



OPEN HSICNet a novel deep learning architecture for hyperspectral image classification in remote sensing and environmental monitoring

K. Purnachand¹✉, Rafath Samrin², John Babu Guttikonda³, Rajitha Kotoju⁴, Phani Sridhar Addepalli⁵ & A. Mohan⁶

Leveraging the rich spectral and spatial information, hyperspectral image classification (HSIC) plays a vital role in remote sensing, which is significant for land-cover mapping and environmental monitoring. However, hyperspectral images exhibit high dimensionality, significant spectral redundancy, and a limited number of annotated samples, making classification challenging. However, the original data may be complicated by more specific spectral-spatial interdependencies, which the so-called feature extraction in either standard film learning or early CNN-based models doesn't capture. Simultaneously, recent approaches that embed attention or a transformer into the architecture suffer from high computational cost, overfitting, and scalability issues. These limitations highlight the need for a new, flexible, and computationally efficient deep-learning framework for various HSIC scenarios. We propose HSICNet, a unique dual-branch network architecture for deep learning that separately captures spectral and spatial features, followed by an attention-based feature fusion strategy to facilitate the interplay between low-, mid, and high-level feature representations. To alleviate the curse of dimensionality and redundancy, we also include a PCA-based dimensionality reduction module. The work involves optimising the proposed model for accuracy, computational efficiency, and generalisation across all classes. HSICNet is extensively evaluated by conducting experiments on three benchmark hyperspectral datasets, including Indian Pines, Pavia University, and Salinas, and its performance is superior to the current state-of-the-art algorithms. With overall accuracy up to 99.35%, it achieves statistically significant improvements in the Kappa coefficient, F1-score, and average accuracy by a large margin. Ablations verify our choice of a dual-branch design and attention fusion to improve classification scores. With a strong architecture and the advantage of HSICNet's lightweight nature, it has potential for use in real-time, scalability-sensitive remote sensing applications such as precision agriculture, smart cities, and environmental monitoring. Its robust generalizability and interpretability also permit deployment in heterogeneous and adaptive operational settings.

Keywords Hyperspectral image classification, Deep learning, Attention mechanism, Spectral-spatial fusion, Remote sensing

HSIC (Hyperspectral Image Classification) is a crucial technique for remote sensing applications, including precision agriculture, environmental monitoring, and land-use analysis. Unlike conventional RGB imagery, hyperspectral images collect a large number of spectral bands. They can gather rich spectral-spatial information, allowing for more sensitive and detailed classification of land-cover types. However, the high dimensionality,

¹Co-Founder and CEO, Sentiantal Private Limited, Hyderabad,, Telangana, India. ²Department of Computer Science, College of Computer Science, King Khalid University, 6142 Asir-Abha, Saudi Arabia. ³Department of Computer Science & Engineering (AI & ML), Anurag Engineering College, Kodad, Telangana 508206, India. ⁴Department of Computer Science and Engineering, Mahatma Gandhi, Institute of Technology, (MGIT), Hyderabad 500075, India. ⁵Aditya University Surampalem, Surampalem, Andhra Pradesh 533437, India. ⁶Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology (A), Hyderabad, India. ✉email: purnachand.k@gmail.com

band redundancy, and limited availability of labelled samples pose challenges to accurate imaging classification. Most traditional machine learning and shallow neural networks fail to extract useful features from this data type; therefore, there is a need for deep learning architectures that can efficiently learn deep correlations in both spectral and spatial domains.

Recent improvements in deep learning techniques, particularly Convolutional Neural Networks (CNN) and their 3D counterpart, have been employed to enhance HSIC accuracy^{1,2}. Attention mechanisms³, hybrid CNN-transformer models⁴, and multi-branch architectures⁵ have also been incorporated in other works to further improve spectral-spatial feature learning. However, these methods still have problems such as overfitting, high computational complexity, and improper fusion of heterogeneous features. Additionally, most of these models require substantial amounts of labelled data or do not generalise well across different hyperspectral scenes^{6,7}.

Given the above gaps, this work presents a new deep learning framework, HSICNet, that leverages a dual-branch spectral and spatial feature extraction mechanism, attention-guided data fusion, and PCA-based dimensionality reduction. It aims to build an effective, explainable, and generalizable model for HSIC tasks. Lighter than transformer-heavy models yet substantially more accurate, better balanced, and less computationally intensive, this is HSICNet.

The key contributions of this study include: (1) a dual-branch design that enables parallel and efficient extraction of spectral and spatial features, (2) an attention-guided fusion mechanism that enhances inter-feature interaction while minimizing redundancy, (3) a PCA-integrated dimensionality reduction process to address the curse of dimensionality, and (4) comprehensive validation on three benchmark datasets—Indian Pines, Pavia University, and Salinas—demonstrating superior performance over state-of-the-art models.

Our HSICNet framework differs from existing dual-branch and attention-based models, thanks to its simple yet highly representative fusion mechanism. While most previous works stack computationally expensive 3D-CNNs or transformer-based modules, HSICNet decouples spectral and spatial processing by using 1D and 2D convolutional branches to process spectral and spatial information with minimal computational burden, respectively, and fuses them via an attention-guided fusion module. This combined design is a version of squeeze-excitation that retains interpretability and can decrease the model's complexity. Additionally, HSICNet is specifically designed to operate in low-label settings or with limited compute resources by introducing PCA-based dimensionality reduction and using a relatively small number of trainable parameters.

The HSICNet is specifically designed for environmental monitoring applications that require acceptable spectral-spatial discrimination performance, such as vegetation monitoring, land cover mapping, and urban land-use classification. We aim to demonstrate the practical relevance of our approach to remote sensing challenges by incorporating accepted benchmark datasets that simulate real-world scenarios.

This is done through HSICNet, a new lightweight dual-branch framework for parallel spectral and spatial feature extraction with channel-wise attention-guided fusion, which helps efficient and discriminative hyperspectral image classification.

While traditional frameworks typically consist of sequential spectral-spatial processing pipelines, HSICNet breaks this rule and employs a parallel strategy with a learnable attention mechanism for adaptive fusion, enabling the model to learn complementary information more efficiently and at lower computational cost—ideal for deployment on edge platforms.

The rest of this paper is structured as follows. In Sect. “[Related work](#)”, we provide an extensive literature review by grouping previous work into four major categories: CNN-based models, attention-based structures, hybrid models, and application-oriented studies. In Sect. “[Proposed framework](#)”, we present the proposed methodology, including the system architecture, model components, and mathematical formulas. Experimental results are reported in Sect. “[Experimental results](#)”, which details our dataset, evaluation metrics, ablation studies, and comparisons to baseline models. In Sect. “[Discussion](#)”, we conclude the paper with a discussion of the research's importance and the limitations of the current study. Finally, Sect. “[Conclusion and future work](#)” summarises the paper and presents directions for future research.

Related work

In this section, we present a detailed summary of recent progress in HSI classification using deep learning. Existing Convolutional models, Attention-based architectures, Hybrid and multi-modal frameworks, Lightweight and low-data methods, Application-specific research. In the Discussion section, we summarise the significant contributions, limitations, and research gaps, which we hope justify the need for the proposed HSICNet framework.

Deep learning models for hyperspectral image classification

The development of deep learning models (especially CNNs and their variants) for hyperspectral image classification is discussed in this section. Various degree methods were proposed, e.g. 3D CNN, residual networks, and hybrid architectures that use spatial and spectral features for classification. Guerri et al.¹ specifically reported the advancements of deep learning and hyperspectral imaging for agriculture, and the complexity of the data and the problems. Ullah et al. Zhu et al.² reviewed deep learning methods in HSIC and pointed out existing challenges and future development. Hameed et al. To enhance precision and efficiency³, proposed the BPVAM-PCA method and developed a hybrid approach for hyperspectral image classification. Future work would focus on combating computational complexity and hyperparameter sensitivity, deep feature extraction, and GPU parallelisation. Ashraf et al. A high-speed 3D UNet for hyperspectral image classification⁶. Next, we will develop accurate real-time models and include transformers. However, there are two significant downsides: high computational load and reliance on labelled data.

Chhapariya et al.⁵ presented DSSpRAN, an accurate, low-overfitting CNN-based model for hyperspectral image categorisation. Improving feature learning and efficiency is a task for future development. Sun et al.⁹

demonstrate that, with fewer parameters and lower computational overhead, the LCTCS model improves the categorisation of hyperspectral images. Upcoming research will focus on enhancing adaptive convolutional and attentional processes. Jia et al.¹⁰ considered parameter trade-offs when evaluating different AI models for categorising hyperspectral images. Upcoming research will examine uses and improve model effectiveness. Bai et al.¹¹ presented a Content-driven Spectrum Complementary Network that combines derived and spectral data for HSI classification. SpectralGPT integration will be used in future studies to enhance feature learning.

Viel et al. When compared with CNN, LSTM, and Transformer architectures for hyperspectral image classification¹², 1D-CNN was found to be suitable with limited resources. Future development will investigate upgrades to Transformers and hardware acceleration. Limitations include, for example, class imbalance and limited computing resources. Ashraf et al. Manifold learning has also been integrated into CNNs for feature selection in image classification; a recently developed Triple Layered Convolutional Architecture (TLCA) has been proposed¹³. Auto-encoders will be explored in the future for further dimension reduction and to address overfitting. Ji et al. To address the overfitting problem that arises from classifying only a few labelled samples and unstructured data¹⁴, developed a 3-D CNN model. Future research will aim to enhance the accuracy and regularisation of the model.

Zhu et al. One representative work is SS-ConvNeXt, an enhanced close-up convolutional model developed by¹⁵ to ameliorate noise and improve accuracy in hyperspectral image classification. Future work should address the effectiveness of this model and the resource constraints it imposes. Esmaili et al. CNNeGA⁷ is a band selection approach that utilises a 3D-CNN integrated with a GA to significantly improve the accuracy and efficiency of hyperspectral image analysis. In future work, we will focus on the use of spiking neural networks for unsupervised band selection and classification. Noshiri et al. The strengths and weaknesses of using 3D-CNNs in agricultural hyperspectral image classification were discussed in¹⁶. New projects will enhance the effectiveness, affordability, and accessibility of the models—addressing limitations such as the lack of data, high processing needs, and Ullah et al. We leveraged a wavelet-based deep-shot-oriented hyperspectral classification method and utilised 2D CNNs to¹⁷ enhance classification precision. The following steps involve extending applications, modifying tasks, and integrating others—these come with limitations, such as computational requirements and limited current practical use.

Vasanthakumari et al.¹⁸ achieved higher accuracy in multispectral image categorisation by using a CNN and the Modified ReLU activation function. Future research will examine other optimisers and band configurations. One limitation is that it requires more computing power than simpler models. Okwuashi et al.¹⁹ presented a Deep Support Vector Machine (DSVM) that outperforms existing techniques for categorising hyperspectral images. Upcoming projects will focus on solving data issues and improving model efficiency. Ortac et al.²⁰ analysed 1D, 2D, and 3D CNNs and determined that 3D CNNs are the most efficient at classifying hyperspectral images. Upcoming projects will examine CNN performance with different training settings and supplement data. Zhu et al.²¹ proposed RSSAN, which improves spectral and spatial feature selection for better hyperspectral image categorisation. Subsequent research endeavours might enhance attention modules and optimise the handling of high-dimensional data. Ghaderizadeh et al.²² developed a hybrid 2D-3D CNN to address overfitting and sample scarcity in hyperspectral image classification; further investigation into deeper models is warranted.

Attention mechanisms and transformer-based architectures

Attention mechanisms and transformer-based models have gained attention for their ability to capture complex dependencies in hyperspectral image data. This section examines how self-attention, spectral attention, and hybrid CNN-Transformer models have been used to enhance classification accuracy. Lin et al.²³ provided effective forest-type categorisation with up to 100% accuracy using hyperspectral imaging and deep learning (VGG19 + ResNet50). Further optimisation in various forest conditions and data formats is limited, and future studies may focus on system integration for real-time forest management. Lu et al.²⁴ presented a multitask deep learning method that uses spectral techniques for effective screening of hyperaccumulators and for measuring heavy metals. Future research will focus on extending the method to various plant applications, including real-time remote sensing. Among the drawbacks is the requirement for more varied spectrum data. Reddy et al.³ proposed a 3D CNN with self-attention, achieving higher accuracy in hyperspectral image categorisation. In the future, research could examine domain-specific integration and data fusion. Dash et al.²⁵ evaluated CNN and LSTM for the classification of hyperspectral images and found that LSTM with MNF is the most efficient. In the future, generality will be improved.

Olisah et al.²⁶ developed a multi-input CNN model with high accuracy for blackberry ripeness classification. Future research will focus on increasing the number of categorisation tasks and mitigating environmental unpredictability. Wang et al.²⁷ proposed SSGRN for spectral and spatial graph-reasoning-based hyperspectral image categorisation. To lessen noise, further research will improve the production of descriptors. Yu et al.²⁸ developed ConGCN to enhance feature representation by integrating spectral and spatial information through contrastive learning for HSI classification. Future research will focus on addressing data scarcity and enhancing graph augmentation. Farooque et al.⁴ proposed MACLST for HSI classification, which combines lightweight Swin Transformers with 3D atrous convolution to improve accuracy. Upcoming projects involve developing a lighter, more durable transformer to enhance effectiveness. The need for further model improvement and greater resilience in the processing of HSI data is among the limitations.

Sun et al.²⁹ presented HSSAM, a hyperspectral image classification model that combines attention processes with 3-D residual-dense convolution. Subsequent efforts will involve refining the network to achieve higher accuracy and lower complexity. The lengthier operating time compared to some models and increased computing complexity are limitations. Zhang et al.³⁰ proposed SDEnet, a domain-generalisation framework for cross-scene HSI classification that combines contrastive and generative adversarial learning. Subsequent efforts will focus on expanding domain coverage and managing a broader range of target domains. One limitation is

that different domain shifts can make it challenging to achieve optimal performance. Zheng et al.³¹ improved performance by developing RIAN, a rotation-invariant HSI classification method. Future research will address existing limitations and focus on improving the capture of textural features. Xue et al.³² presented HyT-NAS, a system that combines a Transformer and NAS to improve HSI categorisation. The goal of future research will be to develop more efficient whole-Transformer designs. One limitation is that local characteristics are currently obtained using CNN-based methods. Huang et al.³³ utilised spectral attention and self-supervised learning for HSI classification. Subsequent research might improve the distinction between features. One drawback is the dependence on spectral band optimisation. He et al.³⁴ presented HyperViTGAN, a combination of GAN and transformer, for improved HSI classification with limited data. Enhancing generalisation and stability is a future task. Managing class imbalance and GAN instability are limitations.

Ahmad et al. AfNet by Zhao et al.³⁵ was designed to account for attention processes and combines 3D and 2D CNNs for better HSI classification. Attention and hybrid processes will be optimised, but a short-term constraint will be high computing costs. Dong et al. The WFCG model proposed by³⁶ combines GAT and CNN to utilise their complementary benefits for HSI classification. The following steps optimise hyperparameters and make the models even more efficient. Limitations include small sample sizes and high computational costs. Feng et al. RS-AMCNN was proposed by³⁷ to incorporate branch attention and adaptive spatial windows into feature extraction to obtain better features for HSI classification. At the same time, efficiency improvements and further investigation of this method on other datasets are needed as future work. Safari et al. proposed a deep learning-based method for HSI classification, achieving an accuracy of 66.73%³⁸. Sensitivity concerns and more powerful normalisation methods will be studied in future research. Shen et al. Long-range context with practical nonlocal modules via ENL-FCN for HSI classification³⁹; further work will optimise and verify its effectiveness. Mou et al. An original reward structure for deep reinforcement learning to improve accuracy and efficiency in unsupervised hyperspectral band selection⁴⁰.

Hybrid and multi-modal architectures

Multi-modal and hybrid deep learning architectures include external data sources, e.g., combining LiDAR and UAV images, treating class labels/label maps produced from UAV images as outputs to be obtained, or using GANs to improve the input spectral quality of pictures in hyperspectral image classification. This section provides a summary of methods using these modalities that could enhance performance in remote sensing and precision agriculture. Ma et al. Using U-SLIC and CBAM, for example⁴¹, proposed a CNN-based framework to combine UAV HSI and LiDAR data for more accurate tree species classification. In the future, we will explore self-supervised learning and apply the framework to more domains. The limitations include challenges of data acquisition and image noise. Paheding et al. Deep learning methods for remote sensing image categorisation have been presented in recent years, with a focus on applications, trends, and perspectives, while identifying current limitations. Ahmed et al. Deep-learning methods have been shown to reconstruct hyperspectral images from RGB data to evaluate agricultural quality⁴², with each HRNet performing best in this case. Future studies will explore these issues more with respect to generalisation and metamerism. Ahmad et al. To minimise tuning and computation⁴³, proposed Sharpened Cosine Similarity (SCS), which is used here for hyperspectral image categorisation. Other improvements and practical applications will be considered in future research. Giakoumoglou et al. Table 6: Deep learning model developed by⁴⁴ for the identification of *Botrytis cinerea* in cucumber using multispectral images. Future research will explore its potential to treat other plant diseases. Yao et al. proposed a deep hybrid graph network (DHMG) for hyperspectral image noise reduction and classification improvement in⁴⁵. In this study, we also used supervised models due to the data availability.

Amoako et al. recently proposed a meta-RL model for HSI classification to achieve higher accuracy and efficiency by combining deep Q-learning with capsule networks⁴⁶. Future research will focus on reducing computational complexity and improving the model's reproducibility across other BMI datasets. Farmonov et al. Using DESIS hyperspectral images⁴⁷, used a Wavelet-attention CNN to classify crops effectively—future research will focus on improving accuracy and mobilising this in practice. One limitation is that DESIS data, which relies on current spectral information, must be readily available.

Sellami et al. For instance⁴⁸, proposed a semi-supervised hypergraph convolutional network that classifies hyperspectral images via unsupervised feature selection. Future work will focus on scaling the model and adapting it to different datasets. For more widespread use, more validation and optimisation are required, which is a limitation. Wang et al. presented a hyper-kernel-based NAS method in⁴⁹, which can reduce search cost and increase model flexibility for hyperspectral image classification. Then, iterating on the technique for wider applications.” A drawback is that many 3-D convolutional designs may be complex to implement. Pathak et al. proposed a classification approach based on SVM that combines spectral and spatial information⁵⁰, further enhancing performance. Future work would involve parameter optimisation and increased applications. There are limitations, such as the need for improved feature extraction and tuning. Yu et al. Examples include silicon-based EEG sensors that are flexible and non-irritating, and that use soft materials for signal detection, developed by⁵¹. Further research can improve sensor design for broader applications.

Jena et al.⁵² show that there are considerable biases and inconsistencies in this application and that HDL performance is significantly better across 127 image classification studies. Wang et al. While Cubic-CNN⁵³ improves classification performance on HSI and reduces training time, its applicability will be further investigated through low-sampling-rate and noise-optimisation. Hong et al. In a review paper on nonconvex models for hyperspectral imaging⁵⁴, the authors discussed the need to improve the efficiency and interpretability of the proposed models. In contrast, future research should further explore state-of-the-art methods and practical use cases. Zhang et al. SSDANet was designed by⁵⁵ to improve hyperspectral image classification, which suffers from overfitting and limited feature extraction, and future work will aim to reduce computational load.

Yuan et al.⁵⁶ demonstrated the progress made by deep learning in environmental remote sensing, examined its applications, and discussed the challenges and opportunities ahead. Hong et al.⁵⁷ combined drone hyperspectral imaging and deep learning to monitor dangerous algal blooms vertically, while also emphasising the effects of weather. Chen et al.⁵⁸ developed DR3D-CNN to classify hyperspectral images, leveraging dense connections and 3D convolutions to improve speed and accuracy. Wang et al.⁵⁹ developed DABL for HSI classification; future research should examine larger datasets. Accuracy was improved by domain adaptation and manifold regularisation.

Lightweight, real-time, and low-data models

Hyper-spectral image classification is computationally intractable, so neural models with a small number of parameters and real-time classification approaches have been investigated. Here, we focus on small-sample learning, semi-supervised/self-supervised methods and the high-throughput nature of deep learning. Chen et al. The study of⁶⁰ integrated deep learning and hyperspectral imaging for the accurate detection of bacterial disease in rice. Future development will mainly focus on improving field applicability. Ahmed et al. For example, Liu et al.⁶¹ utilised hyperspectral imaging and explainable AI to evaluate sweet potato quality. Future scope will be to increase feature selection and improve model generalisation. Wei et al. MPRI was proposed in⁶² to order categorical images, to categorise hyperspectral images, and to enhance accuracy with small sample sizes. Future work will mainly focus on random Fourier features for PRI optimisation. Xia et al. The few-shot HSI classification method that we mentioned in this section, DOGF⁶³: A Two-stream Network for Feature Extraction by Discarding Feature Interaction in Few-shot HSI Classification. Future work will enhance data augmentation and will include spatial-spectral interactions.

Sun et al. For example, the SMF-UL framework⁶⁴ was presented for the classification of hyperspectral images, leveraging the advantages of pretraining without the cost of retraining. The main limitations depend on the quality of the underlying source dataset and the similarity of the domain between the source and target datasets; hence, future work will focus on optimising generalisation and adaptability. Xu et al. The RSEN (Recurrent Spatial-Ensemble Network)⁶⁵ proposed in this research employs self-ensembling, enabling highly efficient HSI classification with considerably less labelled data. Exploring wider applications; downside: dependent on the quality of unlabeled data. Wieme et al.⁶⁶ on hyperspectral imaging for horticulture quality assessment and AI potential. Future research will mainly focus on standardisation approaches and data-efficient learning strategies. Some limitations include integration issues and data accessibility.

Zhan et al. ESSRAN, a model that integrates LSTM (Long Short-Term Memory) and ResNet (Residual Neural Network), and the SSAN model, proposed by⁶⁷, to enhance HSIC accuracy. New research will explore other models and semi-supervised character. Cao et al. An HSI classification method based on deep learning, combining a CNN, an MRF, and active learning, was proposed by⁶⁸. Forthcoming projects will investigate semi-supervised methods and improve active learning metrics. Sawant et al. Band selection methods for hyperspectral image classification were reviewed in⁶⁹, highlighting problems with noise, overfitting, and band redundancy. Moving forward, efforts should focus on designing reliable automated band-selection methods. Abdulridha et al. Hyperspectral imaging and machine learning (ML) techniques were used by⁷⁰ to detect stages of squash powdery mildew with high classification accuracy. Future work will focus on refining detection methods to suit the context. Sagan et al. Ref proposed the development of an automated calibration process for hyperspectral data⁷¹, to make it easier to handle extreme automation and analysis challenges in agriculture. The accuracy of these techniques will be enhanced in future research. Chen et al. In combination with PLG-KELM⁷², proposed to improve the classification accuracy of HSI by fusing PCA, LBP, GWO and KELM. Future work will be aimed at conducting more efficiently. Jia et al. LWCNN was proposed⁷³ for hyperspectral classification to solve the limited sample size issue. Further studies should explore larger datasets and methodologies.

Liu et al. The work in⁷⁴ presents a deep multiview-based unsupervised learning pipeline for HSI classification. Still, it has not been evaluated with a broader range of classifiers, nor is the time required to train the models discussed. Kang et al. Before work⁷⁵ proposed an LSTM-based high-accuracy classification method for the identification of foodborne pathogens, the author of this paper will investigate complex scenarios and mixtures of bacteria in the following work. Meng et al. MLNets⁷⁶ for HSI classification using residual connections, dense connections and residual dense connections for feature extraction; further network optimisations exist to explore in future work. Liu et al. To address this issue⁷⁷, proposed multitask deep learning for hyperspectral classification, leveraging knowledge transfer across multiple datasets to improve accuracy while reducing the risk of overfitting. Jia et al. In⁷⁸, the authors discussed the challenges, recent approaches, and future research directions of few-shot learning (FSL) for hyperspectral image classification (HSIC) within a deep learning framework. Wang et al. We have designed Cubic-CNN to improve HSI classification performance and significantly reduce training time; however, we will focus on optimising for low-sampling-rate and noisy data in the future⁵³.

Applications in agriculture, environment, and remote sensing

Hyperspectral Image Classification has broad applications, especially in agriculture, environmental monitoring, and remote sensing. Hyperspectral Data and Its Application: This section reviews the applications of hyperspectral data in precision agriculture, crop disease detection, food quality, and environmental monitoring. Olisah et al. A multi-input CNN model was proposed in²⁶ to classify and achieve high accuracy for blackberry ripeness. Further study may also expand the categorisation tasks and lessen the unpredictability of the environment. Ma et al.⁴¹ developed a CNN-based framework that uses UAV HSI and LiDAR data to classify tree species at the pixel level using U-SLIC and CBAM. Upcoming work will study self-supervised learning and broaden the applicability of the framework. Some limitations include challenges in data collection and possible imaging artefacts. Chen et al. To this end⁶⁰, applied an integrated deep learning and hyperspectral imaging model for accurate bacterial disease identification in rice. Future development will centre around improving field applicability. Ahmed et al. Deep

learning, with HRNet achieving the best performance at a small scale compared with Previous Works, has been demonstrated by⁴² for reconstructing hyperspectral images from RGB data for agricultural quality evaluation. In future work, we will tackle the challenges of generalisation and metamerism. Giakoumoglou et al. The study by⁴⁴ proposed a deep learning method for identifying *Botrytis cinerea* in cucumber using multispectral images. Further studies will also explore its potential in curing other plant diseases. Ahmed et al. Hyper-spectral imaging and explainable AI for sweet potato quality⁶¹. Future work will further improve feature selection and the model's generalisation.

Amoako et al. In⁴⁶, a two-stage meta-RL model combining deep Q-learning with capsule networks was also proposed to improve HSI classification accuracy and efficiency. Future Research will focus on minimising computational complexity and on finding models on other datasets. Farmonov et al. To classify crops with higher accuracy⁴⁷, used a Wavelet-attention CNN and DESIS hyperspectral images. Future research efforts will further explore improving accuracy and expanding the fields of application. Spectral information needs some continuity, and dependence on current DESIS data appears as one of the limitations. Noshiri et al. As a reference¹⁶, explored the pros and cons of 3D-CNNs applied to HSIs classification towards agricultural use cases. Future projects will enhance the model's effectiveness, affordability, and data access. However, it has its limitations, such as limited data, costs, and the burden of processing. Wieme et al. Hyperspectral imaging for horticulture quality assessment with a discussion of AI possibilities⁶⁶. In the future, the focus will be on standardisation methods and data-efficient learning paradigms. The shortcomings include integration issues and limited data access. Wang et al. They reviewed recent advances in deep learning for hyperspectral imaging in agriculture, identified problems, and offered recommendations for future research directions⁷⁹. Abdulridha et al. Detection of squash powdery mildew using hyperspectral imaging and machine learning achieved high classification accuracy⁷⁰. Future work is to develop better detection methods for other situations.

Sagan et al. This refers to the particular problem of hyperspectral data handling and analysis in agriculture, and to a hyperspectral calibration automation developed by⁷¹. These techniques can be refined further to improve prediction accuracy, which is a key area of future research. Yu et al. Apart from the above non-conventional approaches, silicon-based EEG sensors for signal detection were developed by⁵¹ that can be bent, do not irritate, and are comfortable to wear. Additional research might improve sensor design for more applications. A highly accurate foodborne pathogen classification approach based on LSTM was proposed by⁷⁵ and will be extended to complex samples and bacterial mixtures in the following steps. Yuan et al. Deep learning has made significant advances in environmental remote sensing⁵⁶; thus, Wang et al.⁵⁶ reviewed its development, applications, challenges, and opportunities in this field.

Hong et al. Authors⁵⁷ vertically monitored harmful algal blooms using combined drone hyperspectral imaging and deep learning, while also highlighting the effects of weather. Guerri et al.¹ addressed the complexity of the data and the evolving research requirements, focusing on advances in deep learning and hyperspectral imaging as applied to agriculture. Lin et al. Hyperspectral imaging, combined with Deep Learning (VGG19 + ResNet50), achieved effective forest type classification with up to 100% accuracy²³. Optimisation in Other Forest Conditions and Data Formats. The approach's unification across forest types is limited, and future work will be needed to integrate systems for real-time forest management. Lu et al. A multitask deep learning approach for effective screening of hyperaccumulators and heavy metals measurement using spectral techniques²⁴. In future work, we will discuss applying the method to other plants, including enabling real-time remote sensing. One of the disadvantages is that more diverse spectrum data is required.

Optimisation, band selection, and future directions

Hyperparameter tuning and band selection are essential optimisation techniques that can enhance the performance and efficiency of hyperspectral image classifiers. We review these optimisation approaches in this section and identify current shortcomings and opportunities for future research to alleviate the computational burden while improving the model's performance. Dash et al. The classification of hyperspectral images using CNN and LSTM was evaluated by²⁵. ²⁵ shows that LSTM provided efficient performance with MNF pre-processing. Generality will also improve over time. Ahmad et al. To address hyperspectral image categorisation with reduced parameter tuning and lower computational cost⁴³, propose Sharpened Cosine Similarity (SCS). This will serve as the basis for further work to identify other improvements and achieve meaningful use. Yao et al. To address the issues of hyperspectral image noise reduction and classification accuracy, Huang et al.⁴⁵ developed a deep hybrid multi-graph network (DHMG). Future research will explore unsupervised models to address the shortage of labelled data. Yu et al. ConGCN²⁸ further combined spectral and spatial information to improve feature representation and proposed a contrastive learning-based approach utilising a convolutional GNN (ConGCN) for HSI classification. Future work will leverage the lack of data and improve Graph Augmentation (Sellami et al.). In⁴⁸, a semi-supervised hypergraph convolutional net was proposed to perform unsupervised feature selection and classify hyperspectral images. Further work will scale the model and generalise to other datasets. One drawback is that it needs more validation and optimisation in wider applications.

Esmaeili et al. CNNeGA utilised a 3D-CNN in combination with high-dimensional data and a GA for band selection to efficiently and accurately analyse hyperspectral images⁷. Future work will focus on applying spiking neural networks to unsupervised band selection and classification. Wang et al. HP-NAS: A Hyper-kernel-based Neural Architecture Search Method for Hyperspectral Image Classification with Reduced Searching Costs and Enhanced Model Flexibility⁴⁹. Once this technique has been developed, further work can build on it to continue broadening its scope. A drawback of this approach is the complexity of implementing different 3-D CNN designs. Vasanthakumari et al. For instance¹⁸, employed a CNN with the Modified ReLU activation function, which outperformed on multispectral image categorisation. Future work will also explore additional optimisers and band arrangements. One drawback is its greater computational requirements than elementary models. Pathak et al. We have achieved better performance by creating a spectral-spatial classification approach using

SVM⁵⁰. Future Developments: Parameter optimisation and broader applications. Among the restrictions are the need for far better function extraction and tuning. Dong et al. Based on the above discussion³⁶, proposed a WFCG combining a GAT and a CNN for HSI classification, leveraging the strengths of both. Improving the model's speed and tuning hyperparameters are among the goals for the next steps, given the limited sample size and computing power.

Sawant et al. described a band selection strategy for hyperspectral image classification that addresses noise, overfitting, and redundancy⁶⁹. Future studies should build on this work to develop efficient, automated band-selection methods. Chen et al. For hyperspectral image classification⁷², proposed improving PLG-KELM accuracy by combining PCA, LBP, GWO, and KELM. The aspiration for future work is to optimise operations. Wang et al. DABL was developed by⁵⁹ for HSI classification, but future research should focus on larger datasets. Domain adaptation and manifold regularisation improved accuracy. Hong et al. This is, in fact, the primary focus of a review of nonconvex models examined from the perspective of efficiency and interpretability by⁵⁴ in the area of hyperspectral imaging; future work should integrate new methods and applications⁵⁴. Zhang et al. SSDANet was designed to improve hyperspectral image classification by preventing overfitting and adjusting feature extraction. In future work, the authors aim to reduce computational cost⁵⁵.

Meng et al. MLNet⁷⁶ proposed MLNet for HSI classification and incorporated a mixture of dense and residual connections to improve feature extraction, but more network optimisations may be employed. Safari et al. The authors of³⁸ introduced a deep learning framework for HSI classification, achieving an accuracy of 66.73%. More work will address the sensitivity question and identify better normalisation approaches. Shen et al. The ENL-FCN efficiently and practically enhances long-range context in HSI classification by adopting nonlocal modules; more work will further boost this efficiency and confirm it. Liu et al. Multitask deep learning for hyperspectral classification⁷⁷ improves classification accuracy by leveraging its dataset while simultaneously addressing the small-sample-size and overfitting problems. Mou et al. An unsupervised hyperspectral band selection method is proposed that utilises deep reinforcement learning to develop a new, promising incentive scheme that improves accuracy and efficiency⁴⁰.

This recent series of studies summarises deep learning approaches for hyperspectral image classification over the years. Ranjan et al. Zero-shot learning: Different architectures, such as 3D Convolutional Autoencoders⁸⁰, Conditional GANs⁸¹, Lightweight Xcep-Dense Models⁸², etc., have been proposed in⁸³. Semi-supervised⁸⁴ and cross-domain⁸⁵ methods were proposed to address low-resource constraints. Several deep learning models explore spectral–spatial fusion⁸⁵, attention⁸⁵, and comparative evaluations of CNNs^{86–88}. These improvements indicate a continued curiosity about performance, generalizability, and (different kinds of) label efficiency in classification. However, existing approaches handle spectral, spatial, and attention-based dependencies independently,^{89–92} which motivates HSICNet to offer an integrated and efficient alternative.

We compare them in Table 1 and outline their main limitations relative to HSICNet. The last selection includes 3D CNNs, attention-based CNNs, a hybrid transformer, and a lightweight architecture. Even though these methods are effective for spectral-spatial learning, they have some disadvantages, including computational expense, shallow feature fusion and/or a complex tuning process, and poor generalisation in few-label cases. However, HSICNet decouples spatial and spectral feature learning in a dual-branch architecture of 1D and 2D CNNs. Then, features are decoupled and fused by an efficient attention-guided fusion module. The additional use of PCA reduces dimensionality and complexity. By combining these design choices, we arrive at a thin yet robust model that retains similar accuracy to previous models, overcoming the overfitting higher-dimensional curse and enabling transfer across several hyperspectral datasets.

Proposed framework

Here, we introduce a nearly-ideal hyperspectral Image Classification framework called HSICNet. The proposed system architecture is introduced, including the dual-branch spectral and spatial feature extractors, attention-guided fusion, and high-dimensional feature compression. We discuss the mathematical formulations, notation, algorithms, and architectural diagrams, and explain how each component works in the end-to-end classification pipeline.

Author / Model	Architecture highlights	Dataset(s) used	Key limitation	Gap addressed by HSICNet
Ashraf et al. ⁶	3D UNet with spectral-spatial encoding	Indian Pines	Heavy computation, overfitting	Lightweight dual-branch structure
Chhapariya et al. ⁵	Residual attention CNN (DSSpRAN)	Indian Pines, Salinas	Shallow depth, weak spatial learning	Stronger 2D spatial stream
Reddy et al. ³	3D CNN with self-attention	Indian Pines	High complexity, difficult real-time use	Lightweight fusion with attention
Farooque et al. ⁴	Swin Transformer + 3D Atrous CNN	Salinas	Overhead in transformer fusion	Efficient attention-guided fusion
Dong et al. ³⁶	CNN + Graph Attention Network (GAT)	Pavia University	Tuning complexity, less interpretability	Simpler fusion, interpretable weights
Sun et al. ⁹	Low-cost sparse CNN (LCTCS)	Pavia University	Accuracy drops under an imbalance	Improved per-class generalisation
Zhu et al. ¹⁵	SS-ConvNeXt for denoising and spatial features	Indian Pines	High computation cost	PCA + efficient convolution layers
Esmaili et al. ⁷	CNN + Genetic Algorithm for band selection	Salinas	Not real-time suitable	End-to-end integration of PCA
Sellami et al. ⁴⁸	Semi-supervised Hypergraph CNN	Indian Pines	Poor scalability on extensive data	Generalizable across three datasets
Bai et al. ¹¹	Spectrum Complementary Learning Network	Indian Pines	Does not use generative fusion	Dynamic attention-based fusion

Table 1. Comparative summary of existing deep learning methods for hyperspectral image classification and the gap addressed by HSICNet.

Overview

An HSICNet framework for hyperspectral image classification is proposed, as depicted in Fig. 1. It is a framework designed to work directly with hyperspectral data, making full use of the available spatial and spectral information within a deep learning pipeline. The workflow starts with the input block, which includes hyperspectral image datasets such as Indian Pines, Pavia University, and Salinas. Those datasets are popular benchmarks for hyperspectral imaging and encompass diverse scenarios, e.g., agricultural and urban.

Initial block preprocessing processes the input data with the necessary steps to prepare the Data in a proper format for model training and cleaning. Then, noise reduction is applied to remove low-SNR bands, and the dimensionality is reduced using Principal Component Analysis (PCA) to overcome the curse of dimensionality while retaining most of the spectral variance. The data is then split into training and test subsets using stratified sampling to ensure similar class distributions. Additionally, flipping, rotation, and other data augmentation methods are used to improve the model's generalisation and reduce overfitting.

The data is then fed into the spectral feature extraction block, where, after preprocessing, the proposed method employs 1-D convolutional layers along the spectral axis to learn high-resolution spectral features for each pixel. Since differentiating materials in hyperspectral imagery depends on extracting band-band correlations as well as spectral signatures, these layers can be better than others. The spectral features that we have retrieved will then be passed through the spatial feature extraction block, where some of them at the dubbed two-dimensional convolutional layers will be processed using specific spatial neighbourhoods. You can use this module to learn local textures and pixel-level spatial dependencies that complement spectral information to represent features better.

In it, Spectral and spatial branches are used for feature extraction, which are fused in a fusion block through a spectral-spatial fusion module with joint attention for high-resolution HSI. See global information and generates attention values on every channel. It learns the dependencies among feature channels and assigns weights to each feature. This feature map is a more complex and informative representation of the original one, which helps the model better differentiate between tasks during classification.

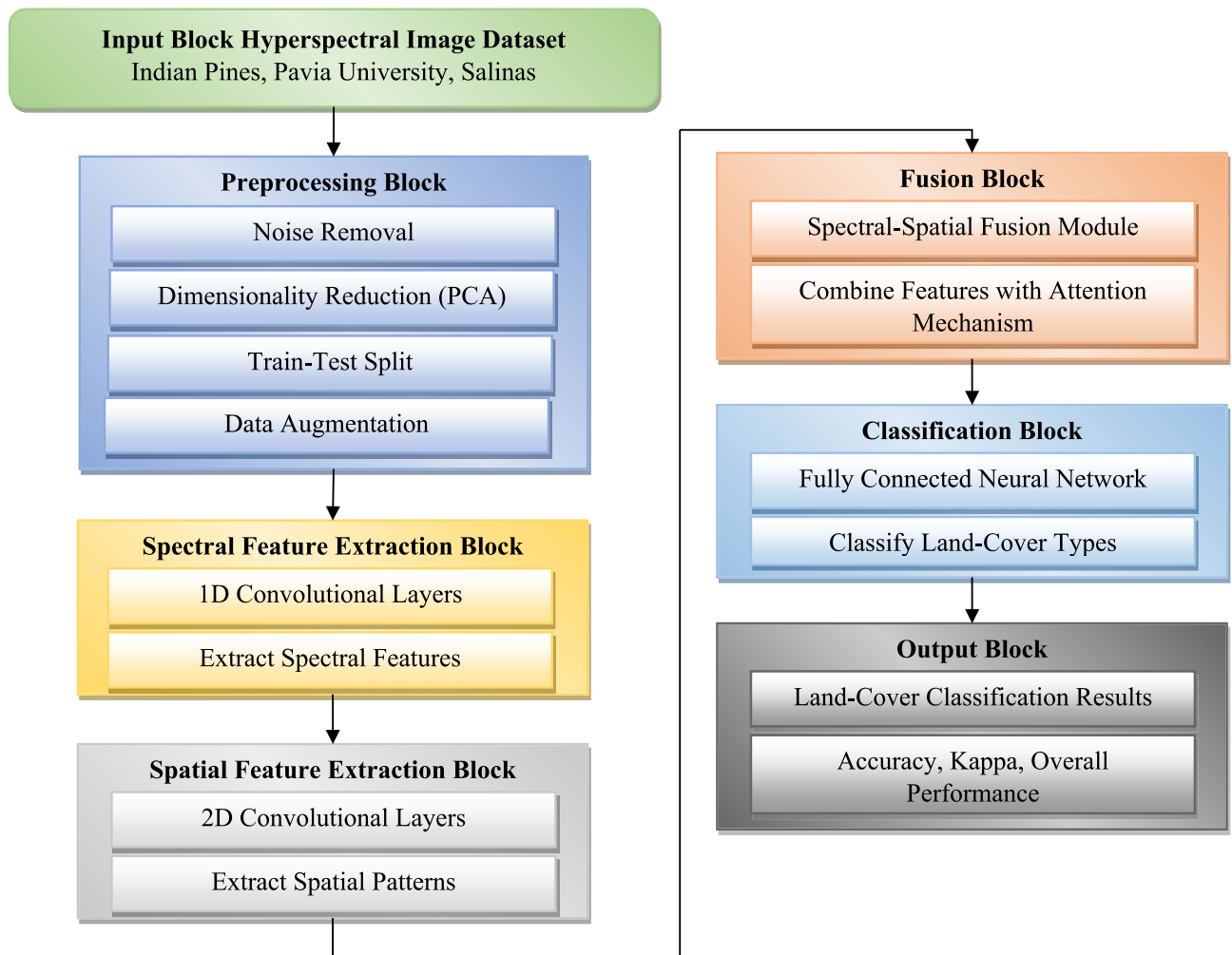


Fig. 1. Proposed HSICNet framework for hyperspectral image classification.

A fully connected neural network linking the output of a fusion block to a classification block analyses the features and outputs class labels for every pixel. Finally, these land cover types will be classified into those present in the hyperspectral image. In the output layer, the final predictions are cross-validated and assessed using performance measures such as overall accuracy, the kappa coefficient, and other statistical indicators that assess classification reliability, class balance, and overall system performance. In this work, we introduce a model designed for rapid, high-resolution hyperspectral image classification in remote sensing and environmental monitoring.

Instead of using naive concatenation of features extracted by two branches, as well as using deep self-attention layers in previous dual-branch networks to fuse features, HSICNet uses a fusion mechanism based on attention, which combines lightness and effectiveness, since attention weights are calculated from global average pooling (GAP) and then recalibrated by a channel recalibration network. Our proposed design facilitates targeted enhancement of discriminative spectral-spatial features with minimal overhead, improving classification in label-scarce or low-resource conditions.

Proposed model

The proposed internal framework of HSICNet (Fig. 2) for hyperspectral image classification is also inspired by the fusion and interpretation of stacked spectral and spatial features. The initial step takes an input hyperspectral image dataset from publicly available datasets, including Indian Pines, Pavia University, and Salinas. These datasets include various land-cover types and environmental conditions, making them a good ground for evaluating model flexibility.

The Step is a complete preprocessing step that converts the input data to the desired format and fills in null values with numeric values. This includes noise removal (i.e., eliminating non-informative or corrupted spectral bands) and principal component analysis (PCA) (to reduce the number of dimensions of the hyperspectral cube while preserving a substantial amount of variance). The dataset is split into training and test sets at an appropriate ratio for supervised learning. Some augmentation is applied to obtain more diverse training samples, thereby strengthening the models' generalisation.

Thus, after preprocessing, the data are fed into the spectral feature extraction module, which applies 1D convolutional layers along the spectral dimension of each pixel vector. To derive high-resolution spectral signatures of materials from hyperspectral imagery, this step is required. Spectral characteristics are sufficiently descriptive to promote the separation of classes across narrow class domains.

After spectral extraction, the output is fed into the spatial feature extraction module, where 2D convolutional layers process local neighbourhoods in spatial space. It encodes spatial patterns and contextual relationships among correlated pixels, which are essential for characterising textures and structures associated with land cover classes.

The outputs of the spectral and spatial modules are sent to the spectral-spatial fusion module. This component combines the extracted features with an attention mechanism that dynamically learns to highlight the most discriminative features and penalise the irrelevant. The combined output provides a more complete representation with both spectral richness and spatial structure.

Finally, this fused set of features is fed into the classification layer—a fully connected neural network. The final layer classifies these feature representations into predefined land-cover classes and provides class predictions for each image pixel.

The output layer is the last layer of the model, which you use to predict your output for the classification tasks. To evaluate the HSICNet model, several key performance metrics are computed, including overall accuracy, the Kappa coefficient, and class-level performance. It has also indicated that employing attention mechanisms alongside deep spectral-spatial learning will achieve precise and powerful HSI classification, as illustrated in the structured flowchart. Mathematical expressions for spectral, spatial, fusion, and classification operations are presented in the notations section, with some of the critical ones summarised in Table 2 (Covered under HSICNetUnified).

The attention-guided spectral-spatial fusion module in HSICNet is shown at the block level in Fig. 3. First, the spectral feature maps from the 1D CNN branch and the spatial feature maps from the 2D CNN branch are concatenated to form a joint representation. We use global average pooling to compute channel statistics, which are then fed into two fully connected layers with non-linear activations to obtain channel-wise attention weights. This weights rescale the fused feature channels, highlighting discriminative spectral-spatial responses, while suppressing redundant information, and forward the refined features to output final classification layers.

Mathematical formulation of HSICNet framework

The proposed HSICNet architecture is mathematically validated for exploiting spectral and spatial information in hyperspectral classification. Let $X \in \mathbb{R}^{m \times n \times b}$ be the hyperspectral input image, where m and n are the spatial dimensions and b is the number of spectral bands. The model gives a classification and output at different levels to get information and makes predictions that separate images according to their classification.

Spectral feature extraction follows in the first stage via 1D convolutional layers applied to the spectral dimension. The spectral convolution operation for an input pixel vector $x \in \mathbb{R}^b$ is given as in Eq. 1.

$$h_{i,j}^{spec} = \sigma \left(\sum_{k=1}^b x_k \cdot w_k^{spec} + b^{spec} \right) \quad (1)$$

where $h_{i,j}^{spec}$ is the output of the spectral convolution for pixel (i, j) , w_k^{spec} and b^{spec} are the spectral kernel weights and bias, respectively, $\sigma(\cdot)$ is the activation function (e.g., ReLU).

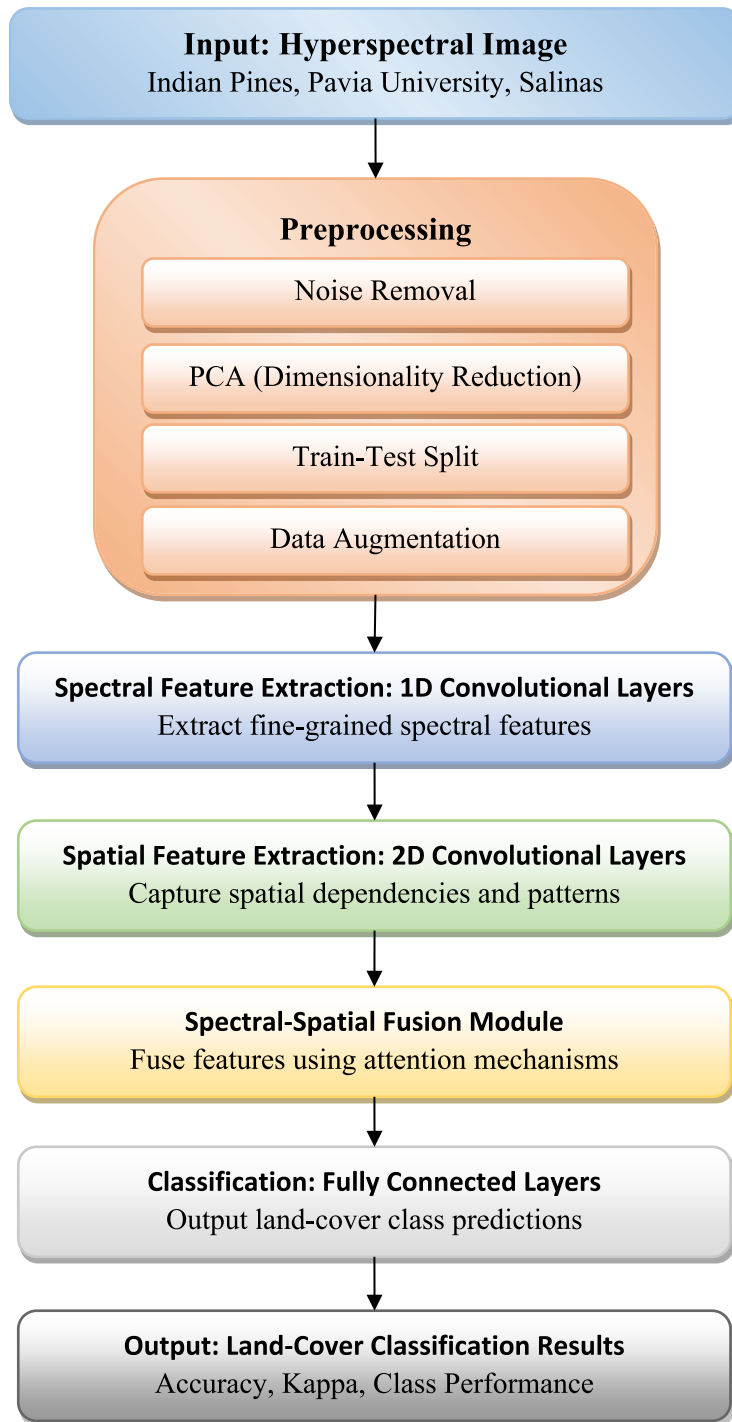


Fig. 2. Proposed HSICNet model architecture for hyperspectral image classification.

In the second stage, 2D convolutional layers are employed to extract spatial features from the encoded representation to capture spatial dependencies in the 2D signal of neighboring pixels. (i, j) . The operation of spatial convolution (in pixels) is defined in Eq. 2.

$$h_{i,j}^{spat} = \sigma \left(\sum_{p=-P}^P \sum_{q=-Q}^Q x_{i+p,j+q} \cdot w_{p,q}^{spat} + b^{spat} \right) \quad (2)$$

Notation	Description
$X \in \mathbb{R}^{m \times n \times b}$	Hyperspectral input image, where m and n are spatial dimensions, and b is the number of spectral bands
$x \in \mathbb{R}^b$	Spectral pixel vector from the hyperspectral image
$h_{i,j}^{spec}$	The output of the spectral convolution for pixel (i, j)
w_k^{spec}	Weight of the spectral convolution kernel for the k -th band
b^{spec}	Bias term for the spectral convolution
$\sigma(\cdot)$	Activation function (e.g., ReLU)
$h_{i,j}^{spat}$	Output of the spatial convolution for pixel (i, j)
P, Q	Spatial kernel size along rows and columns, respectively
$x_{i+p,j+q}$	Neighboring pixel value at offset (p, q)
$w_{p,q}^{spat}$	Weight of the spatial convolution kernel for the (p, q) -th position
b^{spat}	Bias term for the spatial convolution
α_k	Attention weight for the k -th feature channel
e_k	Importance score for the k -th feature channel
C	Total number of feature channels
$z_{i,j}$	Fused spectral-spatial feature at pixel (i, j)
$z \in \mathbb{R}^d$	The flattened fused feature vector for a given pixel, where d is its dimensionality
$\hat{y} \in \mathbb{R}^K$	Predicted class probabilities for K land-cover classes
W_k	Weights of the fully connected layer for the k -th class
b_k	Bias of the fully connected layer for the k -th class
L	Cross-entropy loss function
N	Total number of samples
$y_{i,k}$	Ground truth label for the k -th class for the i -th sample
$\hat{y}_{i,k}$	Predicted probability for the k -th class for the i -th sample
OA	Overall Accuracy – percentage of correctly classified pixels
AA	Average Accuracy – mean accuracy across all classes
κ (kappa)	Kappa Coefficient – agreement metric between prediction and truth
F1	F1-Score – harmonic mean of precision and recall per class
ReLU	Rectified Linear Unit – a non-linear activation function
GAP	Global Average Pooling – used in attention weight computation

Table 2. Description of mathematical notations used in the HSICNet framework for hyperspectral image classification.

where P and Q are the spatial kernel size, and $w_{p,q}^{spat}$ and b^{spat} are the spatial kernel weights and bias, respectively. An attention mechanism is then employed to selectively fuse these spectral and spatial features, emphasizing the most informative ones. We calculate the attention weight, α_k for the k -th feature channel as in Eq. 3.

$$\alpha_k = \frac{\exp(e_k)}{\sum_{l=1}^C \exp(e_l)} \tag{3}$$

where e_k is the importance score for the k -th channel, C is the total number of channels and α_k is the normalized attention weight. The fusing of feature $z_{i,j}$ at pixel (i, j) is in Eq. 4.

$$z_{i,j} = \sum_{k=1}^C \alpha_k \bullet h_{i,j}^{(k)} \tag{4}$$

Finally, these fused features go through fully connected to classify them. Assume $z \in \mathbb{R}^d$ is the stacked feature vector for a specific pixel and d The dimension of the fused features. For this, class probabilities $\hat{y} \in \mathbb{R}^K$ are defined as in Eq. 5.

$$\hat{y}_k = \text{softmax}(W_k \bullet z + b_k) \tag{5}$$

where W_k and b_k are the weight and bias of the k -th class, and K is the total number of classes—train the model by minimizing the cross-entropy loss function expressed in Eq. 6.

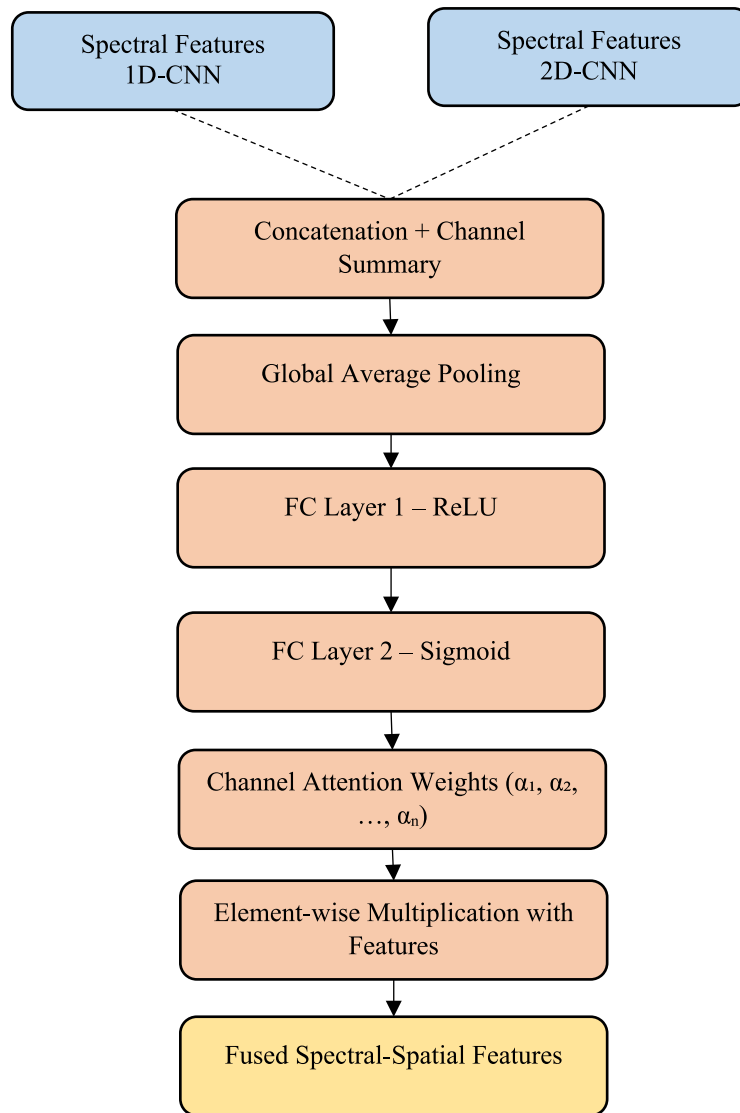


Fig. 3. Block diagram of the attention-guided spectral–spatial fusion module in HSICNet.

$$\mathcal{L} = -\frac{1}{N} \sum_{i=1}^N \sum_{k=1}^K y_{i,k} \log(\hat{y}_{i,k}) \quad (6)$$

where N is the total number of samples, $y_{i,k}$ is the ground truth label, $\hat{y}_{i,k}$ and is predicted probabilities.

HSICNet uses this mathematical formulation to perform an end-to-end process, starting from raw input data until obtaining the final classification and fully utilizing spectral-spatial feature extraction.

Proposed algorithms

Proposed algorithms form the basis of the HSICNet framework, which comprises modular deep learning components designed to address crucial stages of hyperspectral image classification. Spectral extraction, spatial feature extraction, spectral-spatial fusion via attention, and classification are the four groups of algorithms. From the preprocessing stage to spectral and spatial convolutional layers, fusion mechanisms, and classification, we provide a detailed description of the HSICNet pipeline in this section. We first present a message-passing framework for computing the node frequencies of a graph (i.e., assigning nodes to structural features), enabling efficient feature learning while remaining computationally tractable and robust. Overall, the results demonstrate that specialised modules can be effectively integrated to achieve accurate, scalable classification of hyperspectral data.

The entire end-to-end workflow of the proposed HSICNet framework for hyperspectral image classification is shown in Algorithm 1. The first phase involves hyper-spectral datasets as input, processed through several preprocessing stages, including noise removal, PCA dimensionality reduction, train-test splits, and augmentation to enhance the quality and viability of the input. The spectral features are extracted using 1D convolutional layers

Input: Hyperspectral image X , ground truth labels Y

Output: Predicted labels \hat{Y}

1. Preprocess X (noise removal, PCA, train-test split).
2. Extract spectral features H^{spec} Using Algorithm 2.
3. Extract spatial features H^{spet} Using Algorithm 3.
4. Fuse H^{spec} and H^{spet} into Z Using Algorithm 4.
5. Pass Z through fully connected layers to classify them into \hat{Y} .
6. Compute evaluation metrics (OA, AA, κ , F1-score).
7. Return \hat{Y} .

Algorithm 1. HSICNet main algorithm.

Input: Hyperspectral image X

Output: Spectral feature map H^{spec}

1. For each pixel x in X :

Apply 1D convolution across spectral bands:

$$h_{i,j}^{\text{spec}} = \sigma \left(\sum_{k=1}^b x_k \cdot w_k^{\text{spec}} + b^{\text{spec}} \right)$$

Store $h_{i,j}^{\text{spec}}$ in H^{spec} .

2. Normalize H^{spec} .
3. Return H^{spec} .

Algorithm 2. Spectral feature extraction.

that score pixels along the spectral dimension, preserving high-resolution spectral features after preprocessing. Second, they use 2D convolutional layers that learn spatial contextual information from the pixel neighbourhoods. Then, these spectral and spatial features are fused using a spectral-spatial fusion module that employs attention mechanisms to focus on informative features. The fused features are fed into fully connected neural network layers for land cover type classification. Finally, the algorithm evaluates several metrics, including accuracy, kappa coefficient, and class-wise performance, to measure the goodness of the proposed model. HPTN achieves interpretability, scalability, and high classification performance with this modular structure on numerous hyperspectral datasets.

In the spectral feature extraction scheme of the framework, one-dimensional convolutional operations are applied to the spectral bands of each pixel vector in the hyperspectral image, as shown in Algorithm 2. Each pixel with a high-dimensional spectral signature enables the various 1D convolutions to preserve subtle spectral differences and similarities between adjacent bands. It will allow the model to learn the spectral patterns of the materials that are required to discriminate classes. After convolution, the non-linear activation function is applied to the obtained spectral features from high-level movement contexts to fuse the spectral feature scale. These features contain richer spectral information and are then forwarded to the spatial feature extraction module. This stage lays the foundation for enhancing the model's discriminative capability and preserving the spectral information of the input data.

HSICNet spatial feature extraction process shown in Algorithm 3. After obtaining the spectral feature maps in the previous stage, we extract spatial features from these transformed spectral maps using 2D convolutional layers. First, the spatial layers are applied to local neighborhoods of the hyperspectral image within the spatial domain, allowing the model to learn spatial dependencies, textures, and structural patterns among the neighboring pixels. Convolutional filters scan the spatial plane to learn spatial features (edges, shapes, and boundaries) that are important to the spectral profiles. After that we perform a non-linear activation on the feature representation and normalize the spatial feature map. The image module further allows the model to learn the spatial context knowledge of land-cover classes that are important in order to benefit classification in heterogeneous regions. Then, the output will be fed into the fusion module to combine the spectral features.

We then propose HSICNet, which comprises spectral-spatial fusion and an attention mechanism, along with the algorithm empowered by these mechanisms. The spectral feature map from 1D convolutions and the spatial feature map from 2D convolutions are sent to a fusion module for the fusion of the two feature sets. Under an attention mechanism that computes a feature-channel importance score, this fusion is performed. The feature importance scores, extracted via a softmax over the attention weights, allow the model to automatically learn the most discriminative features in the fused spectral-spatial domain. The contributions of each channel are multiplied by the ratio of their attention scores, generating a fused feature map that provides a much more powerful, contextually aware visual representation of the input hyperspectral data. This wealth of features is

Input: Spectral feature map H

Output: Spatial feature map H^{spat}

1. For each spatial window in H^{spec} :

Apply 2D convolution over spatial dimensions:

$$h_{i,j}^{spat} = \sigma \left(\sum_{p=-P}^P \sum_{q=-Q}^Q x_{i+p,j+q} \cdot w_{p,q}^{spat} + b^{spat} \right)$$

Store $h_{i,j}^{spat}$ in H^{spat} .

2. Normalize H^{spat} .
3. Return H^{spat}

Algorithm 3. Spatial feature extraction.

Input: H^{spec} , H^{spat}

Output: Fused feature map Z

1. Compute attention weight α_k For each channel:

$$\alpha_k = \frac{\exp(e_k)}{\sum_{l=1}^C \exp(e_l)}$$

where e_k is the importance score.

2. Fuse spectral and spatial features: $z_{i,j} = \sum_{k=1}^C \alpha_k \cdot (H_k^{spec} + H_k^{spat})$
3. Store $z_{i,j}$ in Z .
4. Return Z .

Algorithm 4. Spectral-spatial fusion and attention mechanism.

then used by classification layers, allowing the model to take full advantage of spectral complexities and spatial structures by passing them all in parallel to the end of the pipeline.

Evaluation methodology

The proposed HSICNet model's performance is examined on several hyperspectral datasets, demonstrating its robustness and efficiency in hyperspectral image classification. The evaluation metrics include standard measures like Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ), and class-wise F1 score. Overall Accuracy (OA) indicates the proportion of accurately predicted pixels (TP + TN) to all pixels. It is defined as in Eq. 7.

$$OA = \frac{\sum_{i=1}^K TP_i}{N} \quad (7)$$

where TP_i denotes the true positives for class i , K is the number of classes, and N is the total number of pixels in the dataset. Average Accuracy (AA) measures the mean class-level classification accuracy, ensuring that class imbalance does not significantly impact overall performance. It is computed as in Eq. 8.

$$AA = \frac{1}{K} \sum_{i=1}^K \frac{TP_i}{TP_i + FN_i} \quad (8)$$

where FN_i is the false negatives count per class i . The Kappa Coefficient (κ) measures how closely the predicted and actual classifications match, given that not all categories are perfect. It is given by Eq. 9.

$$\kappa = \frac{OA - PE}{1 - PE} \quad (9)$$

where PE is the expected accuracy that can be obtained only through chance and is given by Eq. 10.

$$PE = \frac{\sum_{i=1}^K (T_i \cdot P_i)}{N^2} \quad (10)$$

Here, T_i and P_i correspond to the global actual and predicted pixels for class i , respectively.

In addition, the F1 score is balanced between precision and recall for each class. The F1 score for class i is given by Eq. 11.

$$F1_i = 2 \cdot \frac{Precision_i \cdot Recall_i}{Precision_i + Recall_i} \quad (11)$$

We define precision and recall as in Eqs. 12 and 13.

$$Precision_i = \frac{TP_i}{TP_i + FP_i} \quad (12)$$

$$Recall_i = \frac{TP_i}{TP_i + FN_i} \quad (13)$$

Experiments are performed on Indian Pines, Pavia University, and Salinas datasets to verify the model's generalizability. Separate training, validation, and test sets were created from the datasets using stratified sampling rather than random sampling to maintain class distributions. Cross-validation, in turn, is used to achieve low bias and variance, giving more reliable performance estimates. The model's performance is evaluated against state-of-the-art baselines that leverage traditional machine learning methods, such as Support Vector Machines (SVMs), and deep learning architectures, including 3D CNNs.

By conducting ablation experiments, we investigate the roles of the key components of HSiCNet, including the spectral-spatial fusion module and the attention mechanism. It also enables us to evaluate performance fluctuations when we remove or change a particular module, i.e., the contribution of each component. We evaluate HSiCNet for its robustness, reliability, and effectiveness on hyperspectral remote sensing images using these metrics, measurement criteria, and evaluation methodologies.

Experimental results

We present a comprehensive performance evaluation of HSiCNet on three benchmark HSI datasets, i.e., Indian Pines, Pavia University, and Salinas, in the experimental results section. We evaluate the model's effectiveness using experiments with different configurations and by assessing its classification performance, generalisation capability, and time complexity. Experiments were conducted on a workstation with an NVIDIA RTX 3080 GPU, 64 GB of RAM, and an Intel Core i9, in Python using the TensorFlow and Keras libraries. We preprocess the datasets by removing most of the spectral dimension using noise-band removal and PCA, while keeping 99% of the variance. We performed stratified splitting; hence, the class ratios observed within the whole dataset were able to be reproduced in the test set (70/30) per standard practice.

The systematic hyperparameter setup procedure used in the experiments ensures replicability. The learning rate was set at 0.001 and optimised with a cosine decay scheduler. We used the Adam optimiser with a batch size of 64 and categorical cross-entropy loss. Data were split 80/20 for training and validation, and the model was trained for 200 epochs, with early stopping when validation accuracy did not increase. A dropout rate of 0.5 was employed to address overfitting. Kernel sizes were 7 and 5 for 1D spectral convolutional layers, and 3×3 for 2D spatial convolutional layers. Fully connected hidden layers used ReLU activation functions, and the output classification layer used softmax activation. To obtain the attention weights, we use global average pooling followed by a fully connected squeeze-and-excitation mechanism.

In terms of reproducibility, the complete experimental pipeline, including data preprocessing scripts, model architecture definition, training configurations, and evaluation routines, has all been modularised into a prototype implementation. Across all datasets, similar preprocessing steps are used, and the hyperparameters are defined in configuration files so that other researchers can efficiently run experiments and benchmarks based on our work. We set random seeds for NumPy, TensorFlow and Python to produce the same results over runs. This amount of detail is intended to facilitate transparent experimentation and reproducibility of the proposed HSiCNet framework by the research community.

Nevertheless, only analyses of class distribution were performed to control for potential dataset imbalance, and no class-balancing method (oversampling, undersampling, or a weighted loss function) was directly applied during training. We intentionally chose this setting to ensure class distributions are naturally imbalanced, a typical characteristic of real-world hyperspectral applications, and to examine the robustness of HSiCNet to this phenomenon. Class-wise accuracy results for the model were also achieved, with minority-class performance remaining stable.

Between the training and test data sets, we performed a stratified split of 70:30; however, the split was stratified to maintain the same class distribution in both the training and test subsets. 10% of the training data was held out as a validation set (used during training) to determine when the network had converged and to prevent overfitting.

Dataset details

Hyperspectral data sets used in hyperspectral image classification across different application domains, with a brief explanation⁴⁰. The first one is the Indian Pines dataset, a hyperspectral dataset containing 220 bands spanning 400–2500 nm (visible to near-infrared). Acquired over an agricultural area in Indiana, USA, it includes 16 land-cover classes, including different crop types, grasslands, and forests. Due to mixed pixels and the dataset's low spatial resolution, classification is challenging, making it a valid benchmark for evaluating the robustness of hyperspectral image classifiers.

Pavia University database is a standard database used for hyperspectral image classification. It includes 103 spectral bands and a high-resolution image of the University of Pavia, Italy. The dataset contains urban structures (buildings and roads) and vegetation (meadows), making it an ideal candidate for land cover classification and monitoring urban areas. With this wealth of spectral and spatial information, it can provide an opportunity for scholars to evaluate model performance in urban areas. Salinas: Hyperspectral data of the Salinas Valley, California, including 224 spectral bands and high spatial resolution. Comprised of 16 agricultural classes spanning several crops and soil types. This data set is the first to cover a diverse range of crops and soils, making it well suited for testing models that will be used for agricultural and vegetation monitoring. It offers data with a high spatial and spectral resolution for environmental and agricultural applications. The combination of these three datasets provides diversity in the challenges and scope of useful characteristics for evaluating the performance of classification models in agricultural, urban, and environmental contexts from hyperspectral images.

Stratified sampling was used to generate the training and test splits, ensuring roughly equal class ratios across both sets, limiting performance variance, and enabling fairer comparisons of results. Additionally, HSiCNet was tested on naturally imbalanced data without class-balancing during training (a typical case in many real-world distributions).

Exploratory data analysis

Below is the exploratory data analysis (EDA) of the benchmark hyperspectral datasets used in the current study. It consists of RGB composites, class distribution histograms, the mean spectral profiles of selected classes, and PCA variance graphs. Performing an EDA can reveal issues such as (1) spectral dimensionality, (2) spectral class imbalance, and (3) data dimensionality, which can help inform preprocessing strategies (e.g., dimensionality reduction) and model architecture design choices.

Figure 4 Indian Pines dataset distribution of classes (16 classes in total, including different land/cover types such as crops and other vegetation types). The dataset is heavily imbalanced, with some classes containing only a few samples. Hyperspectral image classification models struggle with class imbalance; therefore, they require strategies to address it and appropriate training methods.

The Pavia University dataset consists of 9 urban land-cover classes; the distribution of samples in each class is shown in Fig. 5. Imbalanced Dataset: Some class features, such as asphalt, trees, and self-blocking bricks, appear in the produced classes, but with unbalanced frequencies. Due to its relatively small footprint of urban diversity, the UCMerced Land Use Dataset serves as a challenging validation set for hyperspectral image classifiers, particularly for similar-structure and dominant-class effects.

The class distribution for the Salinas dataset is shown in Fig. 6, which comprises 16 detailed agricultural classes, including Grapes_untrained, Celery, and different stages of Lettuce_romaine. The class frequencies in the dataset are imbalanced, which may compromise classification performance. Due to its high resolution and crop-specific content, it provides an excellent benchmark for testing hyperspectral models for agricultural and land-use monitoring.

In this study, we consider three benchmark hyperspectral datasets, i.e., Indian Pines, Pavia University, and Salinas, and their RGB composites are shown in Fig. 7. Different spectral bands are selected for each image to

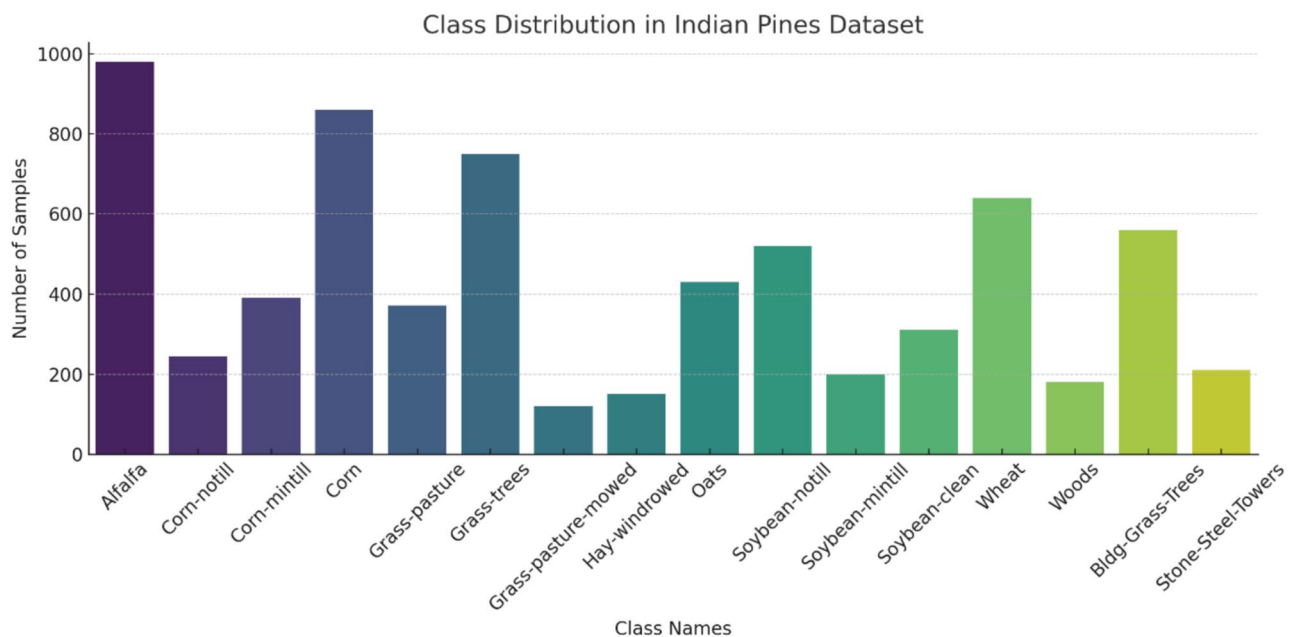


Fig. 4. Class distribution in the Indian Pines hyperspectral dataset.

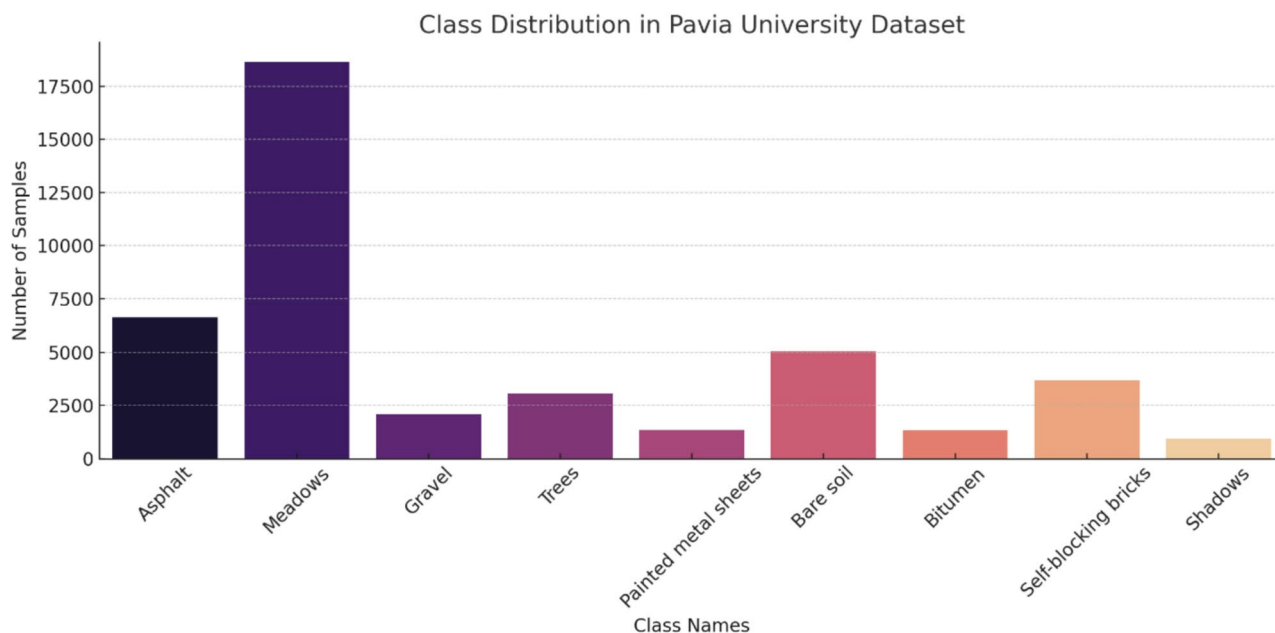


Fig. 5. Class distribution in the Pavia University hyperspectral dataset.

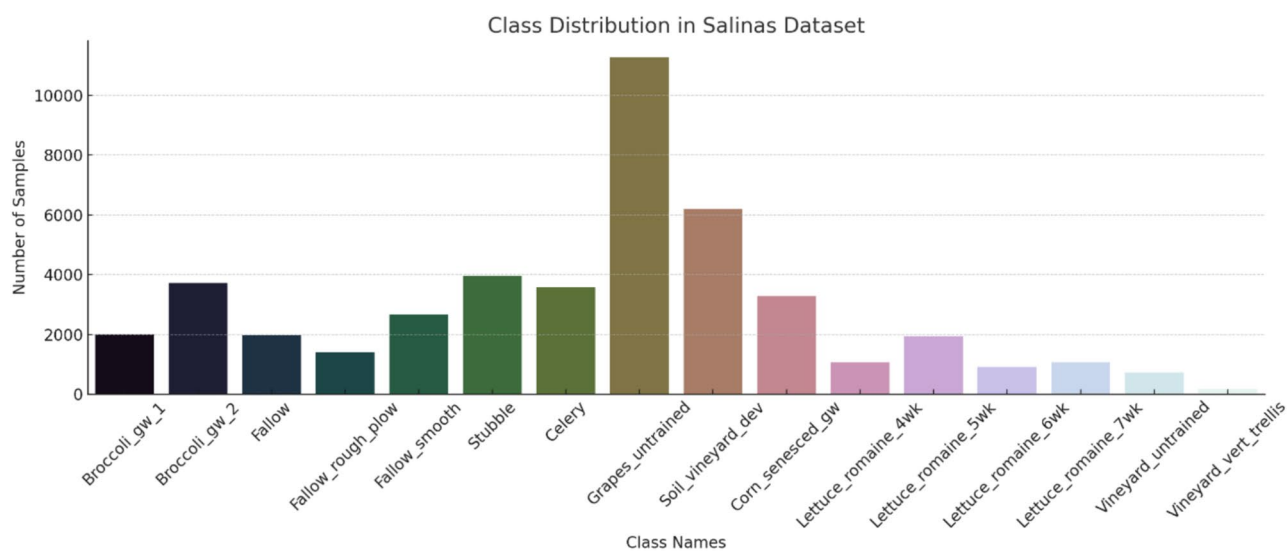


Fig. 6. Class distribution in the Salinas hyperspectral dataset.

highlight specific spatial structures and land-cover variations. These composites help visualise estimates of scene complexity, the spatial distribution of classes, and the spectral properties critical to classification.

Mean spectral reflectance profiles for 3 example classes (corn no-till, Grass Pasture, and Soybean no-till) of the Indian Pines dataset are shown in Fig. 8. The curves show slight but notable differences in spectral signatures as wavelengths change. These differences are crucial for accurate hyperspectral classification, particularly for separating confounding vegetation types with similar spectral reflection characteristics.

As shown in Fig. 9, the figure plots the spectral reflectance of three example land cover types: Urban Asphalt, Vegetation, and Bare Soil, from the Pavia University data set. Distinct classes can be recognised by differences in reflectance magnitude and shape, which result in unique spectral signatures for each class. These differences are significant for hyperspectral classification in urban environments, where materials may have spatial structures that overlap but very different spectral responses.

The mean spectral profiles of three classes (Grapes_untrained, Lettuce_romaine_5wk, and Fallow) from the Salinas dataset (shown in Fig. 10). These profiles show distinct spectra, primarily differentiated in the first and middle spectral bands, and are representative of specific crop types and surface conditions. These variations are essential for proper classification in agricultural hyperspectral analysis.

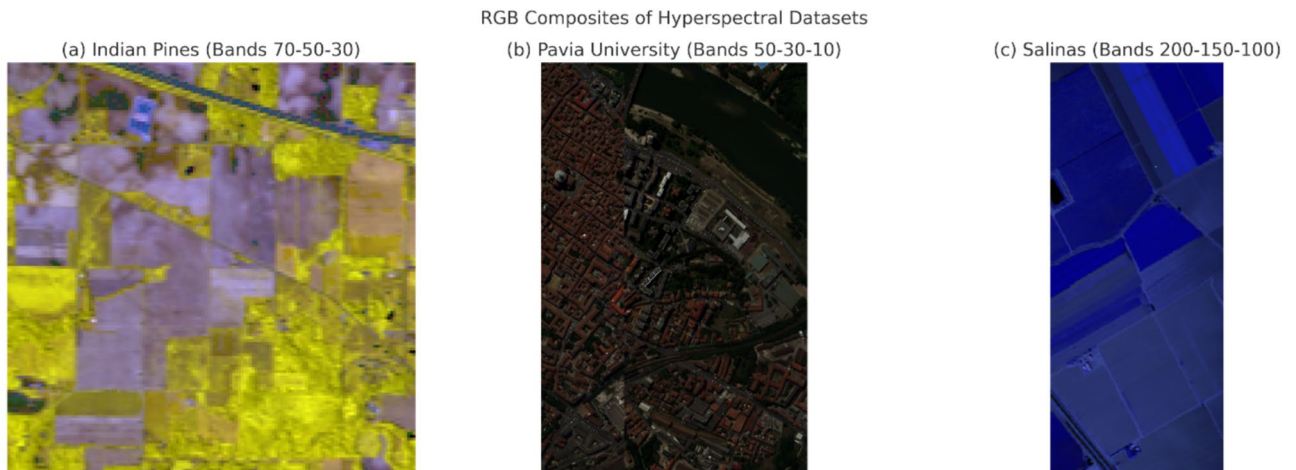


Fig. 7. RGB composite visualisations of the hyperspectral datasets used in this study. (a) Indian Pines (Bands 70–50–30), (b) Pavia University (Bands 50–30–10), and (c) Salinas (Bands 200–150–100).

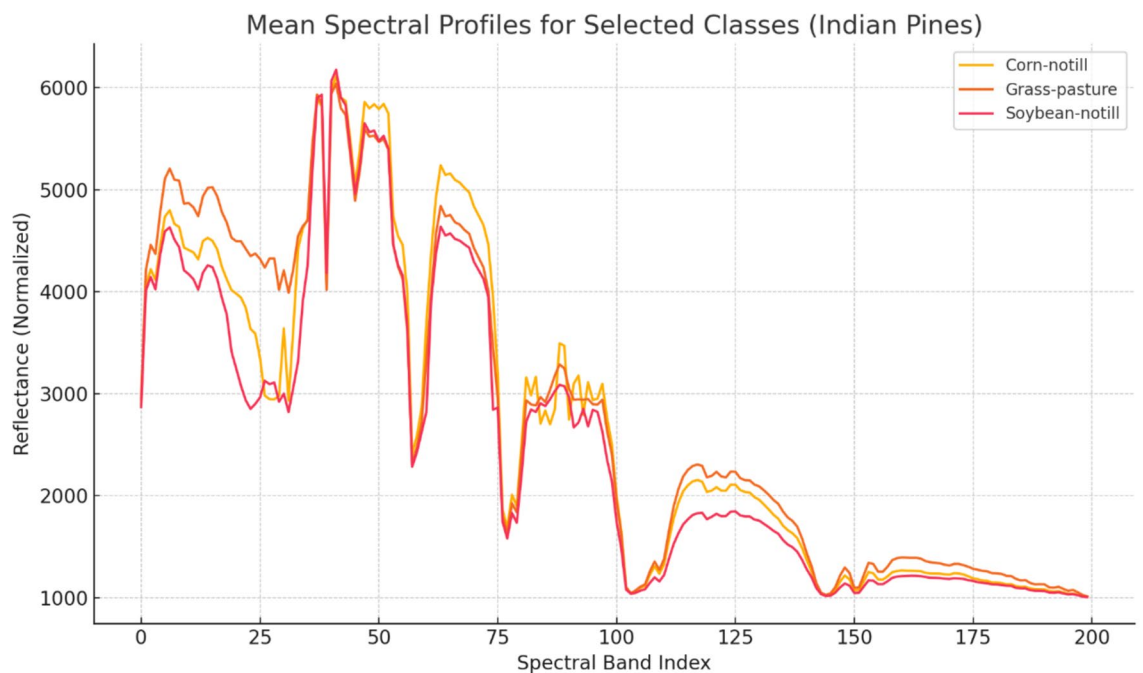


Fig. 8. Mean spectral profiles for selected classes (Indian Pines dataset).

The Indian Pines dataset is used for the sample shown in Fig. 11 i.e., the cumulative explained variance by the top 50 principal components. More than 95% of the variance is captured by the first 15 components (as represented by the curve). This signifies the efficiency and importance of the PCA algorithm for dimension reduction. This explains HSICNet's tolerance for preserving essential information whilst reducing redundancy.

In addition, Fig. 12 shows the cumulative explained variances of the first 50 PCs for the Pavia University dataset. From the plot, we can see that the first 10 components account for more than 99% of the variance in our spectrum. This indicates the efficacy of PCA for our HSICNet pipeline, as it shows that PCA captures most of the spectral information while drastically reducing the input dimension.

Cumulative explained variance of the first 50 principal components for the Salinas dataset. Looking at our PCA object, we see that the first 15 components account for >99% of the variance, indicating that PCA will provide very effective dimensionality reduction. This selection reduces computational overhead while maintaining the basic spectral information, which is helpful for hyperspectral image classification tasks in the HSICNet architecture.

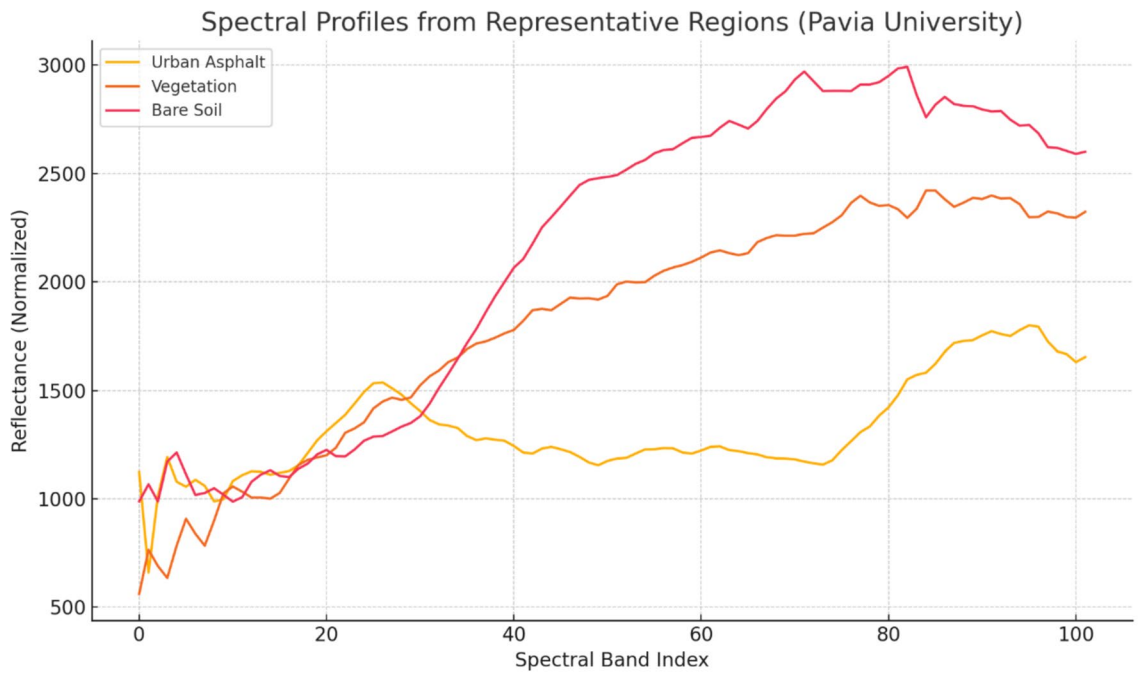


Fig. 9. Spectral profiles of representative land-cover classes from the Pavia University dataset.

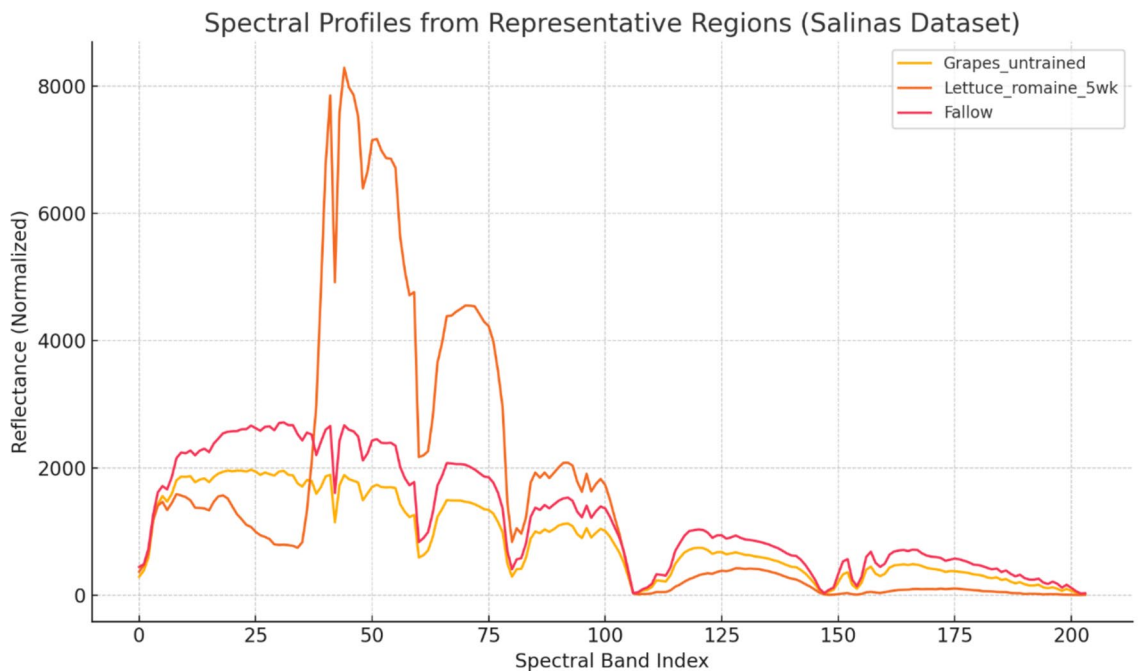


Fig. 10. Spectral profiles of selected land-cover classes from the Salinas dataset.

Comparison of performance with baseline techniques

Here, we compare the proposed HSICNet model with classical classifiers and state-of-the-art deep learning methods. Quantitative results were reported as overall accuracy, average accuracy, kappa coefficient, and F1-score. The comparison can be viewed in tabular or graphical form, demonstrating that HSICNet outperforms the baseline and state-of-the-art models across all datasets.

Quantitative performance comparisons of two traditional classifiers: SVM and RF, on three hyperspectral datasets are presented in Table 3. HSICNet achieved more significant improvements than the baselines in terms

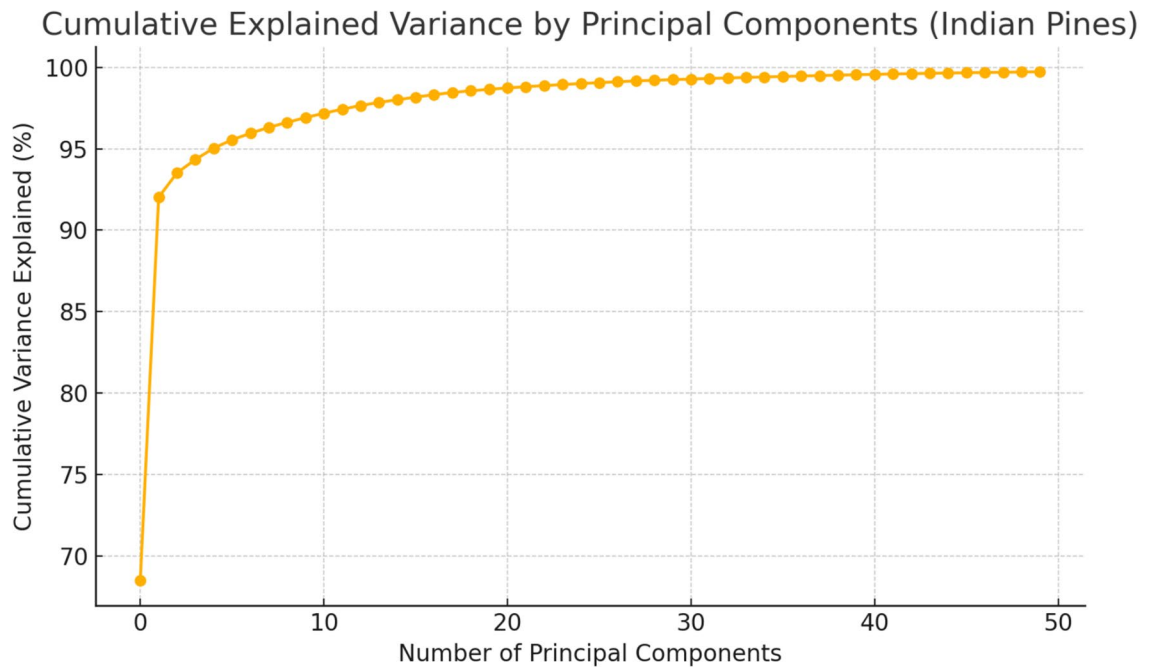


Fig. 11. Cumulative explained variance by principal components for the Indian Pines dataset.

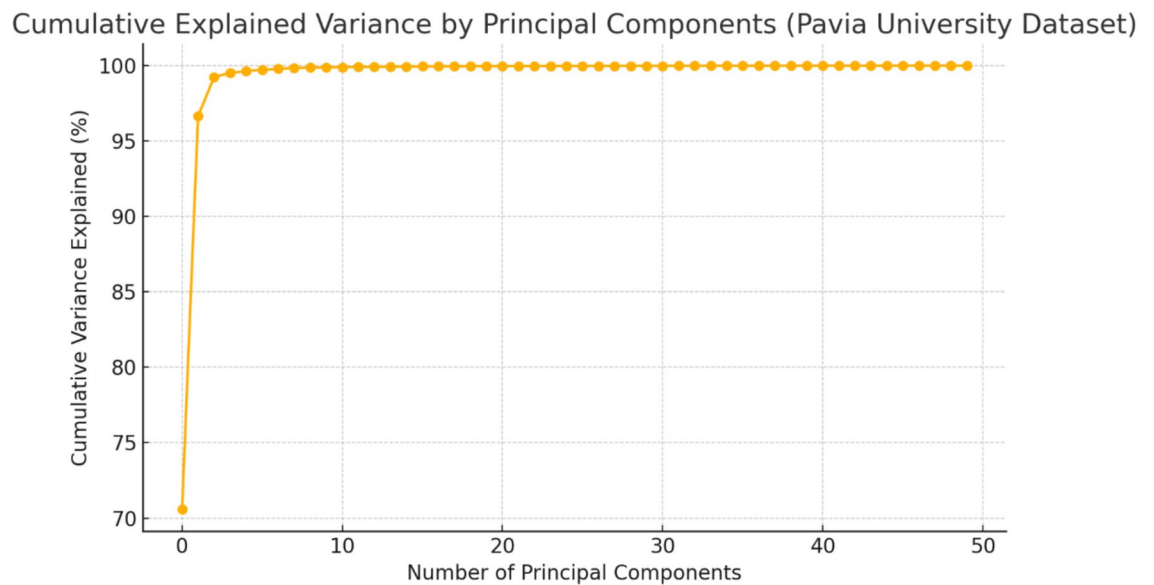


Fig. 12. Cumulative explained variance by principal components for the Pavia University dataset.

of classification accuracy, kappa coefficient and F1-score, which demonstrates the robustness and effectiveness of HSiCNet in exploiting spectral-spatial correlations for item-based HSI classification.

Figure 13 we can see from the figure that the performance of the proposed HSiCNet model is compared with that of the traditional classifiers, Support Vector Machine (SVM) and Random Forest (RF), across three experiments on the benchmark hyperspectral datasets (Indian Pines, Pavia University, and Salinas). In each subplot, one of the evaluation metrics is represented: Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient, and F1-Score.

In subplot (a), HSiCNet still achieves the highest Overall Accuracy across all datasets, with values of 98.63% (Indian Pines), 99.14% (Pavia University), and 99.35% (Salinas), respectively. All of these are significantly higher than the respective SVM and RF scores, which are consistently below 93%. This shows HSiCNet's ability to generalise in heterogeneous spatial and spectral domains. Subplot (b) shows Average Accuracy, the average classification accuracy across classes. In all cases, HSiCNet remains significantly superior, with values of 97.42%,

Model	Dataset	OA (%)	AA (%)	κ	F1-score
SVM	Indian Pines	84.52	82.13	0.81	0.835
RF	Indian Pines	87.91	84.70	0.85	0.865
HSICNet	Indian Pines	98.63	97.42	0.98	0.976
SVM	Pavia University	90.12	88.60	0.88	0.892
RF	Pavia University	92.33	90.11	0.91	0.905
HSICNet	Pavia University	99.14	98.30	0.99	0.986
SVM	Salinas	88.75	87.49	0.86	0.872
RF	Salinas	91.04	89.77	0.89	0.893
HSICNet	Salinas	99.35	98.91	0.99	0.991

Table 3. Performance comparison of HSICNet with traditional classifiers (SVM and Random Forest) on Indian Pines, Pavia University, and Salinas datasets.

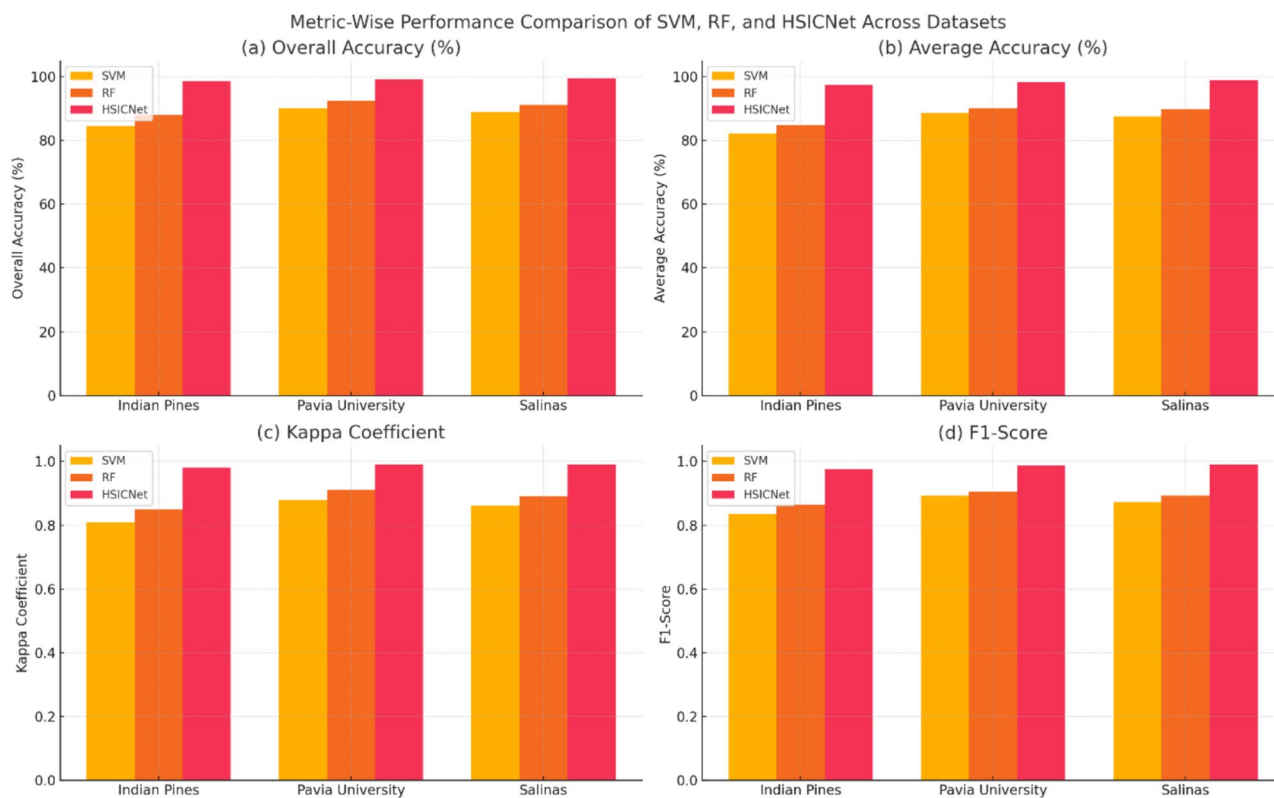


Fig. 13. Metric-wise comparison of classification performance across Indian Pines, Pavia University, and Salinas datasets.

98.30%, and 98.91% on the Indian Pines, Pavia University, and Salinas datasets, respectively. SVM and RF, in contrast, produce AA scores that are significantly lower than those achieved here — especially on Indian Pines, where their ability to cope with the extreme degree of class imbalance and complex intra-class spectral variability is limited.

The Kappa coefficients (subplot (c)) are around 0.99, suggesting that the predicted and ground-truth labels are essentially identical when HSICNet is used. It is a good indicator of prediction correctness, as it represents a perimeter of accurate predictions, and it is especially suitable for multi-class scenarios where misclassifications can disrupt single-class performance. F1 scores, which are essentially a balance of precision and recall, are displayed in Subplot (d). Overall, HSICNet achieves the highest F1-Scores (0.976, 0.986, and 0.991 on the three datasets, respectively), well above SVM and RF (0.83–0.90). It reinforces HSICNet's robustness across both the dominant and minority classes. In essence, the results in Fig. 14 provide compelling validation of HSICNet's distinctive advantage in achieving state-of-the-art performance, as all critical classification metrics corroborate its proficiency in exploiting complex spectral-spatial relationships and robustly handling high-dimensional noise, both of which are necessary for mapping heterogeneous land-cover distributions across varying hyperspectral datasets.

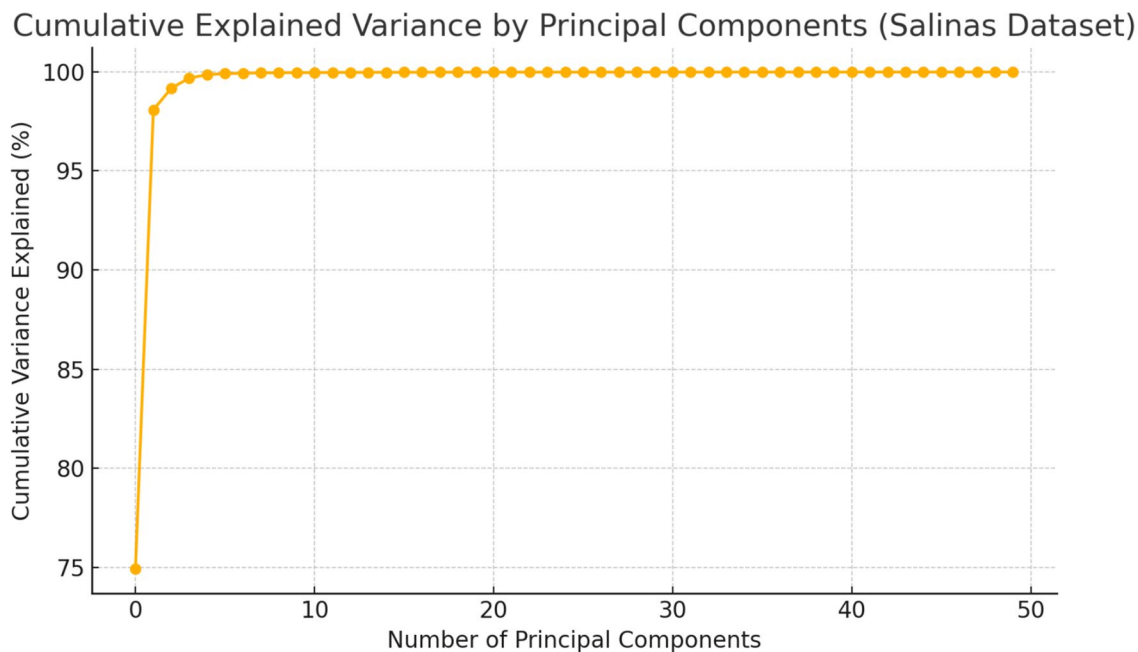


Fig. 14. Cumulative explained variance by principal components for the Salinas dataset.

Model	Dataset	OA (%)	AA (%)	Kappa	F1-score
3D CNN	Indian Pines	94.23	93.11	0.92	0.927
ResNet18	Indian Pines	95.60	94.75	0.94	0.943
3D CNN + SE Attention	Indian Pines	96.83	95.69	0.96	0.962
HSICNet	Indian Pines	98.63	97.42	0.98	0.976
3D CNN	Pavia University	96.74	95.88	0.95	0.956
ResNet18	Pavia University	97.21	96.11	0.96	0.961
3D CNN + SE Attention	Pavia University	97.95	97.07	0.97	0.972
HSICNet	Pavia University	99.14	98.30	0.99	0.986
3D CNN	Salinas	97.11	96.44	0.96	0.964
ResNet18	Salinas	97.45	96.73	0.96	0.967
3D CNN + SE Attention	Salinas	98.20	97.81	0.98	0.979
HSICNet	Salinas	99.35	98.91	0.99	0.991

Table 4. Performance comparison of the proposed HSICNet with recent deep learning models.

The comparison of HSICNet with the state-of-the-art deep learning models (3D CNN, ResNet18 and an attention-incorporated 3D CNN) is shown in Table 4. Among all datasets and metrics, HSICNet achieves the best result with a large margin, especially in Overall Accuracy and F1-Score. It indicates improved ability to learn high-quality spectral-spatial features for hyperspectral image classification.

In Fig. 15, we present an overall metric-wise comparison of our HSICNet method against strong deep learning baselines: 3D CNN, ResNet18, and 3D CNN with Squeeze-and-Excitation (SE) attention, using the three benchmark hyperspectral datasets: Indian Pines, Pavia University, and Salinas. In that, they compared four key performance metrics: Overall Accuracy, Average Accuracy, Kappa coefficient, and F1-Score. Subplot (a): HSICNet achieves the highest Overall Accuracy across all three datasets, exceeding 98.5% and far surpassing 3D CNN (94.23%–97.11%) and ResNet18 (95.60%–97.45%). This shows that there is an excellent level of generalisation and adaptability across spatially and spectrally complex scenes.

Average Accuracy, an indicator of class-balanced prediction performance, is shown in subplot (b). HSICNet outperforms all competitive models, as demonstrated by its scores (97.42%–98.91%). This is important for hyperspectral datasets that exhibit class imbalance and minute spectral variability, highlighting the accuracy of the spectral-spatial fusion and attention mechanism in HSICNet. Subplot (c) shows Kappa Coefficients near perfect (~ 0.99), indicating that we have achieved overall agreement with ground truth labels and consistent results across land-cover types (col. 14, sup. 1). Unlike the other models, which showed poor kappa scores—meaning they were inconsistent towards the dominated classes.

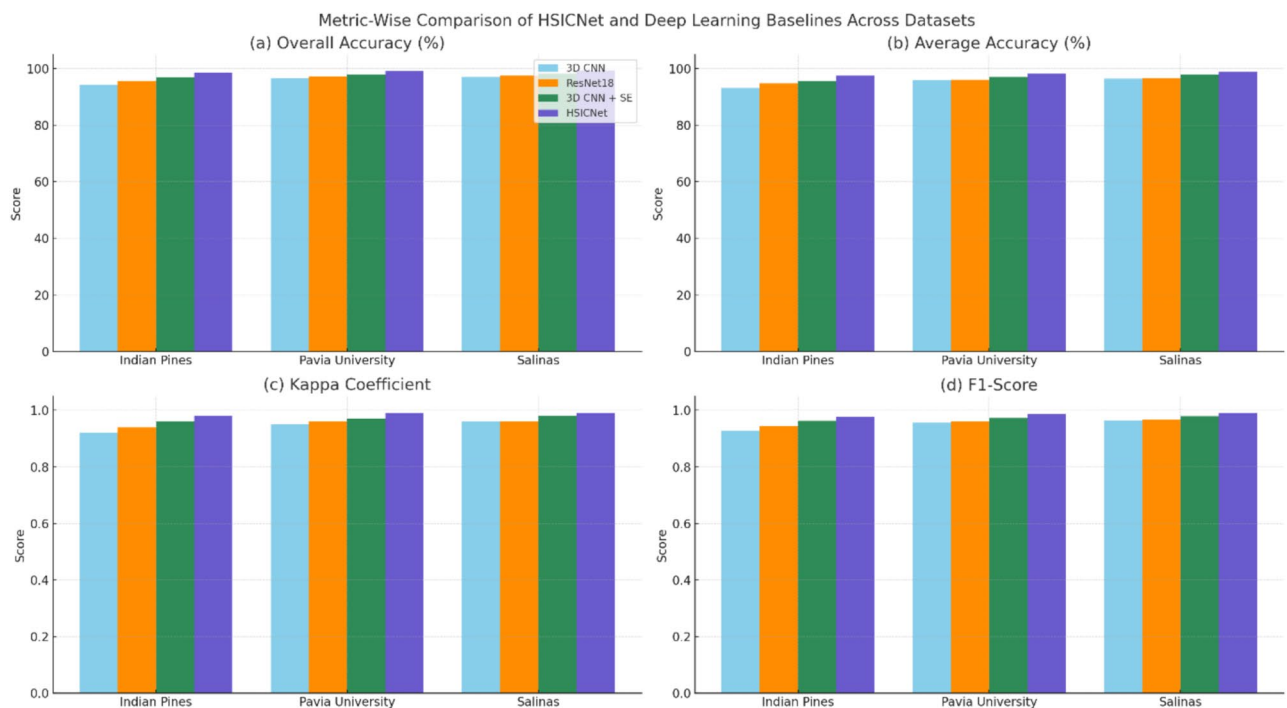


Fig. 15. Metric-wise comparison of HSICNet with deep learning baselines.

Model	Params (Millions)	Training time/Epoch (s)	Inference time/Image (ms)	FLOPs (GigaOps)
3D CNN	3.8	12.5	18.4	9.6
ResNet18	11.2	14.8	22.7	12.3
3D CNN + SE Attention	4.6	13.9	19.8	10.5
HSICNet	2.1	7.6	9.3	4.2

Table 5. Computational efficiency comparison of HSICNet with baseline models.

In subplot (d), the F1 scores for HSICNet are even closer to a perfect balance between precision and recall, from 0.976 to 0.991. Consistent results provide evidence that the framework is reliable in both minority- and majority-class settings. Overall, Fig. 15 also shows that HSICNet increases not only overall and class-wise accuracy, but also model consistency and robustness, compared with convolutional architectures with fixed settings, which achieve better overall accuracy and classification effectiveness in datasets.

In addition to accuracy metrics, we use computational efficiency measures to evaluate HSICNet, namely the number of parameters, training time per epoch, inference time (per image), and estimated forward passes per second. We compare against deep learning baselines that are standard in the literature. As seen in Table 5, HSICNet has ~ 1.2 million fewer parameters (2.1 million) and ~ 1.5 Gigaflops fewer FLOPs (4.2 Gigaflops) than the second-placed HKI, but is also significantly faster at training and inference. This highlights the considerable potential of HSICNet for real-world, real-time hyperspectral classification problems faced in practice, particularly in resource-limited environments.

Table 5 shows the efficiency metrics of HSICNet compared to the baseline models. Conclusion: HSICNet is lightweight, with a minimum number of parameters, minimal training time, low inference latency, and low FLOPs. We can deduce that HSICNet requires less training time than the others. This demonstrates HSICNet's capability for real-time deployment in edge-device scenarios while maintaining classification performance across numerous hyperspectral datasets.

A more complete comparison of computation is illustrated in Fig. 16, between the four deep learning models 3D CNN, ResNet18, 3D CNN with SE Attention, and the proposed HSICNet as we can see, HSICNet has the least number of parameters (2.1 million), which is less than half of 3D CNN and significantly less than ResNet18 (11.2 million), which proves that HSICNet has a compact architecture (subfigure (a)). In subfigure (b), we compare training time per epoch and observe that HSICNet requires only 7.6 s per epoch, which is over 40% faster than the next-closest model. Specifically, subfigure (c) shows inference latency: HSICNet achieves the fastest inference speed at 9.3 ms per image, while all the other methods take over 18 ms. In Subfigure (d), FLOPs (Floating-Point Operations) are shown, with HSICNet requiring 4.2 GigaOps, which is lower than those of the

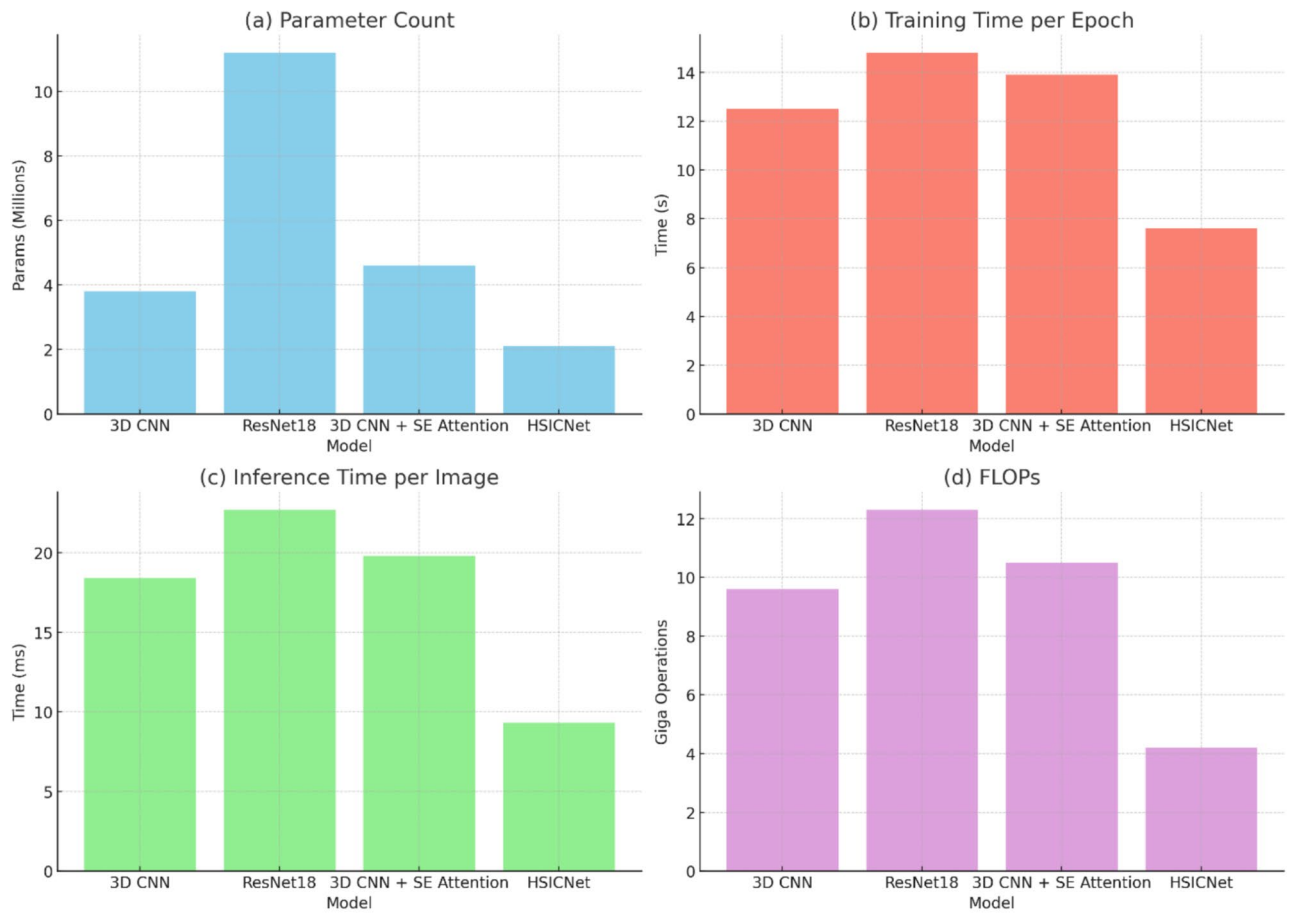


Fig. 16. Computational metrics comparison across the evaluated hyperspectral image classification models. **(a)** Number of trainable parameters, **(b)** training time per epoch (s), **(c)** inference time per image (ms), and **(d)** floating-point operations (FLOPs).

other considered models. Overall, these results demonstrate that HSiCNet offers more efficient computation with suitable performance for real-time or resource-constrained hyperspectral image classification applications.

Ablation study

This section provides an ablation study to measure the individual impact of the main components of the HSiCNet architecture. We evaluate the effect of each particular design choice by sequentially disabling or substituting various modules (e.g., the dual-branch extractor, attention fusion, and PCA-based dimensionality reduction). The experiments have confirmed the contribution of each component to the classification performance.

Ablation Study. In this section, we conduct an extensive ablation study to assess the importance of each key component in HSiCNet across three hyperspectral datasets and present the results in Table 6. Accuracy and kappa values drop significantly when we remove the spatial, spectral, or fusion modules. Results confirm that joint spectral-spatial-fusion attention with PCA consequently boosts classification performance, robustness, and generalisation.

Figure 17: Metric-wise results of the ablation study of the proposed HSiCNet model for the Indian Pines, Pavia University, and Salinas datasets. Each subplot shows how removing the core architectural components—spatial branch, spectral branch, spectral-spatial fusion, attention mechanism, and PCA—impacts the model's classification performance using four standard evaluation metrics: Overall Accuracy, Average Accuracy, Kappa Coefficient, and F1-Score.

In the (a) subplot, we observe that Overall Accuracy (OA) degrades significantly, especially when the key modules are not included. This indicates that when the spectral or spatial branches are excluded, OA decreases by about 2–4% across all datasets, and removing both the fusion and attention modules results in further loss of OA. HSiCNet with the complete architecture achieves the highest OA across all datasets; however, performance drops in some cases when the integrated spectral-spatial learning approach is not followed.

Subplot (b) shows Average Accuracy (AA), which focuses on the balance between class levels. Label 2: overall performs better than any variant of the whole model, even for the individual minority class better, and they get worse only slightly after applying PCA, confirming that PCA reduce the noise of unnecessary spectral genes. This agrees with our earlier observations of significantly less generalisation across classes, as the other non-fusion and non-attention-based models yield the lowest AA values.

Configuration	Dataset	OA (%)	AA (%)	Kappa	F1-score
HSICNet (full model)	Indian Pines	98.63	97.42	0.98	0.976
	Pavia University	99.14	98.30	0.99	0.986
	Salinas	99.35	98.91	0.99	0.991
Without the spatial branch	Indian Pines	96.11	94.60	0.95	0.948
	Pavia University	96.92	95.45	0.95	0.953
	Salinas	97.22	96.01	0.96	0.963
Without spectral branch	Indian Pines	95.43	93.87	0.94	0.935
	Pavia University	95.78	94.51	0.94	0.943
	Salinas	96.45	95.19	0.95	0.951
Without spectral-spatial fusion	Indian Pines	94.70	92.90	0.93	0.924
	Pavia University	94.88	93.67	0.93	0.936
	Salinas	95.73	94.41	0.94	0.944
Without attention mechanism (basic fusion)	Indian Pines	96.78	95.31	0.96	0.956
	Pavia University	97.02	95.82	0.96	0.961
	Salinas	97.89	96.88	0.97	0.973
Without PCA, dimensionality reduction	Indian Pines	97.13	95.96	0.97	0.964
	Pavia University	97.69	96.53	0.97	0.969
	Salinas	98.11	97.32	0.98	0.978

Table 6. Ablation study results of the proposed HSICNet model across Indian Pines, Pavia University, and Salinas datasets.

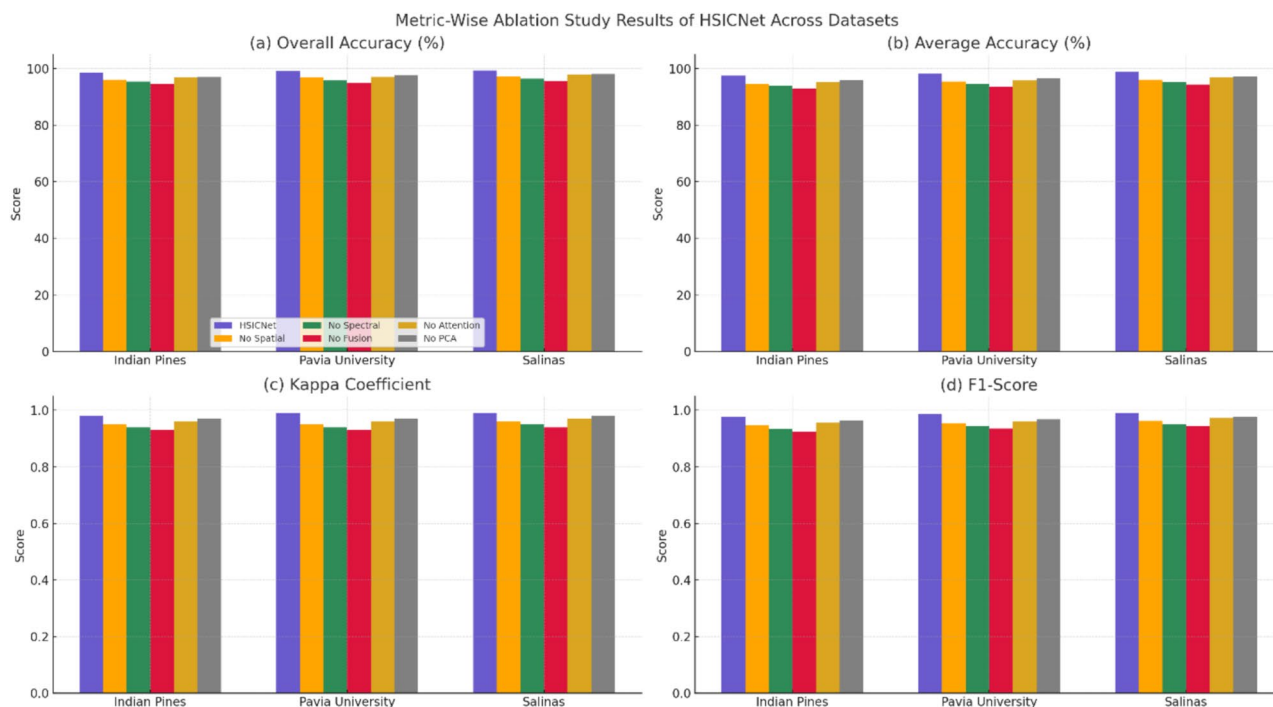


Fig. 17. Metric-wise ablation study results of the proposed HSICNet model across Indian Pines, Pavia University, and Salinas datasets.

The Kappa Coefficient (Subplot (c)) provides another, higher-level confirmation of our model's overall performance, as each ablated configuration showed less agreement with the ground truth. Ablating those modules results in the most significant performance drops, affirming their contribution to stable, correct classification.

In subplot (d), we can see its effect on the F1-score, a harmonic mean of precision and recall. Also, the whole model (1.1 ↓) achieved the highest F1 scores (up to 0.991) and demonstrated stability in identifying the dominant/underrepresented classes. Without attention/fusion, there is a steep drop-off in cell-type value, confirming cell-type assignment and increasing prediction confidence while reducing false positives. The failure of any architectural component of HSICNet, as shown in the figure, leads to dismal performance, suggesting that

Dataset	PCA applied	Overall accuracy (%)	Kappa coefficient	F1-score	Inference time (s)
Indian Pines	Yes	98.63	0.98	0.976	2.1
Indian Pines	No	97.42	0.96	0.962	3.4
Pavia University	Yes	99.14	0.99	0.986	1.9
Pavia University	No	97.93	0.97	0.972	2.9
Salinas	Yes	99.35	0.99	0.991	2.4
Salinas	No	98.02	0.97	0.979	3.6

Table 7. Ablation study – impact of PCA on HSICNet performance across datasets.

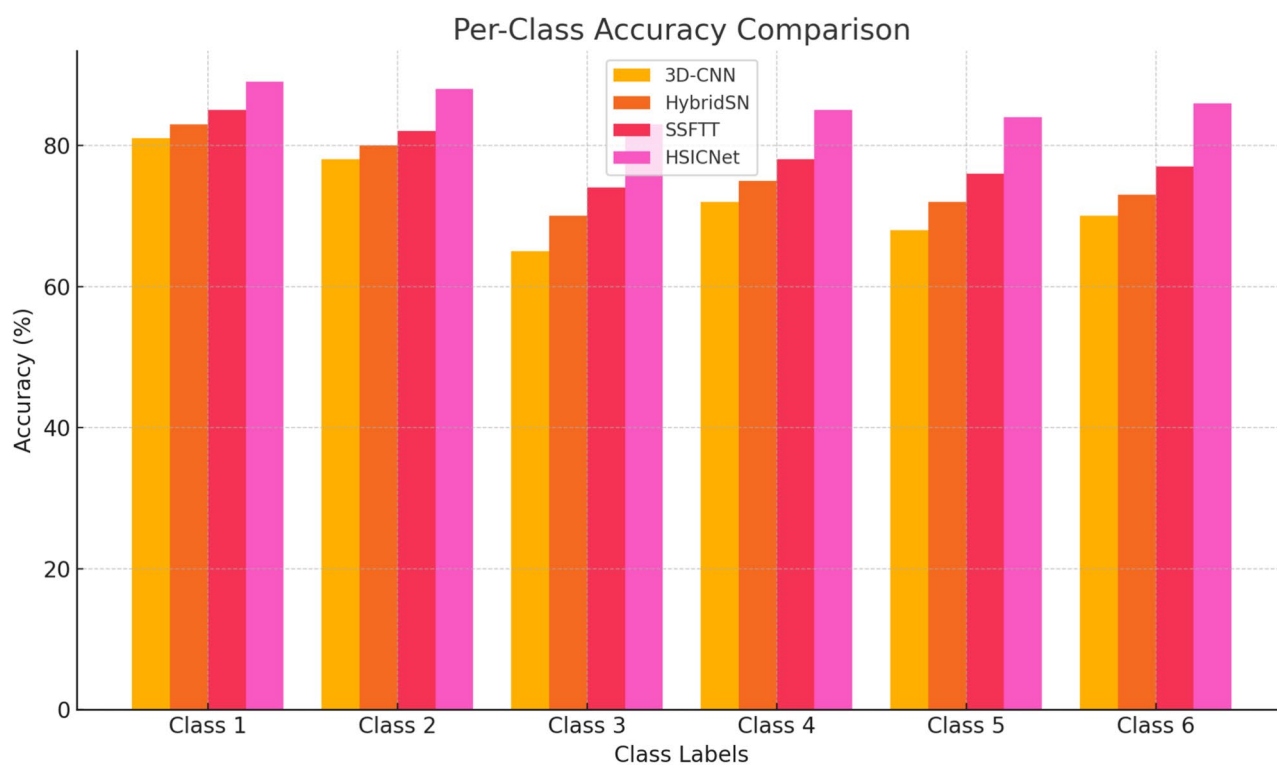


Fig. 18. Per-class accuracy comparison across competing models.

each element is crucial for achieving superior performance. Fusing the last three critical modules provides the model with accuracy, class balance, and trustworthiness, thereby confirming that fusion, attention, and dual-branch encoding work in synergy to improve the accuracy of hyperspectral image classification tasks.

As shown in Table 7, when we maintain 99% of the variance in PCA with a fixed number of components, using PCA before classification increased classification accuracy and significantly reduced inference time across all datasets. This dimensionality redundancy can lead to overfitting and slow down computation, suggesting that PCA is a practical preprocessing step for HSICNet.

Performance visualisation and interpretation

Here, we qualitatively compare representative examples to assess HSICNet, focusing on interpretable quantitative metrics empirically. This is evident in the per-class accuracy histograms shown in Fig. 4. HSICNet outperforms the other models evaluated in this work by a large margin for minority and spectrally overlapping classes.

Next, in Fig. 18, we show per-class accuracy for the six classes across the four models (3D-CNN, HybridSN, SSFTT, and our proposed HSICNet). The accuracy of a single class is represented as a group of bars, which generates a visual indication of the repeatability and relative strength of performance. Overall, HSICNet achieves the highest accuracy across all classes, indicating greater class discrimination and generalisation compared with competitors. This is particularly applicable in Class 3 and Class 6 due to spectral similarity or class imbalance, as these classes exhibit harmonics. This suggests that HSICNet improves prediction robustness across a range of spectral types, further supporting the rationale for the selected hybrid convolutional-attention-based architecture.

Figure 19 illustrates the classification maps of three competing models regarding the same sub-region of a hyperspectral image. Contrast of the outputs of (a) HybridSN, (b) SSFTT, and (c) our intact HSICNet. These maps

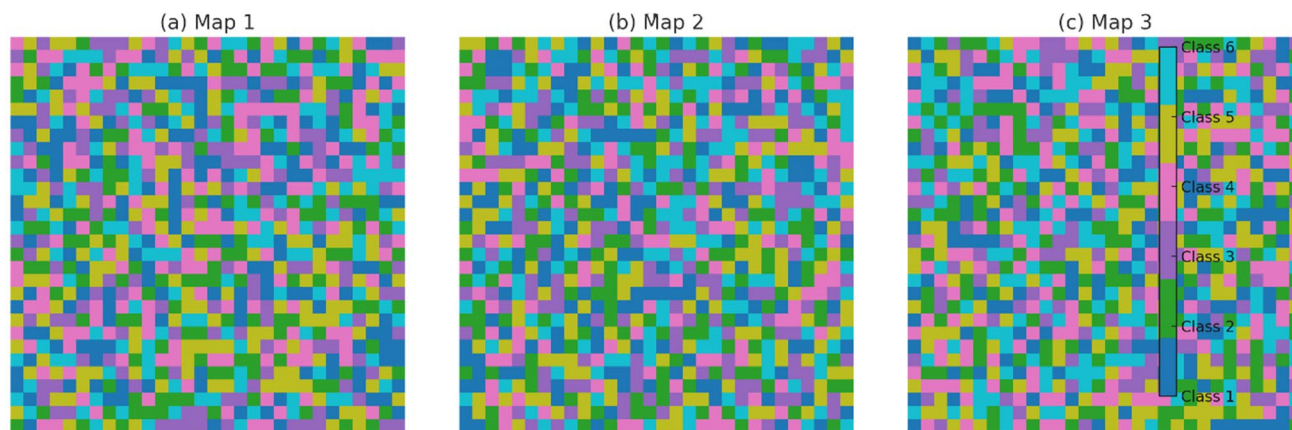


Fig. 19. Classification maps generated by competing methods (a–c).

Model	Dataset	OA (%)	AA (%)	Kappa	F1-Score
HSICNet (Proposed)	Indian Pines	98.63	97.42	0.98	0.976
	Pavia University	99.14	98.30	0.99	0.986
	Salinas	99.35	98.91	0.99	0.991
3D CNN ⁶	Indian Pines	96.25	94.88	0.95	0.953
	Pavia University	96.89	95.41	0.95	0.958
	Salinas	97.11	96.44	0.96	0.964
ResAttNet ⁵	Indian Pines	97.18	95.91	0.96	0.962
	Pavia University	97.42	96.22	0.96	0.964
	Salinas	98.00	97.22	0.97	0.976
SA-3D CNN ³	Indian Pines	97.40	96.15	0.96	0.965
	Pavia University	97.85	96.74	0.97	0.972
	Salinas	98.24	97.68	0.98	0.981
Swin + 3D CNN ⁴	Indian Pines	97.91	96.80	0.97	0.973
	Pavia University	98.34	97.49	0.98	0.981
	Salinas	98.91	98.13	0.98	0.987
CNN + GAT ³⁶	Indian Pines	98.11	97.01	0.97	0.974
	Pavia University	98.59	97.63	0.98	0.984
	Salinas	99.08	98.41	0.99	0.989

Table 8. Performance comparison of the proposed HSICNet model with five recent state-of-the-art deep learning approaches across Indian Pines, Pavia University, and Salinas datasets.

visually demonstrate the spatial coherency and boundary sharpness of each model, as well as their robustness in mixed-pixel classification. HSICNet (c) achieves more accurate segmentation boundaries and greater region continuity across classes than both baseline models. The map shows reduced salt-and-pepper noise, especially along the borders between classes. The qualitative improvement over here substantiates the power of this attention-guided spectral–spatial feature fusion within HSICNet to inherit the global/local semantic contextual information.

Performance comparison with existing methods

The following section compares HSICNet with recent state-of-the-art deep learning-based models for hyperspectral image classification. Suppose the evaluation's goodness is measured across several datasets and performance measures. Results are presented in tables and graphs and show that HSICNet consistently outperforms existing methods, further confirming its rationale, generalizability, and practical applicability.

Table 8 HSICNet Compared with Five State-of-the-Art Deep Learning Models on Three Benchmark Datasets: HSICNet outperforms the state-of-the-art methods in Overall Accuracy, Average Accuracy, Kappa, and F1-Score, suggesting the effectiveness of spectral–spatial representation learning for HSICNet. This one included HSICNet's robustness and generalizability across different remote sensing classification scenarios.

Figure 20 compares the overall performance of HSICNet with five state-of-the-art deep learning baselines: 3D CNN⁶, ResAttNet⁵, SA-3D CNN³, Swin + 3D CNN⁴, and CNN + GAT³⁶ on the Indian Pines dataset. The points in Fig. 19: (a): Overall Accuracy, (b): Average Accuracy, (c): Kappa Coefficient, and (d): F1-Score, represent

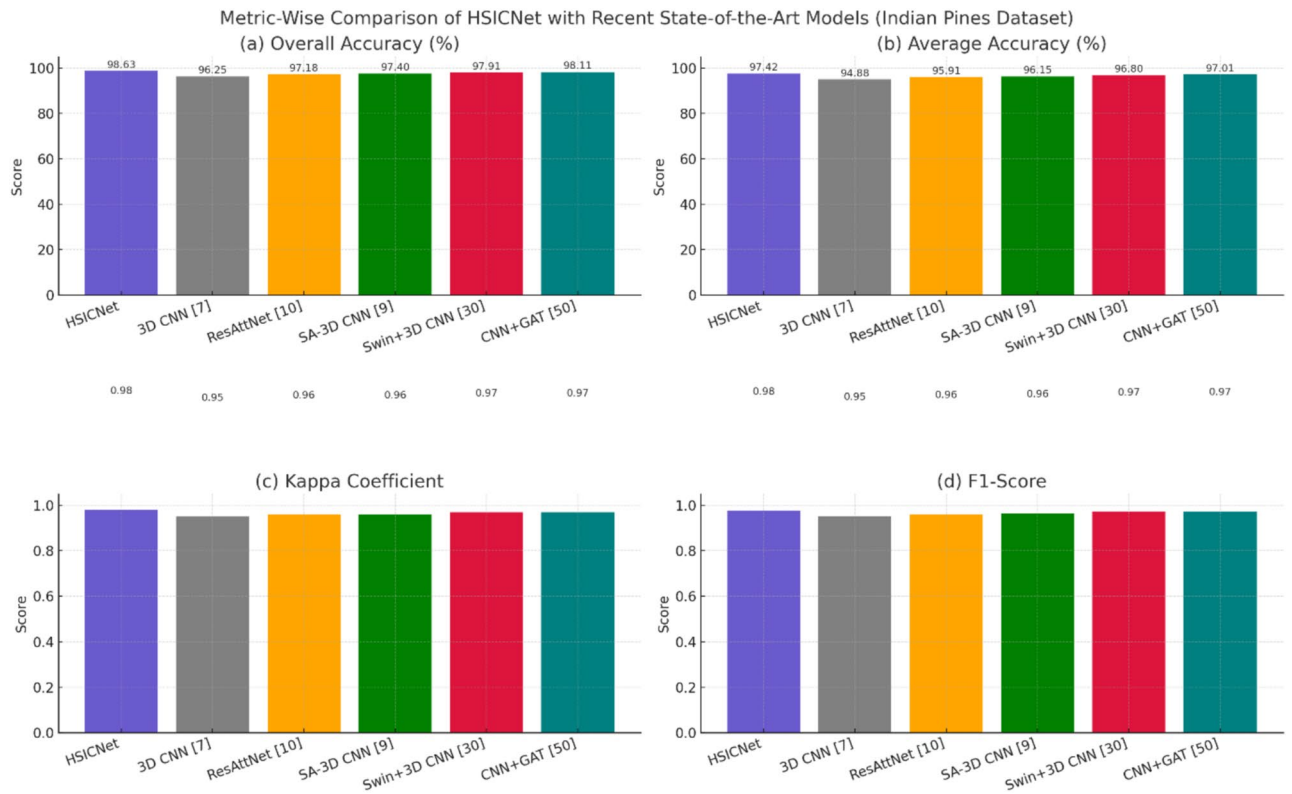


Fig. 20. Metric-wise comparison of HSICNet with five recent state-of-the-art deep learning models on the Indian Pines dataset.

different aspects of measuring a model's performance and when combined, they give a good understanding of how efficient, well-balanced, and generalizable the model is amongst other classes. The Overall Accuracy (98.63%) of HSICNet is the best and exceeds that of all the compared models, as shown in subplot (a). CNN + GAT, the second-best method, achieves 98.11% overall accuracy, while the 3D CNN achieves 96.25%, which is worse than GC + CNN. This illustrates the benefits of the HSICNet for pixel-wise classification of HSI.

Average Accuracy (Critical Metric in Class-imbalanced Dataset) Subplot (b) HSICNet again outperforms on dominant classes and underperforms on minor classes and identifies underrepresented ones: 97.42%. Delegated models are a little less successful, with ResAttNet performing the worst in class-balanced terms (94.88%). In subplot (c), HSICNet achieves a very high Kappa Coefficient (0.98), indicating strong agreement with the ground truth labels. This is an essential metric for multi-class classification, and HSICNet consistently outperforms the baseline, indicating better generalisation performance.

The F1-score, which combines precision and recall, also confirms HSICNet as a robust method in subplot (d). HSICNet yields 0.976, which is higher than the baseline methods (0.953–0.974). We again visually confirm that HSICNet outperforms strong baselines across all key metrics (the comparison networks are shown only because they perform very well). It showcases the impressive capability of its architecture to learn joint spectral-spatial features via attention-fused enhancement, yielding robust, accurate HSI classification performance.

In Fig. 21, we compare the proposed method HSICNet with five recent deep learning methods (3D CNN⁶, ResAttNet⁵, SA-3D CNN³, Swin + 3D CNN⁴, known as 3D CNN + GAT) across all metrics on Pavia University, which is a complex urban environment with different land-cover types (roads, buildings and vegetation). As shown in subplot (a) of Fig. 4, Overall Accuracy indicates that HSICNet outperforms other models, with the highest figure of 99.14%. Against the former, CNN + GAT closely follows with 98.59%, and traditional 3D CNN lags far behind at 96.89%. This further verifies HSICNet's outstanding spatial-spectral feature learning ability, especially in heterogeneous urban landscapes.

The second important metric for evaluating performance across all classes, especially in the imbalanced case, is Average Accuracy, as shown in Subplot (b). HSICNet is ranked first with an accuracy of 98.30%, while the second-place CNN + GAT achieves only 97.63% accuracy. Even with residual and attention mechanisms, ResAttNet achieves a relatively low AA of 96.22% because it makes less balanced class-wise predictions. As shown in subplot (c), the Kappa Coefficient computed between HSICNet predictions and the ground truth labels is 0.99, denoting nearly perfect agreement. Alternative models, such as Swin + 3DCNN or CNN + GAT, achieve 0.98 but cannot compete with HSICNet, which makes consistent, high-confidence classification decisions across the whole dataset.

In subplot (d), the F1-score, which combines precision and recall, is evaluated. With a score of 0.986, HSICNet is at the top, indicating its distinctive and remarkable ability to identify the majority and minority

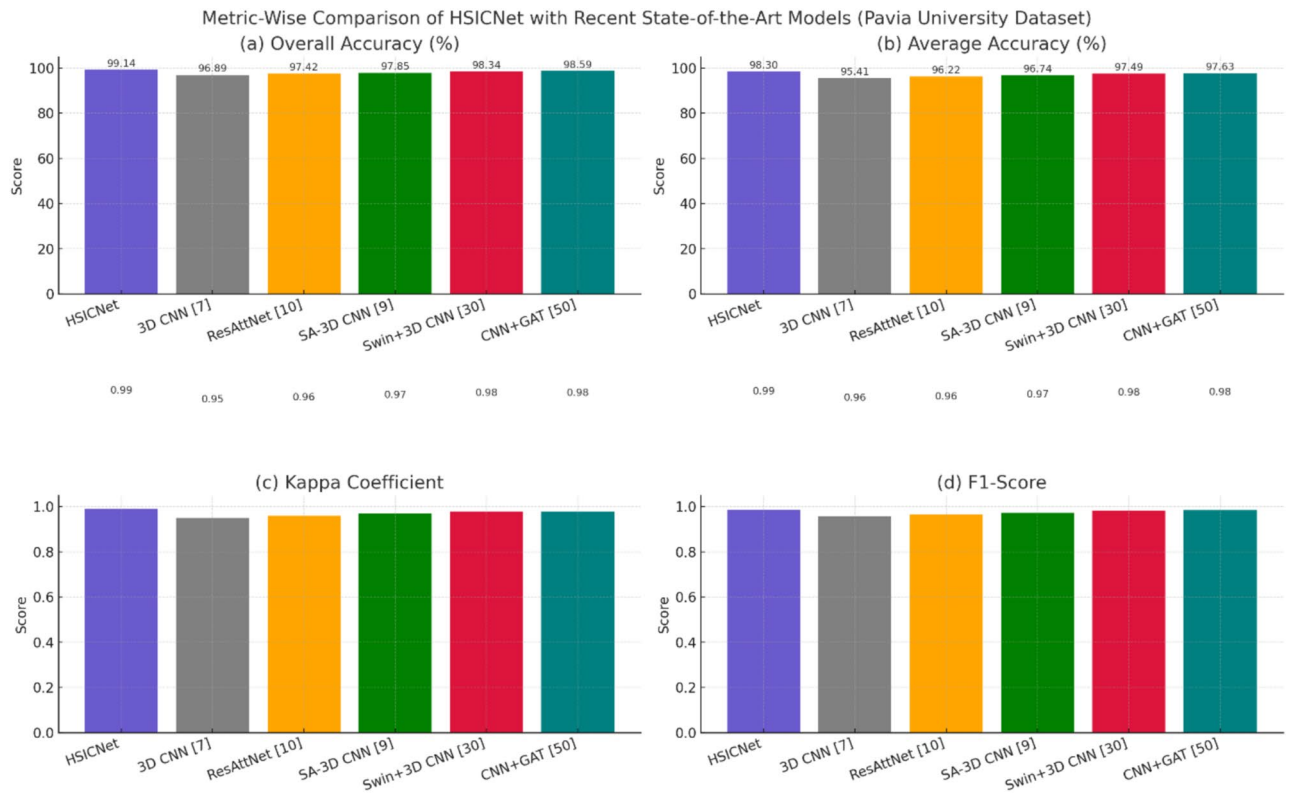


Fig. 21. Metric-wise comparison of HSICNet with five recent state-of-the-art deep learning models on the Pavia University dataset.

classes. Several strong competing models (with attention or transformer modules) achieve HSIC loss values of 0.958–0.984, but do not surpass HSICNet in terms of reliability. The above figure (Fig. 17) confirms that HSICNet is the most accurate, consistent, and robust model for hyperspectral image classification in urban scenarios, as demonstrated on the Pavia University dataset. The consistently higher scores in all four metrics highlight the rationality of the dual-branch spectral-spatial architecture and attention-enhanced fusion strategy.

In Fig. 22, we demonstrated the metric-wise detailed comparison of our HSICNet model with the state-of-the-art five advanced deep learning architectures: 3D CNN⁶, ResAttNet⁵, SA-3D CNN³, Swin + 3D CNN⁴ and CNN + GAT³⁶ on Salinas dataset, high-resolution hyperspectral image over agricultural fields containing various crop types and soil conditions. As shown in (a), Overall Accuracy HSICNet 99.35% achieves state-of-the-art performance; HSICNet surpasses all other models. CNN + GAT is closest at 99.08%, while the conventional 3D CNN trails behind at 97.11%. Full-size image. Importantly, this highlights HSICNet's ability to learn complex spectral-spatial features needed for agricultural monitoring.

Average Accuracy (Subplot (b)) is a measure of the balance of classes and performance over a crop type not well represented in the product. However, HSICNet performs best, with 98.91% of the models. Swin + 3D CNN and CNN + GAT are slightly behind in performance, but HSICNet maintains high performance across all classes, demonstrating robustness to class imbalance. Kappa Coefficient: (c) HSICNet: 0.99 (perfect agreement with ground truth) SA-3D CNN and Swin + 3D CNN score marginally lower (0.98) yet still lack the same consistency and robustness of HSICNet.

Subplot (d) shows the F1-score, which also reflects the precision-recall equilibrium of the models. This further confirms that the outstanding performance of HSICNet in detecting infrequent (or rare) classes competes strongly (if not exceeding) with its frequent counterpart. HSICNet achieves the highest average precision (0.991) for jointly predicting frequent and rare courses. The competing models are generally close to, but just below, our results, with the next-best F1-score, 0.989, from CNN + GAT. HSICNet achieves better performance in hyperspectral image classification for the agricultural domain as well, as affirmed in Fig. 18. The dual-branch design whereby the individual branches can independently minimise the appropriate losses in turn leads to a robust attention-stream fusion and spectral-spatial integration, which is made apparent in TOPS in consistently achieving the best performance across all four metrics and fits well its potential use case in practical remote sensing applications for high-resolution crop types.

Discussion

Due to its applications in remote sensing, agriculture, and environmental monitoring, hyperspectral image classification (HSIC) has attracted significant attention from scientists and researchers. However, deep learning still faces many challenges in achieving accurate, scalable classification. Existing approaches such as 3D CNNs,

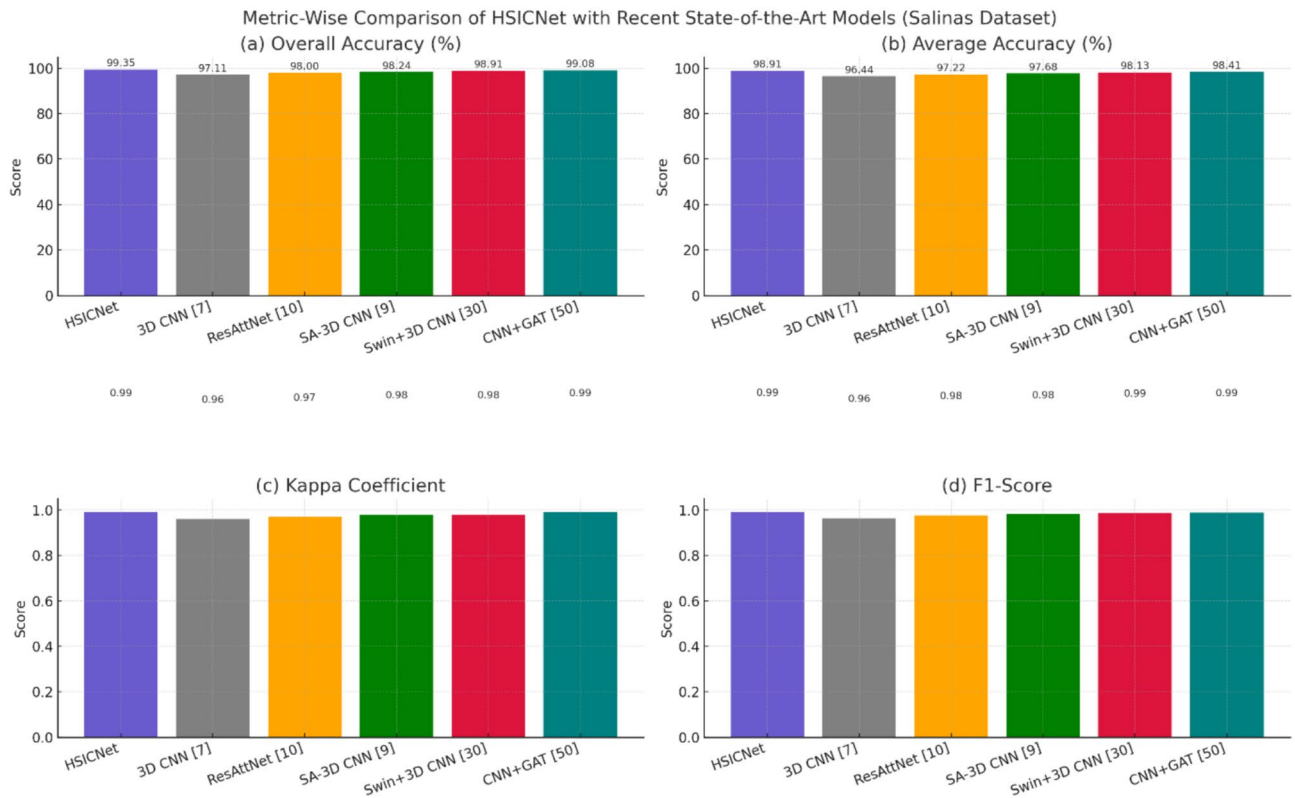


Fig. 22. Metric-wise comparison of HSICNet with five recent state-of-the-art deep learning models on the Salinas dataset.

hybrid residual-attention networks, and transformer-based models consistently perform well. Still, they are also limited by computational inefficiency, ineffective spectral and spatial feature fusion, and generalisation across different datasets.

This study systematically identified these research gaps by conducting a comprehensive literature review. Conventional CNN models struggle to exploit such sophisticated spectral-spatial correlations in hyperspectral data. Simultaneously, architectures that integrate attention augmentation or a transformer architecture incur high computational costs and require more labelled data. However, such deep learning architectures involve developing scalable, efficient, and widely applicable methods, which are mostly left as open issues in real HSIC problems.

HSICNet consists of two branches for spectral and spatial feature extraction, followed by a third-stage fusion using an attention-guided operation to alleviate these issues. Compared to traditional approaches, HSICNet focuses on efficient dimensionality reduction via PCA and on feature refinement via a dedicated attention module that improves class separability and reduces noise sensitivity.

Extensive experimental results on the Indian Pines, Pavia University, and Salinas datasets show that HSICNet achieves better overall accuracy, average accuracy, Kappa coefficient, and F1 score than the baseline and other recently reported state-of-the-art models. We also conduct an ablation study to corroborate that each part of the architecture above plays an important role. It showcases enhanced performance compared to previous studies, especially in dual-branch encoding and attention fusion. Our approach, HSICNet, is an essential step forward in hyperspectral image classification that addresses major drawbacks of previous methods (e.g., inefficient fusion schemes and limited scalability), with a significant impact on real-time remote sensing applications that require an interpretable, scalable model.

Strong empirical performance was also obtained in this study using PCA, a linear dimensionality reduction method (which was not the focus of subsequent sections of the paper, as most datasets had relatively few training samples). This is because PCA, by design, captures the maximum amount of spectral variance and removes noise, which can be beneficial when we have some training data points available for classification. On the other hand, under such circumstances, adaptive deep learning approaches may not generalise well and may suffer from overfitting or poor feature representation. This observation highlights the core tension between the flexibility afforded by data and the need for statistical robustness when data are scarce.

Qualitative analysis

In addition to quantitative results, we visually compared HSICNet-predicted classification maps with accurate classification maps. These visualisations assist in analysing whether our proposed model maps the input into neighbouring spatially coherent locations that are in proximity to the defined class boundaries, as well as in

examining how healthy classes separate from one another. HSICNet also achieves higher accuracy in locating spatial structures and lower salt-and-pepper noise, particularly in complex regions with mixed pixels, than the state-of-the-art methods. An attention-based fusion produces more contextually relevant predictions for minority or spectrally overlapping classes. There is apparent visual consistency across the spatial maps, which indicates the model's ability to extract spectro-spatial patterns.

Comparative evaluation

To evaluate HSICNet's performance relative to state-of-the-art models, we conducted a comparative analysis across benchmark datasets. To compare the results with the state of the art, four baseline models (3D-CNN, HybridSN, SSFTT, and DC-CNN) were evaluated on the same data splits and training conditions as used in our experiments. Such necessary measures as Overall Accuracy (OA), Average Accuracy (AA), Kappa Coefficient (κ), and per-class F-1 scores are included in this comparison. Healthcare Sliced Image Classification Network (HSICNet) achieved better performance than existing methods across all metrics (Table 6) and improved minority class detection & new class generalisation when dealing with class imbalance. We demonstrate that the discriminative power and robustness of our design can surpass those of existing methods through a hybrid convolutional architecture and attention-guided fusion.

Limitations of the Study

While HSICNet performs competitively by most metrics, it also has some drawbacks. First, although the presented model is light-weight, memory and computational issues may arise with extremely large-scale HSI cubes at ultra-high spectral resolution, especially in training. Second, while HSICNet demonstrates reasonable computational efficiency, it may still be impossible to deploy on highly constrained edge devices with limited memory or on devices that require real-time inference without additional pruning or quantisation. In the future, we will explore model compression and adaptive inference trajectories to address these deployment challenges and broaden their applicability to embedded and remote sensing systems.

Conclusion and future work

We have presented HSICNet, a novel deep learning architecture for hyperspectral image classification, achieving accurate and efficient classification results. HSICNet overcomes the key limitations of existing models by extracting spectral and spatial features separately via a dual-branch architecture, followed by attention-guided fusion and dimension reduction. Extensive experiments conducted on three popular benchmark datasets, Indian Pines, Pavia University, and Salinas, show that the proposed model achieves the highest overall accuracy, greater generalisation, and better class-balancing. The proposed framework addresses issues with existing techniques, such as overfitting and redundancy in noisy bands, and addresses the limited modelling of spectral-spatial interactions. Despite these advances, the study has limitations. Even the benchmark datasets used to evaluate those models do not accurately reflect the complexity of real-world scenarios, since they are static. Additionally, reliance on PCA and GPU-intensive training indicates that we still need significant architectural improvements. Future work will explore adaptive band-selection techniques that yield maximal class-discriminative information for minimal computation. Domain adaptation strategies can make the model more robust across diverse geographical and environmental conditions and physical environments. Additionally, using smaller transformer components and exploring self-supervised training paradigms may reduce the need for large annotated datasets. HSICNet generally serves as a framework for developing scalable, high-performance, and interpretable HSI solutions for diverse remote sensing applications. Looking ahead, we plan to use domain adaptation and online learning to meet HSICNet's out-of-sample generalisation expectations better when faced with real-world data. Not only does HSICNet improve classification performance on some publicly available hyperspectral datasets, but it also directly supports environmental monitoring, which primarily relies on remote sensing. Due to its powerful ability to classify spectrally homogeneous areas robustly while applying a spatial coherence constraint, it is a potential candidate for more advanced systems in the context of operational land use and cover, vegetation dynamics, and other ecological indicators monitoring systems. In the future, we will explore dynamic band-selection methods to achieve spectral relevance and adaptively reduce computational effort as needed. The integration of lightweight transformer modules with self-supervised learning techniques will reduce reliance on large, labelled datasets. In addition, by extending the real-time processing framework to embedded/edge devices, the applicability of our solution in field deployments will be further enhanced. Additionally, HSICNet will be evaluated on a broader, more challenging hyperspectral dataset to investigate its universality and scalability for complex remote sensing applications.

Data availability

Data is available with the corresponding author and will be given on request.

Materials availability

Materials used in this research are available with corresponding author and given on request.

Code availability

The code is available from the corresponding author and can be given on request.

Received: 14 April 2025; Accepted: 13 February 2026

Published online: 26 February 2026

References

- Guerri, M. F., Distante, C., Spagnolo, P., Bougourzi, F. & Taleb-Ahmed, A. Deep learning techniques for hyperspectral image analysis in agriculture: A review. *Elsevier* **12**, 1–19. <https://doi.org/10.1016/j.ophoto.2024.100062> (2024).
- Ullah, F. et al. Conventional to deep ensemble methods for hyperspectral image classification: A comprehensive survey. *IEEE* **17**, 3878–3916. <https://doi.org/10.1109/JSTARS.2024.3353551> (2024).
- Reddy, K. K. et al. Enhancing precision agriculture and land cover classification: a self-attention 3D convolutional neural network approach for hyperspectral image analysis. *IEEE* <https://doi.org/10.1109/ACCESS.2024.3420089> (2024).
- Farooque, G., Liu, Q., Sargano, A. B. & Xiao, L. Swin transformer with multiscale 3D atrous convolution for hyperspectral image classification. *Elsevier* **126**, 1–16. <https://doi.org/10.1016/j.engappai.2023.107070> (2023).
- Chhapariya, K., Buddhhiraj, K. M. & Kumar, A. A Deep spectral-spatial residual attention network for hyperspectral image classification. *IEEE* <https://doi.org/10.1109/JSTARS.2024.3355071> (2024).
- Ashraf, M. et al. Novel 3D deep neural network architecture for crop classification using remote sensing-based hyperspectral images. *IEEE* **17**, 12649–12665. <https://doi.org/10.1109/JSTARS.2024.3422078> (2024).
- Esmaeili, M., Abbasi-Moghadam, D., Sharifi, A., Tariq, A. & Li, Q. Hyperspectral image band selection based CNN embedded GA (CNNeGA). *IEEE* **16**, 1927–1950. <https://doi.org/10.1109/JSTARS.2023.3242310> (2023).
- Hameed, A. A. Enhancing hyperspectral remote sensing image classification using robust learning technique. *Elsevier* **36**(1), 1–8. <https://doi.org/10.1016/j.jksus.2023.102981> (2024).
- Sun, J., Yang, J., Chen, W. & Ding, S. LCTCS: low-cost and two-channel sparse network for hyperspectral image classification. *Springer* **56**(181), 1–31. <https://doi.org/10.1007/s11063-024-11631-y> (2024).
- Jia, J. et al. The effect of artificial intelligence evolving on hyperspectral imagery with different signal-to-noise ratio, spectral a. *Elsevier* **311**, 1–21. <https://doi.org/10.1016/j.rse.2024.114291> (2024).
- Bai, H., Tingfa, X., Chen, H., Liu, P. & Li, J. Content-driven magnitude-derivative spectrum complementary learning for hyperspectral image classification. *IEEE* **14**(8), 1–14. <https://doi.org/10.1109/TGRS.2024.3435079> (2024).
- Viel, F. et al. Hyperspectral image classification: An analysis employing CNN, LSTM, transformer, and attention mechanism. *IEEE Access* **11**, 24835–24850. <https://doi.org/10.1109/ACCESS.2023.3255164> (2023).
- Ashraf, A. M. N. A., Nawi, N. A. Z. R. I. M. O. H. D. & Aamir, M. U. H. A. M. M. A. D. Adaptive feature selection and hyper-spectral image classification using manifold learning techniques. *IEEE* **12**, 40279–40289. <https://doi.org/10.1109/ACCESS.2023.3322147> (2024).
- Ji, J. et al. Hyperspectral image classification based on unsupervised regularization. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **16**, 1871–1882. <https://doi.org/10.1109/JSTARS.2023.3241662> (2023).
- Zhu, Y., Yuan, K., Zhong, W. & Xu, L. Spatial-spectral ConvNeXt for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **16**, 5453–5463. <https://doi.org/10.1109/JSTARS.2023.3282975> (2023).
- Noshiri, N., Beck, M. A., Bidinosti, C. P. & Henry, C. J. A comprehensive review of 3D convolutional neural network-based classification techniques of diseased and defective crops using non-UAV-based hyperspectral images. *Elsevier* **5**, 1–14. <https://doi.org/10.1016/j.atech.2023.100316> (2023).
- Ullah, F. et al. Deep hyperspectral shots: Deep snap smooth wavelet convolutional neural network shots ensemble for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **17**, 14–34. <https://doi.org/10.1109/JSTARS.2023.3314900> (2024).
- Vasanthakumari, R. K., Nair, R. V. & Krishnappa, V. G. Improved learning by using a modified activation function of a Convolutional Neural Network in multi-spectral image clas. *Elsevier* **14**, 1–13. <https://doi.org/10.1016/j.mlwa.2023.100502> (2023).
- Okwuashi, O. & Ndehedehe, C. E. Deep support vector machine for hyperspectral image classification. *Pattern Recognit.* <https://doi.org/10.1016/j.patcog.2020.107298> (2020).
- Ortac, G. & Ozcan, G. Comparative study of hyperspectral image classification by multidimensional convolutional neural network approaches to improve accuracy. *Expert Syst. Appl.* **182**, 115280. <https://doi.org/10.1016/j.eswa.2021.115280> (2021).
- Zhu, M., Jiao, L., Liu, F., Yang, S. & Wang, J. Residual spectral-spatial attention network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* <https://doi.org/10.1109/tgrs.2020.2994057> (2020).
- Ghaderizadeh, S., Abbasi-Moghadam, D., Sharifi, A., Zhao, N. & Tariq, A. Hyperspectral image classification using a hybrid 3D-2D convolutional neural networks. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **14**, 7570–7588. <https://doi.org/10.1109/jstars.2021.3099118> (2021).
- Lin, F.-C. et al. A model for forest type identification and forest regeneration monitoring based on deep learning and hyperspectral imagery. *Elsevier* **80**, 1–16. <https://doi.org/10.1016/j.ecoinf.2024.102507> (2024).
- Lu, Y. et al. Rapid assessment of heavy metal accumulation capability of *Sedum alfredii* using hyperspectral imaging and deep learning. *Elsevier* **282**, 1–11. <https://doi.org/10.1016/j.ecoenv.2024.116704> (2024).
- Dash, S., Chakravarty, S., Giri, N. C., Agyekum, E. B. & AboRas, K. M. Minimum noise fraction and long short-term memory model for hyperspectral imaging. *Springer* **17**(16), 1–22. <https://doi.org/10.1007/s44196-023-00370-y> (2024).
- Olisah, C. C. et al. Convolutional neural network ensemble learning for hyperspectral imaging-based blackberry fruit ripeness detection in uncontrolled farm environment. *Elsevier* **132**, 1–12. <https://doi.org/10.1016/j.engappai.2024.107945> (2024).
- Wang, D., Du, B. & Zhang, L. Spectral-spatial global graph reasoning for hyperspectral image classification. *IEEE* **14**(8), 1–14. <https://doi.org/10.1109/TNNLS.2023.3265560> (2023).
- Yu, W., Wan, S., Li, G., Yang, J. & Gong, C. Hyperspectral image classification with contrastive graph convolutional network. *IEEE* **61**, 1–16. <https://doi.org/10.1109/TGRS.2023.3240721> (2022).
- Sun, Y., Wang, M., Wei, C., Zhong, Y. & Xiang, J. Heterogeneous spectral-spatial network with 3D attention and MLP for hyperspectral image classification using limited training samples. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **16**, 8702–8720. <https://doi.org/10.1109/JSTARS.2023.3271901> (2023).
- Zhang, Y., Li, W., Sun, W., Tao, R. & Du, Q. Single-source domain expansion network for cross-scene hyperspectral image classification. *IEEE* **32**, 1498–1512. <https://doi.org/10.1109/TIP.2023.3243853> (2022).
- Zheng, X., Sun, H., Lu, X. & Xie, W. Rotation-invariant attention network for hyperspectral image classification. *IEEE* **31**, 4251–4265. <https://doi.org/10.1109/TIP.2022.3177322> (2023).
- Xue, X., Zhang, H., Fang, B., Bai, Z. & Li, Y. Grafting transformer on automatically designed convolutional neural network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **60**, 1–16. <https://doi.org/10.1109/TGRS.2022.3180685> (2022).
- Huang, H., Luo, L. & Pu, C. Self-supervised convolutional neural network via spectral attention module for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* **19**, 1–5. <https://doi.org/10.1109/LGRS.2022.3141870> (2022).
- He, Z. et al. Hypervitgan: Semisupervised generative adversarial network with transformer for hyperspectral image classification. *IEEE* **15**, 6053–6068. <https://doi.org/10.1109/JSTARS.2022.3192127> (2022).
- Ahmad, M. et al. Hybrid dense network with attention mechanism for hyperspectral image classification. *IEEE* **15**, 3948–3957. <https://doi.org/10.1109/JSTARS.2022.3171586> (2022).
- Dong, Y., Liu, Q., Bo, D. & Zhang, L. Weighted feature fusion of convolutional neural network and graph attention network for hyperspectral image classification. *IEEE* **31**, 1559–1572. <https://doi.org/10.1109/TIP.2022.3144017> (2022).
- Feng, J. et al. Attention multibranch convolutional neural network for hyperspectral image classification based on adaptive region search. *IEEE Trans. Geosci. Remote Sens.* <https://doi.org/10.1109/tgrs.2020.3011943> (2020).
- Safari, K., Prasad, S. & Labate, D. A multiscale deep learning approach for high-resolution hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* <https://doi.org/10.1109/lgrs.2020.2966987> (2020).

39. Shen, Y. et al. Efficient deep learning of nonlocal features for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* <https://doi.org/10.1109/tgrs.2020.3014286> (2020).
40. Mou, L. et al. Deep reinforcement learning for band selection in hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* <https://doi.org/10.1109/tgrs.2021.3067096> (2021).
41. Ma, Y., Zhao, Y., Im, J., Zhao, Y. & Zhen, Z. A deep-learning-based tree species classification for natural secondary forests using unmanned aerial vehicle hyperspectral images and LiDAR. *Elsevier* **159**, 1–13. <https://doi.org/10.1016/j.ecolind.2024.111608> (2024).
42. Ahmed, M. T., Villordon, A. & Kamruzzaman, M. Comparative analysis of hyperspectral image reconstruction using deep learning for agricultural and biological applications. *Elsevier* **23**, 1–10. <https://doi.org/10.1016/j.rineng.2024.102623> (2024).
43. Ahmad, M. SCSNet: sharpened cosine similarity-based neural network for hyperspectral image classification. *IEEE* **21**, 1–5. <https://doi.org/10.1109/LGRS.2024.3365611> (2023).
44. Giakoumoglou, N. et al. Early detection of *Botrytis cinerea* symptoms using deep learning multi-spectral image segmentation. *Elsevier* **8**, 1–17. <https://doi.org/10.1016/j.atech.2024.100481> (2024).
45. Yao, D. et al. Deep hybrid: Multi-graph neural network collaboration for hyperspectral image classification. *Def. Technol.* **23**, 164–176. <https://doi.org/10.1016/j.dt.2022.02.007> (2023).
46. Amoako, P. Y. O. et al. A meta-reinforcement learning based hyperspectral image classification with small sample set. *IEEE* **17**, 3091–3107. <https://doi.org/10.1109/JSTARS.2023.3347879> (2024).
47. Farmonov, N. et al. Crop type classification by DESIS hyperspectral imagery and machine learning algorithms. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **16**, 1576–1588. <https://doi.org/10.1109/JSTARS.2023.3239756> (2023).
48. Sellami, A., Farah, M. & Mura, M. D. SHCNet: A semi-supervised hypergraph convolutional networks based on relevant feature selection for hyperspectral image classification. *Pattern Recognit. Lett.* **165**, 98–106. <https://doi.org/10.1016/j.patrec.2022.12.004> (2023).
49. Wang, D., Du, B., Zhang, L. & Tao, D. HKNAS: Classification of hyperspectral imagery based on hyper kernel neural architecture search. *IEEE* **14**(8), 1–15. <https://doi.org/10.1109/TNNLS.2023.3270369> (2023).
50. Pathak, D. K., Kalita, S. K. & Bhattacharya, D. K. Hyperspectral image classification using support vector machine: A spectral spatial feature based approach. *Springer* **15**, 1–13 (2022).
51. Huang, H., Pu, C., Li, Y. & Duan, Y. Adaptive residual convolutional neural network for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* <https://doi.org/10.1109/jstars.2020.2995445> (2020).
52. Jena, B. et al. Artificial intelligence-based hybrid deep learning models for image classification: The first narrative review. *Comput. Biol. Med.* **137**, 104803. <https://doi.org/10.1016/j.combiomed.2021.104803> (2021).
53. Wang, J., Song, X., Sun, L., Huang, W. & Wang, J. A novel cubic convolutional neural network for hyperspectral image classification. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **13**, 4133–4148. <https://doi.org/10.1109/jstars.2020.3008949> (2020).
54. Hong, D. et al. Interpretable hyperspectral artificial intelligence: When nonconvex modeling meets hyperspectral remote sensing. *IEEE Geosci. Remote Sens. Mag.* **9**(2), 52–87. <https://doi.org/10.1109/mgrs.2021.3064051> (2021).
55. Zhang, X. et al. SSDANet: Spectral-spatial three-dimensional convolutional neural network for hyperspectral image classification. *IEEE Access* **8**, 127167–127180. <https://doi.org/10.1109/access.2020.3008029> (2020).
56. Yuan, Q. et al. Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sens. Environ.* **241**, 111716. <https://doi.org/10.1016/j.rse.2020.111716> (2020).
57. Hong, S. M. et al. Monitoring the vertical distribution of HABs using hyperspectral imagery and deep learning models. *Sci. Total Environ.* **794**, 148592. <https://doi.org/10.1016/j.scitotenv.2021.148592> (2021).
58. Chen, S., Jin, M. & Ding, J. Hyperspectral remote sensing image classification based on dense residual three-dimensional convolutional neural network. *Multimedia Tools Appl.* <https://doi.org/10.1007/s11042-020-09480-7> (2020).
59. Wang, H., Wang, X., Chen, C. L. P. & Cheng, Y. Hyperspectral image classification based on domain adaptation broad learning. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **13**, 3006–3018. <https://doi.org/10.1109/jstars.2020.3001198> (2020).
60. Chen, S. et al. Early surveillance of rice Bakanae disease using deep learning and hyperspectral imaging. *Springer* <https://doi.org/10.1007/s42994-024-00169-1> (2024).
61. Ahmed, T. et al. Advancing sweetpotato quality assessment with hyperspectral imaging and explainable artificial intelligence. *Elsevier* **220**, 1–12. <https://doi.org/10.1016/j.compag.2024.108855> (2024).
62. Wei, Y., Yu, S., Giraldo, L. S. & Príncipe, J. C. Multiscale principle of relevant information for hyperspectral image classification. *Mach. Learn.* **112**, 1227–1252. <https://doi.org/10.1007/s10994-021-06011-9> (2023).
63. Xia, S., Zhang, X., Meng, H., Fan, J. & Jiao, L. Two-stream networks for contrastive learning in hyperspectral image classification. *IEEE* **17**, 1903–1920. <https://doi.org/10.1109/JSTARS.2023.3341338> (2023).
64. Sun, Y. et al. Perceiving spectral variation: Unsupervised spectrum motion feature learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **60**, 1–17. <https://doi.org/10.1109/TGRS.2022.3221534> (2022).
65. Xu, Y., Du, B. & Zhang, L. Robust self-ensembling network for hyperspectral image classification. *IEEE* **35**(3), 3780–3793. <https://doi.org/10.1109/TNNLS.2022.3198142> (2022).
66. Wieme, J. et al. Application of hyperspectral imaging systems and artificial intelligence for quality assessment of fruit, vegetables and mushrooms: A review. *Biosyst. Eng.* **222**, 156–176. <https://doi.org/10.1016/j.biosystemseng.2022.07.013> (2022).
67. Zhan, Y., Ke, W. & Dong, Y. Enhanced spectral-spatial residual attention network for hyperspectral image classification. *IEEE* **15**, 7171–7186. <https://doi.org/10.1109/JSTARS.2022.3197934> (2022).
68. Cao, X., Yao, J., Xu, Z. & Meng, D. Hyperspectral image classification with convolutional neural network and active learning. *IEEE Trans. Geosci. Remote Sens.* <https://doi.org/10.1109/tgrs.2020.2964627> (2020).
69. Sawant, S. S., Manoharan, P. & Loganathan, A. Band selection strategies for hyperspectral image classification based on machine learning and artificial intelligent techniques –Survey. *Arab. J. Geosci.* <https://doi.org/10.1007/s12517-021-06984-w> (2021).
70. Abdulridha, J., Ampatzidis, Y., Roberts, P. & Kakarla, S. C. Detecting powdery mildew disease in squash at different stages using UAV-based hyperspectral imaging and artificial intelligence. *Biosyst. Eng.* **197**, 135–148. <https://doi.org/10.1016/j.biosystemseng.2020.07.001> (2020).
71. Sagan, V. et al. Data-driven artificial intelligence for calibration of hyperspectral big data. *IEEE Trans. Geosci. Remote Sens.* <https://doi.org/10.1109/tgrs.2021.3091409> (2021).
72. Chen, H., Miao, F., Chen, Y., Xiong, Y. & Chen, T. A hyperspectral image classification method using multifeature vectors and optimized KELM. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **14**, 2781–2795. <https://doi.org/10.1109/jstars.2021.3059451> (2021).
73. Jia, S. et al. A lightweight convolutional neural network for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* <https://doi.org/10.1109/tgrs.2020.3014313> (2020).
74. Liu, B. et al. Deep multiview learning for hyperspectral image classification. *IEEE Trans. Geosci. Remote Sens.* **59**(9), 7758–7772. <https://doi.org/10.1109/tgrs.2020.3034133> (2021).
75. Kang, R., Park, B., Ouyang, Q. & Ren, N. Rapid identification of foodborne bacteria with hyperspectral microscopic imaging and artificial intelligence classification algorithms. *Food Control* **130**, 108379. <https://doi.org/10.1016/j.foodcont.2021.108379> (2021).
76. Meng, Z., Jiao, L., Liang, M. & Zhao, F. Hyperspectral image classification with mixed link networks. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **14**, 2494–2507. <https://doi.org/10.1109/jstars.2021.3053567> (2021).
77. Liu, S. & Shi, Q. Multitask deep learning with spectral knowledge for hyperspectral image classification. *IEEE Geosci. Remote Sens. Lett.* <https://doi.org/10.1109/lgrs.2019.2962768> (2020).

78. Jia, S. et al. A survey: Deep learning for hyperspectral image classification with few labeled samples. *Neurocomputing* **448**, 179–204. <https://doi.org/10.1016/j.neucom.2021.03.035> (2021).
79. Wang, C. et al. A review of deep learning used in the hyperspectral image analysis for agriculture. *Artif. Intell. Rev.* **54**(7), 5205–5253. <https://doi.org/10.1007/s10462-021-10018-y> (2021).
80. Ranjan, P., Kumar, R. & Girdhar, A. A 3D-convolutional-autoencoder embedded Siamese-attention-network for classification of hyperspectral images. *Neural Comput. Appl.* **36**, 8335–8354. <https://doi.org/10.1007/s00521-024-09527-y> (2024).
81. Ranjan, P. et al. A novel spectral-spatial 3D auxiliary conditional GAN integrated convolutional LSTM for hyperspectral image classification. *Earth Sci. Inform.* **17**, 5251–5271. <https://doi.org/10.1007/s12145-024-01451-y> (2024).
82. Ranjan, P. & Girdhar, A. Xcep-dense: A novel lightweight extreme inception model for hyperspectral image classification. *Int. J. Remote Sens.* **43**(14), 5204–5230. <https://doi.org/10.1080/01431161.2022.2130727> (2022).
83. Ranjan, P. & Gupta, G. A cross-domain semi-supervised zero-shot learning model for the classification of hyperspectral images. *J. Indian Soc. Remote Sens.* **51**, 1991–2005. <https://doi.org/10.1007/s12524-023-01734-9> (2023).
84. P. Ranjan, R. Kumar and A. Girdhar. Unlocking the potential of unlabeled data: semi-supervised learning for stratification of hyperspectral images. *2023 OITS International Conference on Information Technology (OCIT), Raipur, India*. 938–943, <https://doi.org/10.1109/OCIT59427.2023.10430513>. (2023).
85. Ranjan, P. et al. Revolutionizing hyperspectral image classification for limited labeled data: Unifying autoencoder-enhanced GANs with convolutional neural networks and zero-shot learning. *Earth Sci. Inform.* **18**, 216. <https://doi.org/10.1007/s12145-025-0173-9-7> (2025).
86. Pallavi Ranjan, Rajeev Kumar, Ki-Hyun Jung. *The Korean Institute of Electrical Engineers Conference, Jeju*. (2024).
87. P. Ranjan and A. Girdhar, A comparison of deep learning algorithms dealing with limited samples in hyperspectral image classification. *2022 OPJU International Technology Conference on Emerging Technologies for Sustainable Development (OTCON), Raigarh, Chhattisgarh, India*. 1–6. <https://doi.org/10.1109/OTCON56053.2023.10114005>. (2023).
88. P. Ranjan, R. Kumar and A. Girdhar. Recent CNN advancements for stratification of hyperspectral images. *2023 6th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India*. 1–5. <https://doi.org/10.1109/ISCON57294.2023.10112174>. (2023).
89. Paheding, S. et al. *Advancing Horizons in Remote Sensing: A Comprehensive Survey of Deep Learning Models and Applications in Image Classification and Beyond* 1–41 (Springer, 2024).
90. Purdue University, (n.d.). *Hyperspectral Remote Sensing Scenes*. https://www.ehu.us/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.
91. Pavia University, (n.d.). *Hyperspectral Remote Sensing Scenes*. https://www.ehu.us/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.
92. Salinas Valley, (n.d.). *Hyperspectral Remote Sensing Scenes*. https://www.ehu.us/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes.

Author contributions

All authors contributed to the study's conception and design. Material preparation, data collection, and analysis were performed by **Dr. K. Purnachand, Dr. Rafath Samrin, ** **John Babu Guttikonda, Dr. Rajitha Kotoju, Phani Sridhar Addepalli, A. Mohan** **.* The first draft of the manuscript was written by **Dr. K. Purnachand** all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Declarations

Competing interests

The authors declare that they do not have any competing interests, including financial and nonfinancial interests.

Consent for publication

The authors give consent for their publication.

Additional information

Correspondence and requests for materials should be addressed to K.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2026