Exploring Applications of Convolutional Neural Networks in Analyzing Multispectral Satellite Imagery: A Systematic Review

Charishma L[1], Mr K Janardhan[2], Dr D Venkatesh[3]

[123]computer Science and Engineering

---***---

Abstract - Multispectral image classification is a fundamental task in remote sensing, supporting applications such as land-cover mapping, agricultural monitoring, and environmental surveillance. Conventional classification methods, including Maximum Likelihood Classifier (MLC), Support Vector Machine (SVM), Decision Tree (DT), and Multi-Layer Perceptron (MLP), often face limitations when dealing with the high dimensionality and complex spectral characteristics of multispectral data. Recent advancements in deep learning have significantly improved the capability of remote sensing systems by enabling automatic extraction of high-level and discriminative features from raw data.

In this study, we investigate the use of Deep Neural Networks (DNNs) for pixel-wise classification of multispectral satellite imagery. DNNs can effectively learn hierarchical feature representations, making them suitable for complex image analysis tasks. We propose a lightweight DNN architecture composed of six layers: an input layer representing spectral reflectance values across multiple bands, a fully connected layer, a batch normalization layer, a Rectified Linear Unit (ReLU) activation layer, a second fully connected layer, and a SoftMax output layer for multi-class classification. In the proposed approach, each pixel is represented as a vector of spectral reflectance values corresponding to different spectral bands.

***Key Words*:** Remote Sensing, Multispectral Image Classification, Deep Neural Networks, Landsat Imagery, Pixel-wise Classification.

1.INTRODUCTION

Multispectral imaging has become an essential technology in the field of remote sensing, enabling the acquisition of detailed information about the Earth's surface by capturing data across multiple spectral bands. Unlike conventional RGB imagery, which records only three visible bands (red, green, and blue), multispectral images cover a wider portion of the electromagnetic spectrum, including visible, near-infrared (NIR), and shortwave infrared (SWIR) regions. This extended spectral coverage provides richer information about surface materials and significantly improves the ability to distinguish between different land cover types. As a result, multispectral imagery is widely used in applications such as agricultural monitoring, land-use and land-cover classification, forestry management, environmental change detection, and disaster assessment.

In remote sensing analysis, pixel-based classification is one of the most commonly used approaches for interpreting multispectral imagery. In this method, each pixel is treated as an individual data point and represented as a vector of spectral reflectance values across different bands. However, multispectral image classification presents several challenges. The high dimensionality of spectral data, along with high inter-class similarity and intra-class variability, makes accurate classification difficult for traditional machine learning algorithms.

Several conventional machine learning techniques have been widely used for pixel-based classification of multispectral images, including Maximum Likelihood Classifier (MLC), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbors (k-NN), and Artificial Neural Networks (ANN). In this approach, a set of training samples is selected from homogeneous regions representing different land-cover categories. These samples are used to train the classification model, which is then applied to classify all pixels in the image scene. Early neural network models, particularly Multi-Layer Perceptron (MLP) architectures trained using backpropagation algorithms, were introduced in the late 1980s for pixel-based classification of multispectral imagery. Compared with some traditional statistical methods such as MLC, neural networks demonstrated improved efficiency and adaptability.

In recent years, deep learning has significantly advanced the field of image analysis by introducing models capable of learning complex and non-linear feature representations directly from raw input data. Earlier neural network models that relied on gradient descent with sigmoid activation functions often suffered from the vanishing gradient problem, where gradients became extremely small during backpropagation, leading to slow convergence and inefficient training. Modern deep learning architectures address this limitation by using Rectified Linear Unit (ReLU) activation functions and entropy-based loss functions, which facilitate faster training and improved performance.

Another challenge in training deep learning models is overfitting, particularly when the available training dataset is small. To overcome this issue, several regularization techniques such as dropout, bagging, and data augmentation are commonly applied. When trained on sufficiently large datasets, deep learning models are capable of achieving significantly higher classification accuracy compared with traditional machine learning approaches.

Deep learning-based pixel classification for multispectral imagery involves designing architectures that can effectively process spectral information at the pixel level. By leveraging deep learning techniques, it becomes possible to extract highly abstract and robust feature representations, thereby improving classification accuracy. These models learn hierarchical representations of input data through multiple layers, enabling them to capture complex patterns and relationships within the data.

Among deep learning architectures, Convolutional Neural Networks (CNNs) have demonstrated exceptional performance in various computer vision tasks, including image classification, object detection, and semantic segmentation. In the context of multispectral image classification, deep learning provides several advantages. It eliminates the need for manual feature engineering, effectively captures spectral and spatial dependencies, and scales efficiently to large and high-dimensional datasets. Furthermore, advanced architectures such as three-dimensional CNNs (3D CNNs), Recurrent Neural Networks (RNNs), and hybrid deep learning models have been proposed to further exploit both spectral and spatial information present in multispectral imagery.

2. Body of Paper

Proposed Approach

A. Data Collection and Pre-processing

In this study, two multispectral scenes acquired from the Landsat-8 Operational Land Imager (OLI) sensor were used. The first scene represents the New Orleans region, located at latitude 30.55580° N and longitude 89.92440° W. The second scene corresponds to the Mississippi River bottomland area, located at latitude 34.65710° N and longitude 90.40900° W.

For both scenes, five spectral bands (Bands 2, 3, 4, 5, and 7) were selected for analysis because they exhibited significant spectral variance across land-cover categories. These bands include the visible and near-infrared regions, which are particularly useful for distinguishing vegetation, water bodies, and land surfaces.

Training samples were obtained by visually inspecting the satellite scenes and selecting homogeneous regions representing different land-cover classes. These samples were used to construct the training dataset for the deep neural network (DNN) model.

For the New Orleans scene, a total of 1200 training samples were collected, consisting of 400 samples from each class. For the Mississippi River bottomland scene, the training dataset contained 3600 samples, with 900 samples collected for each class. Each pixel was represented by a feature vector consisting of reflectance values from the five selected spectral bands.

To train and evaluate the model, the dataset was randomly divided into two subsets. Seventy percent of the samples were used for training, while the remaining thirty percent were used for testing. Prior to model training, the spectral reflectance values were normalized to a range between 0 and 1 toensure stable and efficient learning.

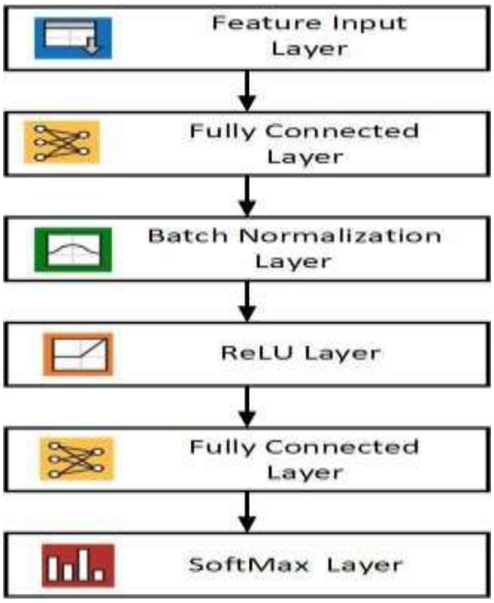

Fig. 1. Deep Neural Network (DNN) architecture.

B. Deep Neural Network (DNN) Architecture

The classification model used in this study was a simple Deep Neural Network (DNN) implemented using MATLAB's Deep Learning Toolbox. The network consisted of six layers designed to process multispectral feature vectors and classify pixels into different land-cover categories based on their spectral properties. The overall architecture of the model is illustrated in Fig. 1.

The layers of the proposed DNN architecture are described as follows:

1) Feature Input Layer

The first layer of the network is the feature input layer. In this layer, each pixel's feature vector, consisting of five spectral reflectance values corresponding to the selected spectral bands, is fed into the network. This layer accepts data with dimensionality equal to the number of spectral bands used in the analysis.

2) Fully Connected Layer

The second layer is a fully connected (dense) layer that applies learnable weights to the input features. This layer enables the network to learn relationships among spectral bands and begins the process of extracting meaningful patterns from the input data.

3) Batch Normalization Layer

A batch normalization layer follows the first fully connected layer. This layer improves training stability and accelerates convergence by normalizing the outputs of the previous layer. Batch normalization ensures that the input to subsequent layers remains centered and scaled, thereby reducing internal covariate shift and improving generalization performance.

4) ReLU Activation Layer

The next layer is a Rectified Linear Unit (ReLU) activation layer. ReLU introduces non-linearity into the network, allowing it to learn complex relationships within the spectral data. The ReLU function is applied element-wise to the outputs of the batch normalization layer, allowing only positive values to propagate through the network. Additionally, ReLU helps mitigate the vanishing gradient problem commonly observed in earlier neural network architectures.

5) Fully Connected Layer

The fifth layer is another fully connected layer that further combines and abstracts the features learned in previous layers. This layer typically contains more neurons than the earlier dense layer, enabling the model to capture more complex feature interactions within the multispectral data.

6) SoftMax Output Layer

The final layer is the SoftMax classification layer. The SoftMax function converts the raw outputs (logits) from the previous layer into a probability distribution across multiple classes. Each output value represents the probability that a given pixel belongs to a specific class, with the probabilities summing to one.

The model was trained using labeled samples extracted from homogeneous regions of the satellite images. During training, network parameters were optimized using backpropagation with stochastic gradient descent (SGD). The categorical cross-entropy loss function was used because the objective was multi-class pixel classification.

To monitor the training process and prevent overfitting, a portion of the training data was reserved for validation.

Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Training was conducted for a fixed number of epochs, with the learning rate adjusted dynamically to improve convergence. Early stopping was also employed, terminating training once validation accuracy stopped improving over several epochs.

The performance of the trained model was evaluated using several metrics, including overall accuracy, precision, recall, F1-score, specificity, and Receiver Operating Characteristic (ROC) curves. The predicted class labels were compared with the ground-truth labels from the test dataset.

Fig. 1 illustrates the proposed Deep Neural Network architecture used for multispectral image classification.

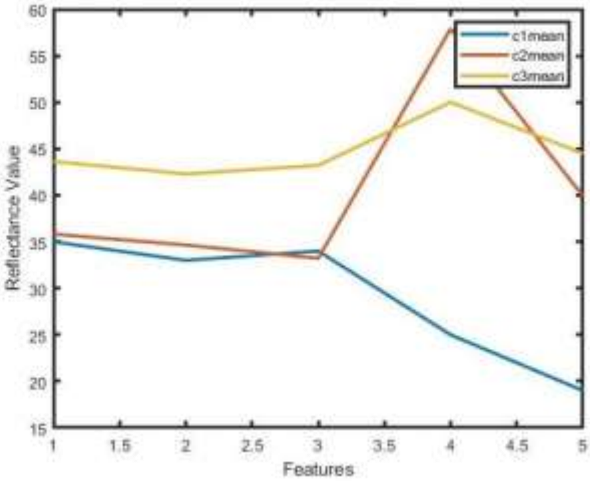

Fig. 2. Spectral signatures (New Orleans scene).

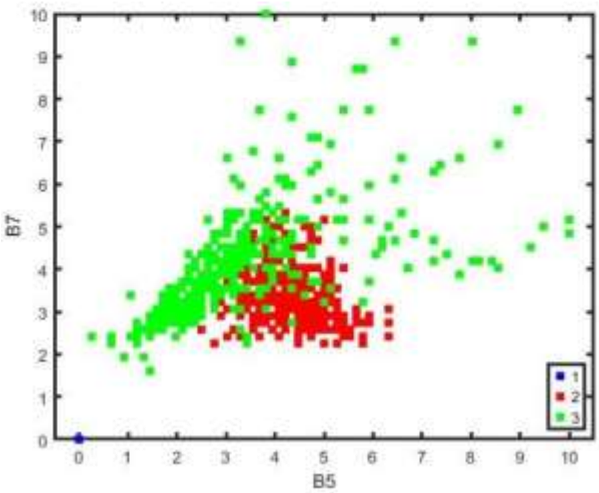

Fig. 3. Scatter plot (New Orleans scene).

Implementation and Results

In this research, the proposed DNN model was implemented using MATLAB scripts with the Deep Learning Toolbox. Two Landsat scenes were analyzed to evaluate the effectiveness of the proposed approach.

A. Example 1: New Orleans Scene

The first dataset consisted of a Landsat-8 OLI scene representing the New Orleans area. The corresponding path and row numbers were 22 and 39, respectively. A subset of the scene with a spatial resolution of 512 × 512 pixels was used for the experiments.

To generate the training dataset, three homogeneous regions were manually selected representing three land-cover classes: water, land, and vegetation. The training dataset contained 1200 samples, with 400 samples collected for each class.

Five spectral bands (Bands 2, 3, 4, 5, and 7) were used to represent each pixel. The spectral signatures obtained from the mean vectors of the selected classes are illustrated in Fig. 2, while the scatter plot of the dataset is presented in Fig. 3.

The dataset was randomly divided into training and validation sets, with 70% of the samples used for training and 30% used for validation. The proposed DNN model achieved an overall classification accuracy of **97.5%** for this scene.

The training progress curve is shown in Fig. 4, while the confusion matrix representing classification performance is shown in Fig. 5. The ROC curves and the final classified output image are presented in Fig. 6 and Fig. 7, respectively. Table I summarizes the evaluation metrics for the dataset, including recall, precision, F-score, overall accuracy, and specificity for each class.

The dataset was split into training and validation sets using a 70%–30% ratio. The proposed DNN model achieved an overall classification accuracy of **95.74%** for this scene.

3. CONCLUSIONS

In this study, a simple Deep Neural Network (DNN) model was proposed and implemented for pixel-wise classification of multispectral satellite imagery. The model was developed using MATLAB and evaluated on two Landsat-8 scenes: one representing the New Orleans region and the other representing the Mississippi River bottomland area. The Landsat scenes used in this research are publicly available datasets, making the approach easily reproducible.

Training samples were generated by selecting small homogeneous regions within the satellite scenes, representing different land-cover categories. Each pixel was represented by a feature vector containing reflectance values from selected spectral bands, which were used as input to train the DNN model. The New Orleans dataset consisted of three classes, while the Mississippi River bottomland dataset contained four classes.

Experimental results demonstrated that the proposed DNN model achieved high classification performance, with overall accuracies of **97.44%** for the New Orleans scene and **95.74%** for the Mississippi River bottomland scene. These results indicate that deep neural network models provide an effective and powerful alternative to traditional machine learning techniques for analyzing multispectral remote sensing data.

It should be noted that the current study focuses only on spectral information and does not incorporate spatial relationships between neighboring pixels. Integrating both spectral and spatial features could potentially improve classification performance. Future research will focus on extending the proposed approach to hyperspectral imagery with a larger number of spectral bands and evaluating the model using widely recognized benchmark datasets.

Furthermore, the proposed DNN architecture can be easily adapted for hyperspectral data by increasing the dimensionality of the input layer and the capacity of the fully connected layers. Although convolutional neural network (CNN)-based hybrid models have been widely used for pixel-based classification, they are generally more complex and computationally intensive. Therefore, future work will also involve comparing the performance of the proposed DNN model with CNN-based hybrid architectures to assess their relative efficiency and accuracy.

.

ACKNOWLEDGEMENT

The heading should be treated as a 3rd level heading and should not be assigned a number.

REFERENCES

[1] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 3, pp. 247–259, 2011.

[2] A. D. Kulkarni and A. Shrestha, "Multispectral image analysis using decision trees," *International Journal of Advanced Computer Science and Applications*, vol. 8, no. 6, pp. 11–18, 2017. doi: 10.14569/IJACSA.2017.080602.

[3] L. Zhang, L. Zhang, and B. Du, "Deep learning for remote sensing data: A technical tutorial on the state of the art," *IEEE Geoscience and Remote Sensing Magazine*, vol. 4, no. 2, pp. 22–40, 2016. doi: 10.1109/MGRS.2016.2540798.

[4] X. X. Zhu et al., "Deep learning in remote sensing: A comprehensive review and list of resources," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 4, pp. 8–36, 2017. doi: 10.1109/MGRS.2017.2762307.

[5] M. Zhu, Y. He, and Q. He, "A review of research on deep learning remote sensing applications," *International Journal of Geosciences*, vol. 10, pp. 1–11, 2019. doi: 10.4236/ijg.2019.101001.

[6] J. S. Song, S. Gao, Y. Zhu, and C. Ma, "A survey of remote sensing image classification based on CNNs," *Big Earth Data*, vol. 3, no. 3, pp. 232–254, 2019.

[7] E. M. Rewhel, J. Q. Li, and A. A. Hamed et al., "Deep learning methods used in remote sensing images: A review," *Journal of Environmental & Earth Sciences*, vol. 5, no. 1, pp. 33–64, 2023.

[8] X. Wu, D. Hong, and J. Chanussot, "Convolutional neural networks for multimodal remote sensing data

classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–10, 2022. doi: 10.1109/TGRS.2021.3124913.

[9] S. Paheding et al., "Advancing horizons in remote sensing: A comprehensive survey of deep learning models and applications in image classification and beyond," *Neural Computing and Applications*, vol. 36, pp. 16727–16767, 2024.

[10] W. Hu, Y. Huang, L. Wei, F. Zhang, and H. Li, "Deep convolutional neural networks for hyperspectral image classification," *Journal of Sensors*, vol. 2015, Article ID 258619, 2015.

[11] E. Maggiori et al., "Convolutional neural networks for large-scale remote sensing image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 2, pp. 645–657, 2016.

[12] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geoscience and Remote Sensing Letters*, vol. 14, no. 5, pp. 778–782, May 2017.

[13] A. D. Kulkarni, "Multispectral image analysis using convolution neural networks," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 10, pp. 13–19, 2023.

[14] R. Li, S. Zheng, C. Duan, L. Wang, and C. Zhang, "Land cover classification from remote sensing images based on multi-scale fully convolutional network," *Geo-Spatial Information Science*, vol. 25, no. 2, pp. 278–294, 2025.

[15] Z. Lian et al., "Recent advances in deep learning-based spatiotemporal fusion methods for remote sensing images," *Sensors*, vol. 25, no. 4, pp. 1093, 2025.

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Lecture Notes in Computer Science, vol. 9351, Springer, 2015.

[17] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, 2015, pp. 3431–3440.

[18] J. V. Solórzano, J. F. Mas, Y. Gao, and J. A. Gallardo-Cruz, "Land use land cover classification with U-Net: Advantages of combining Sentinel-1 and Sentinel-2 imagery," *Remote Sensing*, vol. 13, no. 18, 2021.

[19] H. M. Sahin et al., "Segmentation of weeds and crops using multispectral imaging and CRF-enhanced U-Net," *Computers and Electronics in Agriculture*, vol. 211, pp. 107956, 2023.

[20] W. Wiratama, J. Lee, and D. Sim, "Change detection on multi-spectral images based on feature-level U-Net," *IEEE Access*, vol. 8, pp. 12279–12289, 2020.

[21] T. Bai et al., "Deep learning for change detection in remote sensing: A review," *Geo-Spatial Information Science*, vol. 26, no. 3, pp. 262–288, 2022.

[22] U.S. Geological Survey, "Landsat Data Access," Available: https://www.usgs.gov/landsat-missions/landsat-data-access