0.8369	0.8406	0.8373	0.8776
AA (20 bands)	0.8615	0.8503	0.8688	0.8750	0.8716	0.8726	0.8980
Kappa (20 bands)	0.7747	0.7585	0.7930	0.7922	0.7965	0.7923	0.8417
OA (15 bands)	0.8321	0.8076	0.8355	0.8415	0.8423	0.8369	0.8785
AA (15 bands)	0.8717	0.8497	0.8679	0.8773	0.8721	0.8735	0.8991
Kappa (15 bands)	0.7864	0.7558	0.7903	0.7978	0.7986	0.7919	0.8428
OA (10 bands)	0.8428	0.8150	0.8419	0.8365	0.8419	0.8366	0.8801
AA (10 bands)	0.8806	0.8554	0.8755	0.8728	0.8737	0.8729	0.9002
Kappa (10 bands)	0.7995	0.7648	0.7983	0.7916	0.7981	0.7916	0.8448
OA (5 bands)	0.8444	0.7825	0.8498	0.8391	0.8369	0.8353	0.8807
AA (5 bands)	0.8816	0.8336	0.8856	0.8758	0.8702	0.8716	0.9007
Kappa (5 bands)	0.8013	0.7254	0.8082	0.7950	0.7920	0.7900	0.8455
OA (1 band)	0.6864	0.7718	0.6951	0.7882	0.6974	0.7746	0.8001
AA (1 band)	0.7737	0.8264	0.7801	0.8379	0.7799	0.8204	0.8310
Kappa (1 band)	0.6129	0.7119	0.6229	0.7324	0.6254	0.7152	0.7443

TABLE X

CLASSIFICATION PERFORMANCE ON THE MUUFL DATA SET (HSI+LIDAR) WITH THE CNN CLASSIFIER. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	LRR_BS [39]	ASPS [56]	OPBS [27]	TRC-OC-FDPC [33]	BSNET [36]	Our Method (Raw Data)	Our Method (Augmented Train)
OA (30 bands)	0.9094	0.9214	0.9042	0.9103	0.9060	0.9179	0.9228
AA (30 bands)	0.8976	0.9138	0.8879	0.8931	0.8962	0.9297	0.9230
Kappa (30 bands)	0.8814	0.8970	0.8748	0.8642	0.8583	0.8930	0.8993
OA (25 bands)	0.9187	0.9027	0.9001	0.8974	0.9028	0.9143	0.9261
AA (25 bands)	0.9030	0.8973	0.8925	0.8796	0.9051	0.9194	0.9346
Kappa (25 bands)	0.8935	0.8729	0.8704	0.8445	0.8548	0.8888	0.9036
OA (20 bands)	0.9073	0.9141	0.9116	0.9089	0.8956	0.9186	0.9261
AA (20 bands)	0.9081	0.9028	0.9036	0.9020	0.9090	0.9304	0.9346
Kappa (20 bands)	0.8769	0.8876	0.8846	0.8609	0.8443	0.8936	0.9036
OA (15 bands)	0.9108	0.9137	0.9106	0.9131	0.9092	0.9175	0.9278
AA (15 bands)	0.9055	0.8923	0.9057	0.8931	0.9011	0.9238	0.9300
Kappa (15 bands)	0.8834	0.8872	0.8834	0.8678	0.8621	0.8924	0.9054
OA (10 bands)	0.9176	0.8989	0.8983	0.8971	0.8912	0.9192	0.9240
AA (10 bands)	0.9121	0.8875	0.8912	0.8970	0.8828	0.9283	0.9306
Kappa (10 bands)	0.8926	0.8679	0.8673	0.8449	0.8364	0.8947	0.9006
OA (5 bands)	0.9033	0.8965	0.9063	0.9114	0.9170	0.9168	0.9171
AA (5 bands)	0.8956	0.8933	0.9051	0.9142	0.9057	0.9115	0.9266
Kappa (5 bands)	0.8740	0.8652	0.8775	0.8662	0.8739	0.8909	0.8937
OA (1 band)	0.8592	0.8382	0.8476	0.8720	0.8178	0.8891	0.8955
AA (1 band)	0.8618	0.8224	0.8566	0.8842	0.7727	0.9009	0.8839
Kappa (1 band)	0.8164	0.7863	0.8015	0.8099	0.7277	0.8562	0.8432

trained on the augmented data has achieved an impressive Overall Accuracy (OA) of 0.9943, an Average Accuracy (AA) of 0.9956, and a Kappa coefficient of 0.9939. The results underscore the effectiveness and robustness of our method in describing hyperspectral images when using fewer selected bands.

2) *Results on the Trento data set:* As shown in Table XII, the comparative assessment of classification performance on the Trento data set presents that our method with the raw training data has better performance than other methods in terms of AA, and are slightly worse on OA and Kappa. Meanwhile, when trained on the augmented dataset, our method achieved the highest classification accuracy. It leads the scoreboard in

OA, AA and Kappa, demonstrating its superior capability to synthesise and classify hyperspectral data with a reduced set of 10 selected bands.

3) *Results on the MUUFL data set:* The comparative results from the MUUFL dataset in Table XIII clearly illustrate the effectiveness of our band selection and fusion method. Our method with the raw training data showcases impressive accuracy, while with method with the augmented training data emerges as the leader, outperforming all other models. Specifically, our method trained with the augmented dataset has achieved the highest OA of 0.9957, AA of 0.9934, and Kappa of 0.9943, confirming its superior classification capabilities when utilising a curated selection of 10 bands, as opposed to



TABLE XI

CLASSIFICATION PERFORMANCE ON THE HOUSTON 2013 DATA SET. ALL OTHER BANDS USED ALL BANDS IN DATA FUSION, WHILE OUR METHOD USED ONLY 10 SELECTED BANDS. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	CNN-HSI [57]	CoupledCNN [58]	Middle_Fusion [13]	EndNet [45]	FusAtNet [49]	CCL [46]	Our Method (Raw Data)	Our Method (Augmented Data)
OA	0.8123	0.8835	0.8576	0.8098	0.7693	0.9215	0.9797	0.9943
AA	0.8309	0.8989	0.8758	0.8166	0.7887	0.9306	0.9795	0.9956
kappa	0.7973	0.8741	0.8461	0.7944	0.7508	0.9151	0.9780	0.9939

TABLE XII

CLASSIFICATION PERFORMANCE ON THE TRENTO DATA SET. ALL OTHER BANDS USED ALL BANDS IN DATA FUSION, WHILE OUR METHOD USED ONLY 10 SELECTED BANDS. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	CNN-HSI [57]	CoupledCNN [58]	Middle_Fusion [13]	EndNet [45]	FusAtNet [49]	CCL [46]	Our Method (Raw Data)	Our Method (Augmented Data)
OA	0.9474	0.9841	0.9873	0.8694	0.9770	0.9917	0.9949	0.9957
AA	0.9298	0.9695	0.9763	0.8525	0.9555	0.9841	0.9915	0.9934
kappa	0.9299	0.9788	0.9831	0.8255	0.9693	0.9890	0.9927	0.9943

TABLE XIII

CLASSIFICATION PERFORMANCE ON THE MUUFL DATA SET. ALL OTHER BANDS USED ALL BANDS IN DATA FUSION, WHILE OUR METHOD USED ONLY 10 SELECTED BANDS. THE BEST RESULTS ARE HIGHLIGHTED USING BLUE COLOR AND THE SECOND BEST RESULTS ARE HIGHLIGHTED USING RED COLOR.

Metrics	CNN-HSI [57]	CoupledCNN [58]	Middle_Fusion [13]	EndNet [45]	FusAtNet [49]	CCL [46]	Our Method (Raw Data)	Our Method (Augmented Data)
OA	0.7384	0.7743	0.7806	0.7915	0.7076	0.8111	0.9192	0.9240
AA	0.7626	0.7640	0.7697	0.7852	0.6773	0.7800	0.9283	0.9306
Kappa	0.6748	0.7133	0.7217	0.7366	0.6377	0.7562	0.8947	0.9006

the full-band approach of the competing models.

All three experiments demonstrate the mechanism of cross-attention in discerning relevant HSI bands that are correlated with LiDAR. This approach shows the potential to enhance remote sensing applications through intelligent band selection guided by LiDAR data, and further benefit for fusion model development.

## V. CONCLUSIONS

In this paper, we have introduced a method for hyperspectral band selection guided by the LiDAR data. Our method integrates band selection with the fusion model and has demonstrated superior classification performance compared to full-band fusion models. Specifically, it has shown a notable increase in the overall accuracy, average accuracy, and Kappa coefficients across multiple data sets. This indicates a robust capability to discern and use the most informative spectral bands, which are paired with LiDAR, thus optimising the balance between data volume and classification performance. By selecting bands first and then applying the fusion model, our approach presents a viable solution for real-time processing applications where speed and accuracy are critical. Our future research will investigate the potential of transfer learning to generalise the application of band selection with hyperspectral and LiDAR data fusion. We plan to test the model in a real-time processing environment to evaluate its effectiveness and efficiency in operational scenarios.

## REFERENCES

- [1] K. P. Vadrevu, T. Le Toan, S. S. Ray, and C. O. Justice, *Remote Sensing of Agriculture and Land Cover/Land Use Changes in South and Southeast Asian Countries*. Springer, 2022.
- [2] T. Lai, "A review on visual-SLAM: Advancements from geometric modelling to learning-based semantic scene understanding using multi-modal sensor fusion," *Sensors*, vol. 22, no. 19, p. 7265, 2022.
- [3] J. Gao, P. Li, Z. Chen, and J. Zhang, "A survey on deep learning for multimodal data fusion," *Neural Computation*, vol. 32, no. 5, pp. 829–864, 2020.
- [4] W. Yu, L. Gao, H. Huang, Y. Shen, and G. Shen, "HI2D2FNet: Hyperspectral intrinsic image decomposition guided data fusion network for hyperspectral and LiDAR classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–15, 2023.
- [5] W. Sun and Q. Du, "Hyperspectral band selection: A review," *IEEE Geoscience and Remote Sensing Magazine*, vol. 7, no. 2, pp. 118–139, 2019.
- [6] F. Jahan, J. Zhou, M. Awrangjeb, and Y. Gao, "Fusion of hyperspectral and LiDAR data using discriminant correlation analysis for land cover classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 11, no. 10, pp. 3905–3917, 2018.
- [7] S. Khan, S. Nazir, I. García-Magariño, and A. Hussain, "Deep learning-based urban big data fusion in smart cities: Towards traffic monitoring and flow-preserving fusion," *Computers & Electrical Engineering*, vol. 89, p. 106906, 2021.
- [8] S. Falahatnejad and A. Karami, "Deep fusion of hyperspectral and LiDAR images using attention-based CNN," *SN Computer Science*, vol. 4, no. 1, p. 1, 2022.
- [9] Z. Wang, D. Ziou, C. Armenakis, D. Li, and Q. Li, "A comparative analysis of image fusion methods," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 6, pp. 1391–1402, 2005.
- [10] H. Kaur, D. Koundal, and V. Kadyan, "Image fusion techniques: a survey," *Archives of Computational Methods in Engineering*, vol. 28, pp. 4425–4447, 2021.

- [11] L. Liu, C.-f. Li, Y.-m. Lei, J.-y. Yin, and J.-j. Zhao, "Feature extraction for hyperspectral remote sensing image using weighted PCA-ICA," *Arabian Journal of Geosciences*, vol. 10, pp. 1–10, 2017.
- [12] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 57, no. 9, pp. 6690–6709, 2019.
- [13] D. Hong, L. Gao, N. Yokoya, J. Yao, J. Chanussot, Q. Du, and B. Zhang, "More diverse means better: Multimodal deep learning meets remote-sensing imagery classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 4340–4354, 2020.
- [14] Z. Xue, X. Tan, X. Yu, B. Liu, A. Yu, and P. Zhang, "Deep hierarchical vision transformer for hyperspectral and LiDAR data classification," *IEEE Transactions on Image Processing*, vol. 31, pp. 3095–3110, 2022.
- [15] J. Zhang, Z. Meng, F. Zhao, H. Liu, and Z. Chang, "Convolution transformer mixer for hyperspectral image classification," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2022.
- [16] C. Debes, A. Merentitis, R. Heremans, J. Hahn, N. Frangiadakis, T. van Kasteren, W. Liao, R. Bellens, A. Pizurica, S. Gautama, W. Philips, S. Prasad, Q. Du, and F. Pacifici, "Hyperspectral and LiDAR data fusion: Outcome of the 2013 GRSS data fusion contest," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 6, pp. 2405–2418, 2014.
- [17] University of Trento, "Theses of the university of trento," [Data set], 2022, original work published 2020. [Online]. Available: <http://data.europa.eu/88u/dataset/theses-of-the-university-of-trento>
- [18] X. Du and A. Zare, "Technical report: Scene label ground truth map for MUUFL Gulfport data set," University of Florida, Gainesville, FL, Tech. Rep. 20170417, Apr. 2017, available: <http://ufdc.ufl.edu/IR00009711/00001>.
- [19] A. M. Kim, F. A. Kruse, and R. C. Olsen, "Integrated analysis of light detection and ranging (LiDAR) and hyperspectral imagery (HSI) data," in *Laser Radar Technology and Applications XXI*, vol. 9832. SPIE, 2016, pp. 271–285.
- [20] C. V. Poulton, M. J. Byrd, P. Russo, E. Timurdogan, M. Khandaker, D. Vermeulen, and M. R. Watts, "Long-range lidar and free-space data communication with high-performance optical phased arrays," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 25, no. 5, pp. 1–8, 2019.
- [21] F. Qamar and G. Dobler, "Atmospheric correction of vegetation reflectance with simulation-trained deep learning for ground-based hyperspectral remote sensing," *Plant Methods*, vol. 19, no. 1, p. 74, 2023.
- [22] S. S. Sawant, P. Manoharan, and A. Loganathan, "Band selection strategies for hyperspectral image classification based on machine learning and artificial intelligent techniques—survey," *Arabian Journal of Geosciences*, vol. 14, pp. 1–10, 2021.
- [23] J.-H. Kim, J. Kim, Y. Yang, S. Kim, and H. S. Kim, "Covariance-based band selection and its application to near-real-time hyperspectral target detection," *Optical Engineering*, vol. 56, no. 5, pp. 053 101–053 101, 2017.
- [24] X. Sun, H. Zhang, F. Xu, Y. Zhu, and X. Fu, "Constrained-target band selection with subspace partition for hyperspectral target detection," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 9147–9161, 2021.
- [25] B. Xu, X. Li, W. Hou, Y. Wang, and Y. Wei, "A similarity-based ranking method for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9585–9599, 2021.
- [26] Z. Zhang, D. Wang, X. Sun, L. Zhuang, R. Liu, and L. Ni, "Spatial sampling and grouping information entropy strategy based on kernel fuzzy C-Means clustering method for hyperspectral band selection," *Remote Sensing*, vol. 14, no. 19, p. 5058, 2022.
- [27] W. Zhang, X. Li, Y. Dou, and L. Zhao, "A geometry-based band selection approach for hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4318–4333, 2018.
- [28] J. Wang, C. Tang, X. Liu, W. Zhang, W. Li, X. Zhu, L. Wang, and A. Y. Zomaya, "Region-aware hierarchical latent feature representation learning-guided clustering for hyperspectral band selection," *IEEE Transactions on Cybernetics*, 2022.
- [29] X. Cao, B. Wu, D. Tao, and L. Jiao, "Automatic band selection using spatial-structure information and classifier-based clustering," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 9, pp. 4352–4360, 2016.
- [30] J. Wang, C. Tang, Z. Li, X. Liu, W. Zhang, E. Zhu, and L. Wang, "Hyperspectral band selection via region-aware latent features fusion based clustering," *Information Fusion*, vol. 79, pp. 162–173, 2022.
- [31] A. Tahraoui, R. Kheddad, A. Bouakache, and A. Belhadj-Aissa, "Affinity propagation for unsupervised classification of remotely sensed images," in *2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*. IEEE, 2017, pp. 1–6.
- [32] Y. Cai, Z. Zhang, Z. Cai, X. Liu, X. Jiang, and Q. Yan, "Graph convolutional subspace clustering: A robust subspace clustering framework for hyperspectral image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 5, pp. 4191–4202, 2020.
- [33] Q. Wang, F. Zhang, and X. Li, "Optimal clustering framework for hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 10, pp. 5910–5922, 2018.
- [34] M. Zhang, J. Ma, and M. Gong, "Unsupervised hyperspectral band selection by fuzzy clustering with particle swarm optimization," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 773–777, 2017.
- [35] J. Wang, C. Tang, X. Zheng, X. Liu, W. Zhang, and E. Zhu, "Graph regularized spatial-spectral subspace clustering for hyperspectral band selection," *Neural Networks*, vol. 153, pp. 292–302, 2022.
- [36] Y. Cai, X. Liu, and Z. Cai, "Bs-nets: An end-to-end framework for band selection of hyperspectral image," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 3, pp. 1969–1984, 2019.
- [37] L. Mou, S. Saha, Y. Hua, F. Bovolo, L. Bruzzone, and X. X. Zhu, "Deep reinforcement learning for band selection in hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–14, 2022.
- [38] C. O. Ayna, R. Mdrafi, Q. Du, and A. C. Gurbuz, "Learning-based optimization of hyperspectral band selection for classification," *Remote Sensing*, vol. 15, no. 18, p. 4460, 2023.
- [39] C. Yu, S. Zhou, M. Song, and C.-I. Chang, "Semisupervised hyperspectral band selection based on dual-constrained low-rank representation," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2021.
- [40] H. Zhao, L. Bruzzone, R. Guan, F. Zhou, and C. Yang, "Spectral-spatial genetic algorithm-based unsupervised band selection for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 59, no. 11, pp. 9616–9632, 2021.
- [41] Y.-L. Chang, J.-N. Liu, Y.-L. Chen, W.-Y. Chang, T.-J. Hsieh, and B. Huang, "Hyperspectral band selection based on parallel particle swarm optimization and impurity function band prioritization schemes," *Journal of Applied Remote Sensing*, vol. 8, no. 1, pp. 084 798–084 798, 2014.
- [42] Q. Wang, J. Lin, and Y. Yuan, "Salient band selection for hyperspectral image classification via manifold ranking," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 6, pp. 1279–1289, 2016.
- [43] F. Jahan, J. Zhou, M. Awrangjeb, and Y. Gao, "Inverse coefficient of variation feature and multilevel fusion technique for hyperspectral and LiDAR data classification," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 367–381, 2020.
- [44] R. Hang, Z. Li, P. Ghamisi, D. Hong, G. Xia, and Q. Liu, "Classification of hyperspectral and LiDAR data using coupled CNNs," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 7, pp. 4939–4950, 2020.
- [45] D. Hong, L. Gao, R. Hang, B. Zhang, and J. Chanussot, "Deep encoder-decoder networks for classification of hyperspectral and LiDAR data," *IEEE Geoscience and Remote Sensing Letters*, vol. 19, pp. 1–5, 2020.
- [46] S. Jia, X. Zhou, S. Jiang, and R. He, "Collaborative contrastive learning for hyperspectral and LiDAR classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 61, pp. 1–14, 2023.
- [47] W. Liao, F. Van Coillie, L. Gao, L. Li, B. Zhang, and J. Chanussot, "Deep learning for fusion of APEX hyperspectral and full-waveform LiDAR remote sensing data for tree species mapping," *IEEE Access*, vol. 6, pp. 68 716–68 729, 2018.
- [48] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22–40, 2016.
- [49] S. Mohla, S. Pande, B. Banerjee, and S. Chaudhuri, "FusAtNet: Dual attention based spectrospatial multimodal fusion network for hyperspectral and LiDAR classification," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2020, pp. 92–93.
- [50] L. Zhou, J. Geng, and W. Jiang, "Joint classification of hyperspectral and LiDAR data based on position-channel cooperative attention network," *Remote Sensing*, vol. 14, no. 14, p. 3247, 2022.
- [51] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, E. Kaiser, and I. Polosukhin, "Attention is all you need," *Advances in Neural Information Processing Systems*, vol. 30, 2017.
- [52] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly *et al.*,

“An image is worth 16x16 words: Transformers for image recognition at scale,” *arXiv preprint arXiv:2010.11929*, 2020.

- [53] D. Hong, Z. Han, J. Yao, L. Gao, B. Zhang, A. Plaza, and J. Chanussot, “SpectralFormer: Rethinking hyperspectral image classification with transformers,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2021.
- [54] D. Hong, B. Zhang, X. Li, Y. Li, C. Li, J. Yao, N. Yokoya, H. Li, X. Jia, A. Plaza *et al.*, “Spectralgpt: Spectral foundation model,” *arXiv preprint arXiv:2311.07113*, 2023.
- [55] X. Zhou, M. Cui, A. Singhanian, and D. Fernández Díaz, Juan Carlos, “Houston 2013 data set,” National Center for Airborne Laser Mapping (NCALM), University of Houston, Jun. 2013, dataset collected over the University of Houston campus and pre-processed with the assistance of the authors. [Online]. Available: [https://hyperspectral.ee.uh.edu/?page\\_id=459](https://hyperspectral.ee.uh.edu/?page_id=459)
- [56] Q. Wang, Q. Li, and X. Li, “Hyperspectral band selection via adaptive subspace partition strategy,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 12, pp. 4940–4950, 2019.
- [57] S. Yu, S. Jia, and C. Xu, “Convolutional neural networks for hyperspectral image classification,” *Neurocomputing*, vol. 219, pp. 88–98, 2017.
- [58] R. Hang, Z. Li, P. Ghamisi, D. Hong, G. Xia, and Q. Liu, “Classification of hyperspectral and LiDAR data using coupled CNNs,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 58, no. 7, pp. 4939–4950, 2020.



**Judy X YANG** (Student Member, IEEE) worked in manufacturing engineering and supply chain management for more than 15 years as an engineer. Her work history includes in an American manufacturing company PennEngineering in Shanghai as a mechanical project engineer and MSI Electrical Company as an OEM product manager. She has been a member of Engineers Australia since 2015. She completed her Master of Informatics at Central Queensland University, Australia, in 2021. She is working toward a Ph.D. at Griffith University, Nathan, QLD, Australia.

Her research interests include data analysis, Machine Learning, Computer Vision, Hyperspectral image analysis, LiDAR analysis, and multisource remote sensing data fusion.



**Jun ZHOU** Jun Zhou (Senior Member, IEEE) received a B.S. degree in computer science and a B.E. degree in international business from Nanjing University of Science and Technology, Nanjing, China, in 1996 and 1998, respectively, the M.S. degree in computer science from Concordia University, Montreal, QC, Canada, in 2002, and the Ph.D. degree in computer science from the University of Alberta, Edmonton, AB, Canada, in 2006. He is currently a Professor with the School of Information and Communication Technology, Griffith University,

Nathan, QLD, Australia. Previously, he had been a Research Fellow with the Research School of Computer Science, Australian National University, Canberra, ACT, Australia, and a Researcher with the Canberra Research Laboratory, National ICT Australia, Sydney, NSW, Australia. His research interests include pattern recognition, computer vision, and spectral imaging with applications in remote sensing and environmental informatics.



learning, computer vision, hyperspectral image analysis, and remote sensing.

**Jing, WANG** Jing Wang received the bachelor’s degree in communication engineering from the Tianjin University of Science and Technology, Tianjin, China, in 2010 and the master’s degree in signal and information processing from the Chinese Academy of Sciences of the University of Beijing, China, in 2013. and in 2020, received the Ph.D. degree from Griffith University, Nathan, QLD, Australia and the Chinese Academy of Sciences University. He is now a scientist at the Agriculture Department. His research interests include pattern recognition and machine



including service as associate editor of SCI-indexed journals and program chair/committee member of international conferences. Assoc Prof. Tian is a senior member of IEEE.

**Hui TIAN** Assoc Prof. Hui Tian is Discipline Head of Computer Science in the School of Information and Communication Technology, Griffith, Australia. She received the PhD degree in Computer Science from Japan Advanced Institute of Science and Technology. Her main research interests include Network Routing and Tomography, Privacy-preserving Computing and Knowledge Discovery. She has published one book and more than 80 research papers and led several research projects in different countries. She has actively engaged in professional activities



informatics stream in IIIS, and is engaging with the Gold Coast University Hospital and Queensland Health on several collaborative projects. He has also engaged actively in professional activities such as on the organizing committee or technical program committee of many conferences, on editorial boards, as assessor for nationally competitive research grants, and reviewers for many international conferences and journals. He is currently serving as an associate editor of IEEE Trans. Fuzzy Systems, Springer Nature Computer Science, and International Journal of Computational Intelligence Systems, Machine Intelligence Research. He is a senior member of IEEE since 2005.

**Alan Wee-Chung Liew** PROFESSOR ALAN LIEW is currently the Head of School, the School of Information Communication Technology, and the Deputy Director of the Institute for Integrated Intelligent Systems (IIIS), at Griffith University, Australia. Prior to joining Griffith University in 2007, he was an Assistant Professor at the Department of Computer Science and Engineering, Chinese University of Hong Kong. His research interest is in the field of AI, machine learning, medical imaging, computer vision, and bioinformatics. He is leading the medical/health