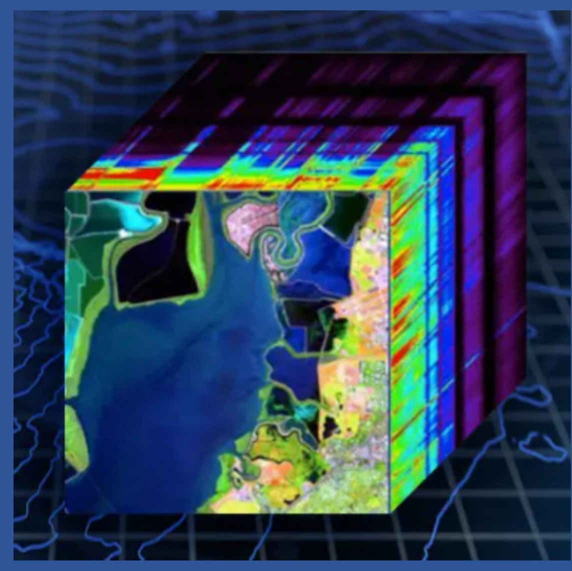
HYPERSPECTRAL IMAGE CLASSIFICATION USING 2D AND 3D CNNS

Pushkar Telavane - 24223351 | Isha Borgaonkar - 24209758

ACM40960 - Projects in Maths Modelling, MSc Data and Computational Science, University College Dublin



Abstract

Hyperspectral Imaging (HSI) provides rich spectral–spatial information across hundreds of contiguous bands, enabling high-precision land-cover classification. However, its high dimensionality, spectral redundancy, and limited labeled samples make accurate classification challenging.

This project compares 2D and 3D Convolutional Neural Networks (CNNs) on the Indian Pines dataset, using Principal Component Analysis (PCA), patch extraction, oversampling, and data augmentation in preprocessing. Performance, measured by Overall Accuracy (OA), Average Accuracy (AA), and the Kappa coefficient, shows 3D CNNs achieve superior OA and Kappa by exploiting joint spectral–spatial features, while 2D CNNs deliver competitive accuracy with lower computational cost—highlighting trade-offs between complexity and efficiency for real-world HSI applications.

Motivation

- Hyperspectral imaging enables detailed spectral analysis beyond conventional RGB images.
- Indian Pines dataset is a benchmark for land-cover classification research.
- Need for accurate 2D & 3D CNN models to handle spectral–spatial complexity.
- High classification accuracy supports real-world applications in agriculture and remote sensing.
- Improves crop monitoring, resource management, and environmental analysis.

Methodology

The Indian Pines hyperspectral dataset (145×145 pixels, 200 spectral bands) with ground truth labels was preprocessed through spectral standardization and Principal Component Analysis (PCA) to reduce redundancy while retaining over 99% variance. Fixed-size spatial–spectral patches were extracted, followed by oversampling of minority classes and data augmentation (flips, rotations) to address class imbalance. Two model architectures were implemented: a 2D CNN extracting spatial features per band independently, and a 3D CNN capturing joint spectral–spatial correlations. Models were trained using categorical cross-entropy loss with the SGD optimizer, and evaluated using Overall Accuracy (OA), Average Accuracy (AA), Kappa coefficient, and inference time. Classification maps were generated to visually compare predicted outputs with the ground truth.

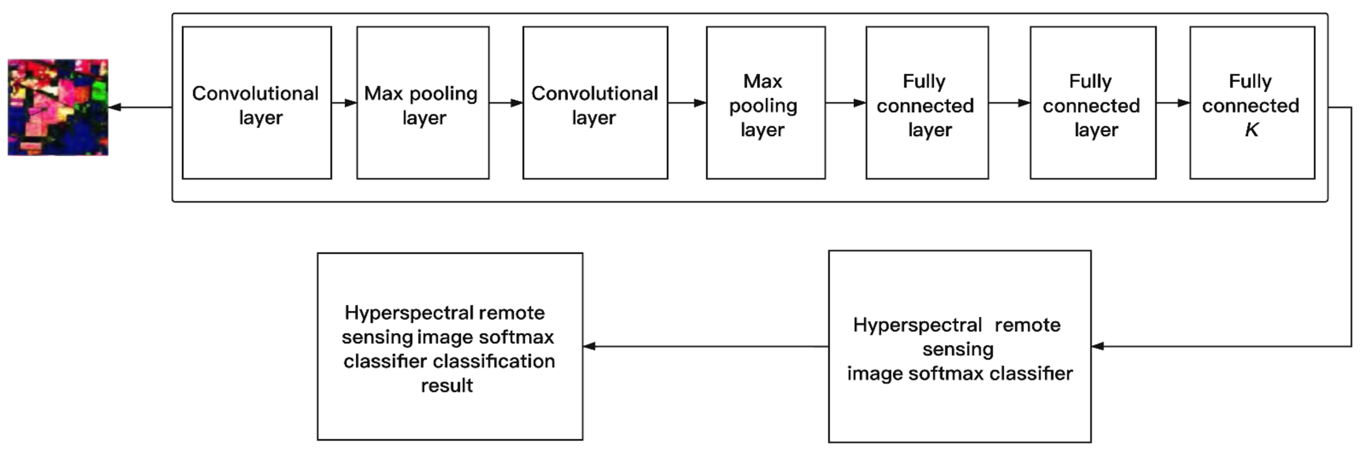


Figure 1 - Flowchart of Hyperspectral Image Classification

Step 1: Convolution Layer for feature extraction

For 2D-CNN: $Y(i,j,k) = \sigma \left(\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} \sum_{c=0}^{C-1} X(i+m, j+n, c) \times W(m,n,c,k) + b_k \right)$

For 3D-CNN: $Y(x,y,z,k) = \sigma \left(\sum_{i=0}^{D-1} \sum_{j=0}^{H-1} \sum_{l=0}^{W-1} X(x+i, y+j, z+l) \times W(i,j,l,k) + b_k \right)$

Step 2: Pooling Layer Reduces spatial size while retaining key features.

$Y(i,j,k) = \max_{(m,n) \in \Omega} X(i+m, j+n, k)$

Step 3: Fully Connected Layer Flattens features for classification.

$y = \sigma (W \cdot x + b)$

Step 4: Softmax Classifier Outputs class probabilities.

$P(y = k \mid x) = \exp(z_k) / \sum_{j=1}^{K} \exp(z_j)$

Step 5: Cross-Entropy Loss Optimization objective for training.

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

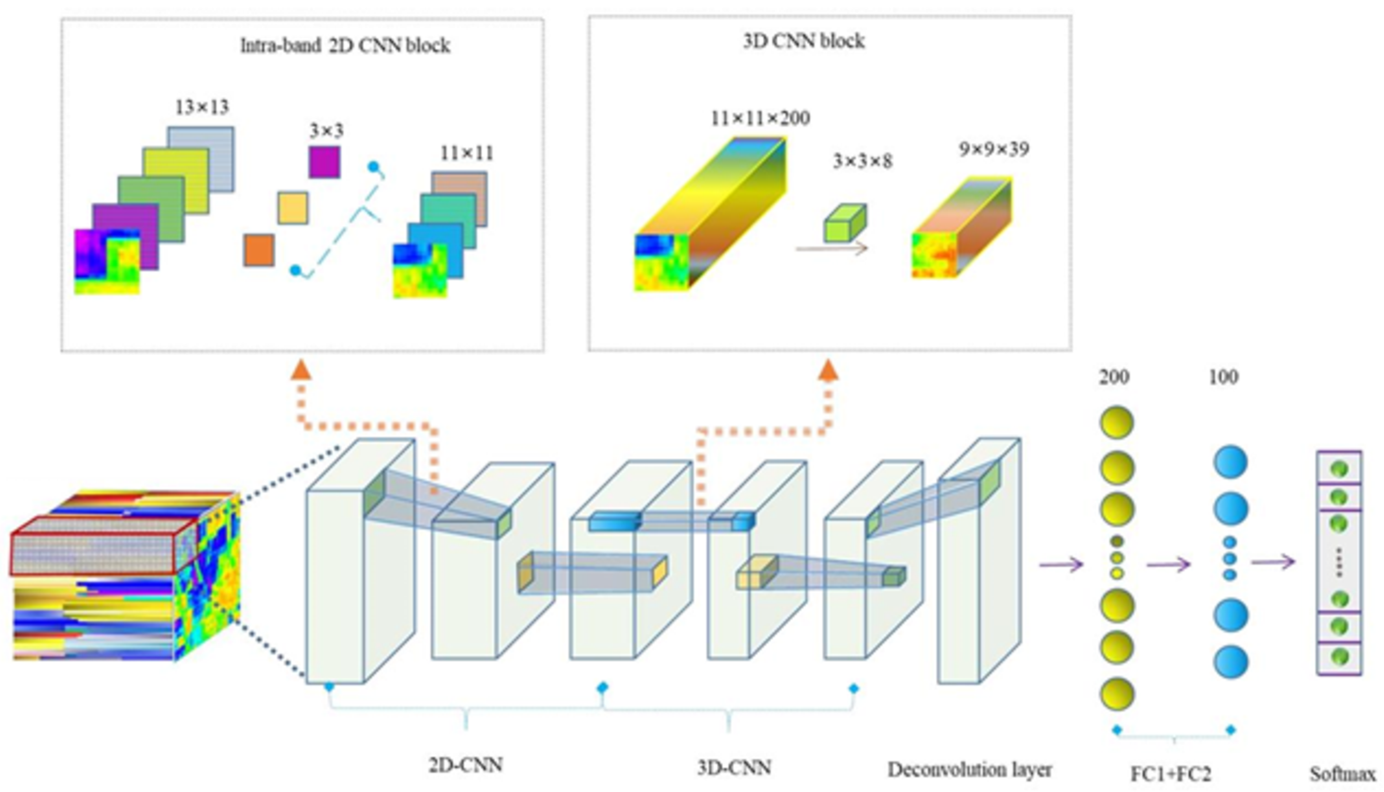


Figure 2 - 2D & 3D CNN HSI Classification Architecture

Results

Model	OA (%)	AA (%)	Kappa (%)	Test Loss	Test Acc (%)	Inference Time (s)
2D CNN	95.85	98.57	95.3	0.1218	95.85	0.28
3D CNN	95.98	91.07	95.42	0.1229	95.98	1.15

Figure 3 - Neural Network Metrices Table

The 3D CNN marginally surpasses the 2D CNN in Overall Accuracy (95.98% vs. 95.85%) and Kappa, reflecting superior spectral–spatial feature extraction.

Despite a longer inference time, its enhanced modeling of hyperspectral data renders it more suitable for accuracy-critical applications in real-world scenarios.

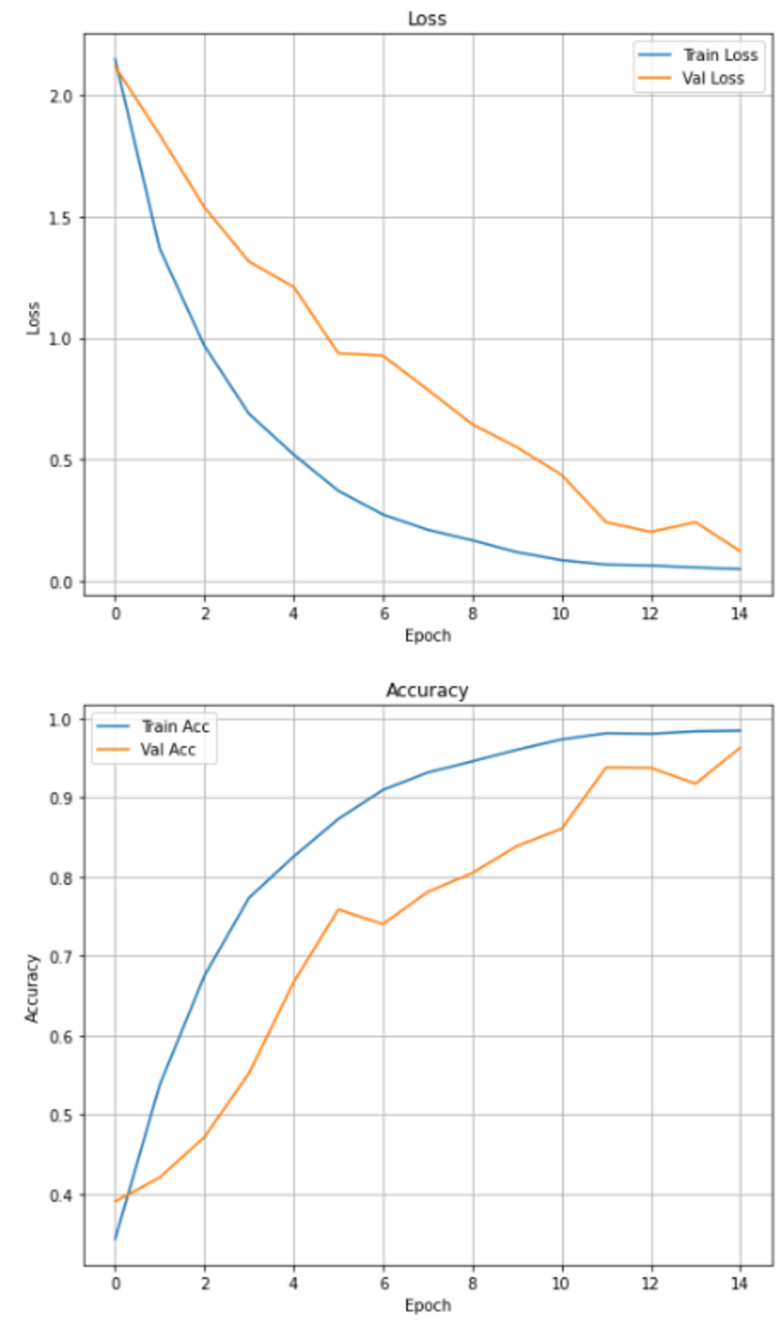


Figure 4 & 5 - Loss and Accuracy Curve for 3D CNN

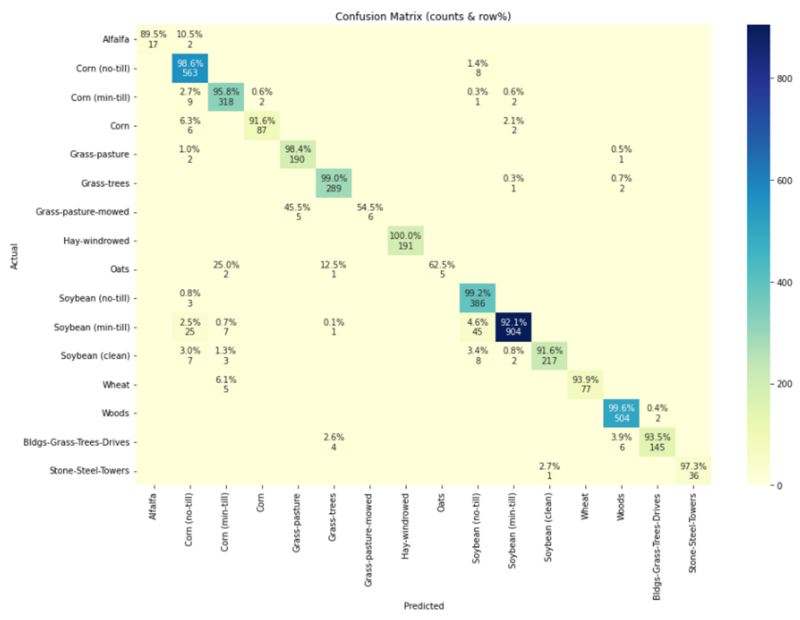


Figure 6 - Confusion Matrix for 3D CNN

The loss–accuracy curves for the 3D CNN show rapid convergence within the initial epochs, with validation accuracy closely matching training accuracy, indicating minimal overfitting. Final validation accuracy exceeds 95%, and loss approaches near-zero, evidencing strong generalization.

The confusion matrix highlights high per-class accuracy, with perfect classification in classes such as Hay-windrowed and Stone–Steel–Towers. Minor misclassifications occur mainly among spectrally similar crops (e.g., Soybean variants), a common challenge in hyperspectral classification. Overall, the model exhibits strong robustness in differentiating both distinct and spectrally close classes.

Figure 7 - Ground Truth Image

Figure 8 - Predicted 2D CNN Image

Figure 9 - Predicted 3D CNN Image

The ground truth map (left) illustrates the true spatial distribution of land-cover classes. The 2D CNN prediction (middle) effectively captures large homogeneous regions but exhibits scattered misclassifications in smaller or spectrally similar areas. In contrast, the 3D CNN prediction (right) produces cleaner maps with sharper class boundaries and reduced noise, closely aligning with the ground truth. By jointly modeling spectral and spatial information, the 3D CNN offers superior spatial fidelity, making it well-suited for accuracy-critical hyperspectral classification tasks.

Conclusion

This project demonstrates that both 2D and 3D CNN architectures can achieve high classification accuracy for hyperspectral imagery, with the 3D CNN achieving superior Overall Accuracy, Kappa coefficient, and class-wise recall. The 3D CNN's ability to jointly exploit spectral and spatial correlations enables more precise boundary delineation and reduced misclassification in spectrally similar regions. While the 2D CNN offers computational efficiency and faster inference, the 3D CNN's enhanced representational capacity makes it better suited for real-world applications demanding high spatial–spectral fidelity and robust generalization.

Future Scope

- Extend the framework to larger, more diverse hyperspectral datasets to improve generalization.
- Investigate hybrid architectures integrating 2D and 3D CNNs for a balance between accuracy and computational cost.
- Incorporate advanced feature selection and attention mechanisms to enhance spectral–spatial information fusion.
- Explore domain adaptation techniques to transfer models across varying sensors, seasons, and geographical locations.
- Optimize inference speed for deployment in time-critical applications such as precision agriculture and disaster monitoring.

References :

1 .Bo Du, Yuxiang Zhang, Liangpei Zhang, and Dacheng Tao, “Sparsity-Regularized Robust Non-Negative Matrix Factorization for Hyperspectral Unmixing,” IEEE Transactions on Cybernetics, 2016.

2 .Wei He, Hongyan Zhang, and Liangpei Zhang, “Robust Non-Negative Matrix Factorization for Hyperspectral Unmixing,” IEEE Transactions on Image Processing, Vol. 25, No. 11, November 2016.

3. Meiping Song, Chunyan Yu, Hongye Xie, and Chein-I Chang, “Progressive Band Selection Processing of Hyperspectral Image Classification (PBSP-HSIC),” IEEE Transactions on Image Processing, 2019.

4 .Ribana Roscher and Björn Waske, “Shapelet-Based Sparse Representation for Landcover Classification of Hyperspectral Images,” IEEE Transactions on Geoscience and Remote Sensing, Vol. 54, Issue 3, March 2016.

5.Jianjun Liu, Zebin Wu, Zhihui Wei, Liang Xiao, and Le Sun, “Spatial-Spectral Kernel Sparse Representation for Hyperspectral Image Classification,” IEEE Transactions on Geoscience and Remote Sensing, 2018.

6. Minchao Ye, Chenxi Ji, Hong Chen, Ling Lei, Huijuan Lu, and Yuntao Qian, “Residual Deep PCA-Based Feature Extraction for Hyperspectral Image Classification,” Springer, 2019.