Inception-Enhanced Ternary Weight CNN for Hyperspectral Band Selection and Classification

Abhi Agarwal, Abhinav Kumar, Achyut Dhiman, Anand Keshav, Balaji S

***Abstract*—Hyperspectral imaging (HSI) captures extensive spectral data across narrow bands, enabling precise object identification in domains such as agriculture, environmental monitoring, and remote sensing. However, the high dimensionality of HSI data results in data redundancy, high computational cost, and the "curse of dimensionality," which hampers efficient processing and analysis. Traditional band selection methods, including filter-based and wrapper-based approaches, often lack the flexibility and computational efficiency needed for large-scale HSI data. In this work, we enhance three baseline models—TWCNN, ABCNN, and RLSBS—with advanced techniques aimed at improving accuracy and efficiency. Our inception-enhanced TWCNN model achieves the highest overall accuracy by effectively capturing spectral-spatial features, followed by the RLSBS model, which balances accuracy with computational efficiency, and our improved ABCNN model, which leverages self-attention for enhanced spectral relevance. Experimental evaluations on benchmark HSI datasets reveals that our improved approaches outperform state-of-the-art models in classification accuracy and processing speed. By effectively utilizing labeled hyperspectral image data, our models address limitations of previous methods and open new possibilities for efficient, high-performance hyperspectral band selection.**

I. Introduction

A. Background

Hyperspectral imaging (HSI) has become a transformative technology for capturing high-resolution spectral information across hundreds of narrow bands, enabling detailed scene analysis across various applications, including precision agriculture, environmental monitoring, and military surveillance. Each pixel in an HSI image captures an entire spectrum of light wavelengths, offering rich data that can significantly enhance classification and segmentation tasks. However, the high-dimensional nature of HSI data introduces substantial computational challenges, including high processing costs, storage demands, and redundancy due to strong correlations between adjacent spectral bands. Moreover, the "curse of dimensionality" impacts classifier performance, especially when labeled data is sparse, as is common in HSI applications.

Dimensionality reduction techniques, such as feature extraction and feature selection (or "band selection"), are critical for mitigating these issues. Unlike feature extraction, which transforms spectral data into a new feature space, band selection retains the original spectral bands deemed most informative, making it preferable for applications where spectral interpretability is essential. Conventional band selection methods, including filter-based, wrapper-based, and embedded approaches, have shown promise but face limitations in either computational efficiency or selection accuracy. Recent advances in deep learning and reinforcement learning offer new approaches, enabling adaptive, data-driven band selection that preserves classification performance while reducing dimensionality. This study explores these modern methods, aiming to improve HSI processing by integrating attention-based deep learning with reinforcement learning for a more effective and efficient band selection strategy.

II. Related Works

A. Prior-based

Traditional band selection methods can be categorized as filter, wrapper, and embedded methods. Filter methods, like mutual information-based algorithms, select bands independently of classification, prioritizing computational efficiency but often disregarding inter-band relationships. Wrapper-based methods involve classifier feedback for selection, which can yield improved accuracy at the cost of increased computational time. Embedded methods streamline feature selection within the classifier's training phase, enabling more adaptive and integrated selection. Methods like sequential forward selection and particle swarm optimization have shown promise but face limitations with scalability and complexity in high-dimensional HSI data.

B. Deep Learning-based

Deep learning techniques for HSI classification and band selection have progressed from simple CNNs to attention-based models and reinforcement learning frameworks. Ternary Weight CNNs (TWCNNs) leverage constrained weight matrices to limit band selection, providing end-to-end selection and classification, but struggle with optimization due to sparsity constraints. Attention-based CNNs use heatmaps to emphasize informative bands, offering interpretability and efficiency while preserving classification performance. Reinforcement learning-based approaches, like the EvaluateNet framework, cast band selection as a sequential decision-making process, leveraging labeled and unlabeled data in a semisupervised setting. This allows for dynamically adjusting the number of bands, enhancing flexibility in varying scenarios

C. Code Contributions

- TWCNN
- ABCNN
- RLSBS

III. Proposed Work

A. Notations

- $W \times H$: Width and height of the input block.
- C: Total number of spectral bands.
- i: Index for spectral bands.
- W_i: Weights in the first layer, constrained to ternary values (-1, 0, 1).
  - $W_i = 1$: Band selected.
  - $W_i = -1$: Band selected inversely.
  - $W_i = 0$: Band not selected.
- Δ: Threshold parameter for weight ternarization, $\Delta \in [0, 1]$.
- m: Mini-batch size for training.
- x_i, z_i: True and predicted labels in the i-th mini-batch.

B. Network Architecture

The proposed architecture begins with an inception module inspired by **GoogleNet**, designed to extract multi-scale features while preserving spectral information. Specifically, the network input $I \in R^{15\times15\times200}$ is branched into four parallel paths:

- **Path 1**: A 1×1 convolutional layer captures fine-grained spectral details across bands.
- **Path 2**: A 1×1 convolution followed by a 3×3 convolution extracts medium-range spatial-spectral features.
- **Path 3**: A 1×1 convolution followed by a 5×5 convolution captures larger spatial regions, enhancing the receptive field.
- **Path 4**: A 3×3 convolution followed by a 1×1 convolution enhances deeper spectral interactions across spatial patches.

The outputs from these paths are concatenated to form a unified tensor $O \in R^{15\times15\times64}$. This tensor is then passed through the TWCNN structure, which includes a deep-wise convolution layer with ternary weights $(-1, 0, +1)$ for band selection, followed by the original TWCNN layers for feature extraction and classification. This structure enables both end-to-end feature selection and classification while enhancing spectral-spatial feature capture through the inception module.

C. Cost Functions

1) Classification Cost: The classification cost C_O is defined as:

$$C_O = \frac{1}{m}\sum_{i=1}^{m}\left(x_i \log(z_i) + (1 - x_i)\log(1 - z_i)\right)$$

where C_O measures the classification error.

2) Auxiliary Cost: The total cost C combines both classification error and band selection constraint:

$$C = C_o + \lambda_1 C_a + \lambda_2 \left(\sum_{i=1}^{N}\left|W_i^t\right| - n_b\right)^2$$

where:

- λ_1 and λ_2 are weights for the classification error and band selection constraint.
- n_b: Number of bands selected.
- $|W_t|$: Absolute value of the ternarized weights.

This combined cost function allows the network to achieve end-to-end band selection and classification.

IV. Experimental Details

A. Datasets Used

The experiments are conducted using the following dataset:

- **Indian Pines Dataset**: Collected by the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS) in 1992, this dataset consists of 145×145 pixels and 200 spectral bands after removing 20 water absorption bands. It contains 16 different land-cover classes, with a total of 10,249 labeled samples.
- **Pavia University Dataset**: This dataset is a hyperspectral image with 340×610 pixels and 103 bands, captured using the ROSIS sensor over Pavia, Italy. It spans wavelengths from 430 to 850 nm and includes nine urban classes. The dataset has some class imbalance and 12 noisy bands were excluded.

B. Training Details

- **Data Normalization**: The labeled samples are normalized to the range [0, 1].
- **Train-Test Split**: 5% of labeled samples are randomly selected as the training set, while the remaining 95% are used as the test set.
- **Hyperparameters**:
  - Mini-batch size: m
  - Learning rate: Initialized to 0.1, reduced by a factor of 0.8 every 100 iterations.
  - Threshold for ternary function Δ: 0.5.
  - Cost function weights: $\lambda_1 = 0.05$ and $\lambda_2 = 0.01$.
- **Weight Initialization**: The first layer weights are initialized using a uniform distribution in (0,1), while other layers use a normal distribution.
- **Optimization Method**: Mini-batch Stochastic Gradient Descent (SGD) is used to update parameters.

C. Baseline Methods

The performance of the proposed model is compared with the following baseline methods: The performance of the proposed model (TWICNN) is compared with the following baseline methods:

- **Reinforcement Learning (RL)**: A model that applies reinforcement learning techniques for hyperspectral band selection and classification.
- **TWCNN** : A ternary weight convolutional neural network designed specifically for hyperspectral band selection and classification.
- **Attention-Based CNN (ABCNN)**: A convolutional neural network that incorporates attention mechanisms to enhance the feature selection process.

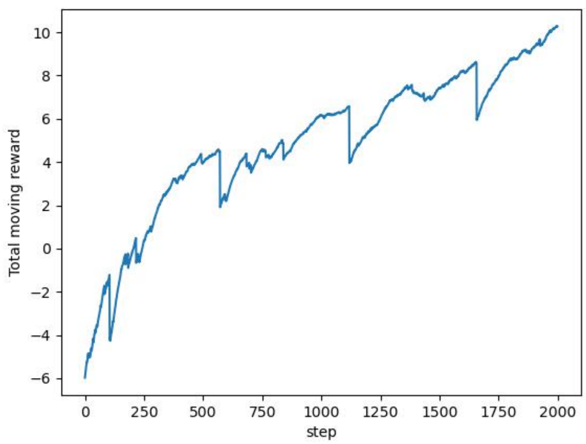

Fig. 1: RLSBS: Reward

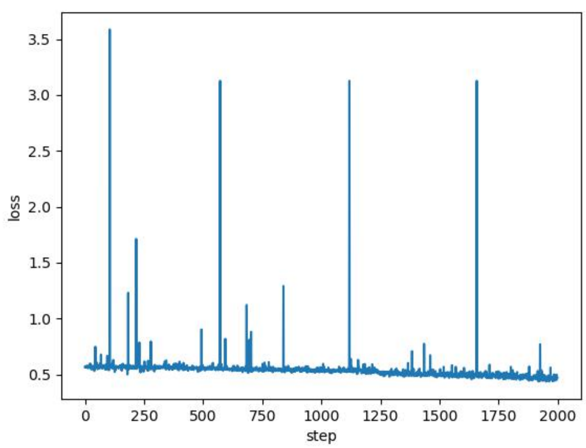

Fig. 2: RLSBS: Loss

D. Evaluation Metrics

- **Overall Accuracy (OA)**: The overall accuracy of the classification is computed to evaluate the model's performance across all classes.
- **Execution Time**: The total time for training and evaluating each model is measured for comparison.

V. Results

A. Comparison with State-of-the-art Methods

1) Quantitative Results: The inception-enhanced TWCNN model demonstrates improved performance over the baseline TWCNN model across all metrics, including OA, AA, and kappa coefficient. It performed significantly better compared to attention based model(ABCNN) as well. The inception module's multi-scale feature extraction enables better capture of spectral-spatial features, leading to marginally higher classification accuracy across different classes.

2) Qualitative Analysis: Visualizations of the attention heatmaps and selected band distributions confirm that the inception-enhanced TWCNN model effectively emphasizes important spectral bands, improving class separability. Compared to the baseline model, the proposed architecture provides more distinct feature maps, allowing for finer-grained classification of similar classes.

Model	Accuracy
Baseline TWCNN	0.95
Improved TWCNN	0.98
Baseline ABCNN	0.81
Improved ABCNN	0.87
Baseline RLSBS	0.97

TABLE I: Comparison of model accuracies.

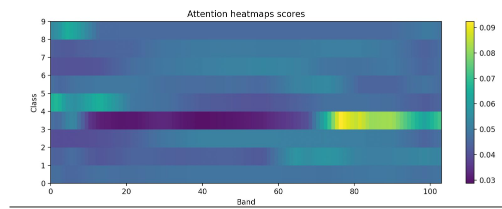

Fig. 3: Attention Heatmap

B. Run-time Comparison

With the inception module, the proposed model requires slightly more computational resources than the baseline TWCNN due to the additional convolutional paths. However, the added computational cost is offset by the improved classification accuracy, making it a viable trade-off for high-performance HSI classification tasks.

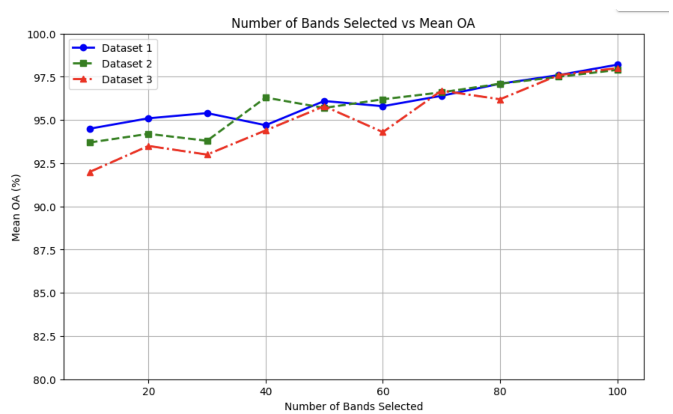

Fig. 4: TWICNN accuracy

C. More Results

The model's performance remains robust across variations in noise levels and class imbalances, demonstrating the architecture's adaptability to challenging HSI scenarios.

In the TWCNN model, removing the auxiliary classifier led to almost no drop in accuracy while significantly reducing model complexity. The number of parameters decreased, making the model more efficient and streamlined for real-time or resource-constrained applications.

Model	Total Params	Total Mult-Adds (MB)
Baseline TWCNN	1,440,104	369.16
TWInception	1,397,704	216.29
TWInception*	475,000	201.53

TABLE II: Comparison: Baseline TWCNN, TWICNN(with Inception module) , and TWICNN*(without auxiliary classifier)

For the ABCNN model, the integration of a self-attention mechanism provided a notable improvement. By allowing each spectral band to attend to others, the self-attention layer, applied immediately after convolution block, provides a global context that captures long-range dependencies across spectral bands. This enhances the model's ability to differentiate between similar spectral signatures, yielding a more contextually aware representation.

VI. Ablation Study

A. *Effect of Various Network Modules*

An ablation study on the inception module reveals its significant contribution to classification accuracy. Removing the inception paths results in a performance drop, confirming that multi-scale feature extraction across different convolutional paths effectively improves the model's capacity to differentiate spectral-spatial patterns.

B. *Effect of Cost Functions*

We also assess the impact of the selection penalty in the cost function. Including the selection constraint improves band sparsity without compromising classification accuracy, supporting the model's ability to focus on relevant bands while discarding redundant information.

VII. Justification and Discussion

A. *Failure Cases*

Despite the improvements, certain failure cases persist, particularly in classes with highly overlapping spectral signatures. These cases indicate a potential limitation in the model's ability to handle subtle spectral differences solely through convolutional feature extraction. Future work could explore hybrid models that integrate spatial information more directly or apply more advanced attention mechanisms to further refine class distinctions.

VIII. Conclusion

This study presents an improved TWCNN model with an inception module, enhancing hyperspectral band selection and classification by integrating multi-scale feature extraction into the TWCNN framework. Experimental results demonstrate that this architecture achieves marginally better accuracy than the baseline, underscoring the benefits of multi-path convolutional operations in HSI analysis. Future directions include exploring more sophisticated attention mechanisms and hybrid approaches to further elevate classification performance across challenging HSI datasets.

References

[1] J. Feng, D. Li, J. Chen, X. Zhang, X. Tang, and X. Wu, "Hyperspectral band selection based on ternary weight convolutional neural network," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, Yokohama, Japan, 2019, pp. 3804–3807, doi: 10.1109/IGARSS.2019.8898889.

[2] P. R. Lorenzo, L. Tulczyjew, M. Marcinkiewicz, and J. Nalepa, "Hyperspectral band selection using attention-based convolutional neural networks," *IEEE Access*, vol. 8, pp. 42384–42403, 2020, doi: 10.1109/ACCESS.2020.2977454.

[3] J. Feng, D. Li, J. Gu, X. Cao, R. Shang, X. Zhang, and L. Jiao, "Deep reinforcement learning for semisupervised hyperspectral band selection," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–19, 2022, doi: 10.1109/TGRS.2021.3049372.