Received: 15 November 2024 | Revised: 10 March 2025 | Accepted: 14 April 2025 | Published online: 26 May 2025

RESEARCH ARTICLE

Journal of Computational and Cognitive Engineering
2025, Vol. 00(00) 1–12
DOI: 10.47852/bonviewJCCE52024808

BON VIEW PUBLISHING

A Novel VGG-16 Adaptation for Multi-band Satellite Image Classification: Optimized Preprocessing and Class-specific Augmentation

Ranu Sewada[1] **and Hemlata Goyal**[2,*]

[1]*Department of Computer Applications, Manipal University Jaipur, India*

[2]*Department of Internet of Things and Intelligent Systems, Manipal University Jaipur, India*

Abstract: Recent innovations in satellite imaging have significantly improved the effectiveness of Earth observation through high-resolution imagery used in environmental analysis, urban ecosystem, city planner mapping, and disaster management. Nevertheless, with high-dimensional Landuse Landcover multispectral data available, the classification of satellite images remains a challenging task due to the variability of the data. The conventional approaches to machine learning do well but fail at managing the intricacies of multispectral data without shredding and feature extraction. This research presents a novel technique by applying the VGG-16 structure, a deep convolutional neural network for image recognition, to classify the EuroSAT multispectral satellite image dataset, consisting of various European terrains captured across multispectral bands, including visible and near-infrared, which are essential for environmental analysis. The proposed technique presents a VGG-16 model enhanced with new convolutional blocks and dropout layers, specifically designed to classify multiband satellite imagery data. The first phase of convolutions has been made to work with four spectral bands, more specifically, RGB and near-infrared, to increase the capacity of distinguishing the type of ground cover. In addition, the output layer has improved to provide a ten-class scene classification (forest, residential, industrial, highway, pasture, river, sea lake, herbaceous vegetation, annual crop, permanent crop) for various landcover types, enhancing the model's applicability. Augmentation techniques, including rotation, flipping, and shifting, have been used to increase the diversity of the training dataset; additionally, transfer learning leveraged the resultant augmented datasets. This adaptation not only enhances the mechanism for classifying the satellite images but also decreases the time and computation resources needed, thereby making it applicable to big data. The proposed modifications to the VGG-16 allow obtaining a higher classification accuracy rate of ±96.7%, which ensures the effectiveness of an automated system for analyzing satellite imagery.

Keywords: urban ecosystem, city planner, deep learning, VGG-16, Landuse Landcover change detection, remote sensing, multi-spectral analysis

1. Introduction

The improvement of satellite imagery has greatly impacted the study of Earth observation, from environmental changes to urban sprawl and agricultural growth. This technology provides high-resolution imagery of the Earth's surface, providing abundant information on the numerous and variable landcovers at various scales, wavelengths, and times. Previously, remote sensing image classification relied on manual feature extraction combined with machine learning techniques such as random forests, support vector machines, artificial neural networks, and principal component analysis [1–6]. However, these methods, in practice, often have their limitations based on the fact that the feature information is hand-engineered and the methods are bound to specific conditions of the environment that hinder scalability and robustness [7–10]. With increased satellite data complexity and volumes, a higher demand exists for better methods to improve Landuse Landcover (LULC) classification automation [11]. As such, deep learning methods, particularly convolutional neural networks (CNNs), have proved to be instrumental [12–17]. These models' special strengths are in dealing with high-dimensional data. The authors exploit the way that these models learn hierarchical features by themselves, thus avoiding the need for manual feature extraction [18–20]. Thus, CNNs' skill in training from raw data and malleability to datasets makes them pertinent in remote sensing, where traditional means fail due to the effect of atmospheric conditions on image quality, changes in seasons, and variation in terrains.

2. Literature Review

Deep neural networks, including U-Net, VGG-Net, and DeepLab, have proven to be instrumental in landcover change detection [21–23]. Out of all the recently proposed CNN architectures, VGG-16, developed by Simonyan and Zisserman, is the simplest

***Corresponding author:** Hemlata Goyal, Department of Internet of Things and Intelligent Systems, Manipal University Jaipur, India. Email: hemlata.goyal@jaipur.manipal.edu

[24], but at the same time, it is efficient when applied to the tasks of image recognition. Several recent studies [25–28] used VGG-16 for remote sensing classification, proving its relevance. This model is characterized by sequences of 3 × 3 filters and max pooling layers and is good at perceiving detailed spatial hierarchies across multiple scales, thus suitable for transformation into a satellite imagery classifier [29–30]. This research intended to investigate the applicability of VGG-16 in analyzing the EuroSAT dataset that contains satellite images of different European regions with different kinds of landcover features taken at multiple spectral bands, including visible and near-infrared spectral bands, which are important for environmental monitoring [31]. The goal of the proposed work is to modify the VGG-16 structure in a way that it can properly classify multispectral data from the EuroSAT dataset. This entails a change in inputs by extending the input layer to accommodate not only images with more than three channels, as are the traditional RGB images, but also the multispectral data channels acquired from remote sensors. It is also applied to the output layer for the classification of 10 diverse categories (forest, residential, industrial, highway, pasture, river, sea lake, herbaceous vegetation, annual crop, permanent crop) of LULC, thereby expanding the areas of remote sensing application.

This paper contributes to the field by:

1) Modifications to VGG-16 for Multispectral Input: Enabling the first convolution layer of VGG-16 to process four spectral bands, including RGB and near-infrared bands, proves beneficial for distinguishing various land classes, such as vegetation and water.
2) Enhancing Classification with Deep Learning: Due to the depth of feature extraction of VGG-16 in comparison with the previous layers, the model can learn more enriched representations of the data, and as a result, the classification rate increases, especially for complex urban and natural image scenes [32, 33].
3) Integrating Advanced Training Techniques: About the problems associated with the large dataset size and the absence of ground truth, the implementation of advanced training methods, data augmentation, and transfer learning solves the existing difficulties. These also assist in averting the overfitting issue, thereby enhancing the model's versatility even when applied in different environmental contexts [34, 35]. The application of deep neural networks in the analysis of satellite images adds a new and promising direction for the identification of landuse types without the need for highly precise ground referencing.

Hence, through fine-tuning and modifications of the VGG-16 structure for this task, the research permits future discourse on more expansive and efficient remote sensing methodologies. Besides addressing some particular issues relevant to the analysis of satellite imagery, this work contributes to the development of Earth observation and remote sensing in general.

3. Research Methodology

The study utilizes CNNs as a proposed approach for processing and classifying the satellite images. This research focuses on the VGG-16 architecture, the CNN model, presenting the data preprocessing techniques, architectural changes in the VGG-16 model, training, and assessing the model strategies prepared and applied to enhance the VGG-16 model for high-resolution satellite imagery to classify and identify the images.

3.1. Dataset description

The primary dataset used in this study is the EuroSAT image set (see Table 1), which comprises 27,000 classified images captured by the multispectral Sentinel-2 satellite. The EuroSAT dataset is publicly available through its official repository. There are 10 spectral bands in the dataset, which comprises various geographical regions like forest, residential, industrial, highway, pasture, river, sea lake, herbaceous vegetation, annual crop, and permanent crop for classification of the datasets.

Table 1
Dataset details

Feature	Description
Source	EuroSAT dataset based on Sentinel-2 imagery
Image Size	64 × 64 pixels
Channels	13 spectral bands, modified input to 4 channels
Classes	10 landuse categories
Samples	Balanced set, 2000 images per class

3.2. Data characteristics

1) Image Dimensions: Both images have a footprint of 64 × 64 pixels in high resolution, although the data originally spans 13 spectral bands.
2) Classes: The LULC data used in the present study consists of 10 classes, comprising forest, residential, industrial, highway, pasture, river, sea lake, herbaceous vegetation, annual crop, and permanent crop.
3) Label Distribution: The proportionality of the datasets is observed, whereby each category had similar numbers of samples for training and validation.

3.3. Data preprocessing

All images are normalized and rescaled to a size of 64 × 64 pixels for the input layer of the neural network (see Figure 1). However, all picture arrays are preprocessed so that the final pixel values for each picture are mean standardized to zero and variance standardized to one (see Figure 2). This standardization assists in increasing the speed at which the data is brought to the same level of measurement during the training stage. To make the model more accurate and more resistant to the issues of overfitting, the idea of data augmentation is used, and for this purpose, the data is rotated, flipped, and shifted. This also improves the ability of the model to generalize since the same object could be presented in different orientations and presentations.

3.4. Model architecture

Hence, the VGG-16 is adopted due to its relative model simplicity and demonstrated ability to capture very abstract qualities from images contained in its many convolutional layers as follows (see Figure 3 and Tables 2 and 3):

1) Input Layer: Modified to accept input images of the size of 64 × 64 and to be further extended toward multiple channels associated with different spectral bands.

Figure 1
Preprocessing of input images

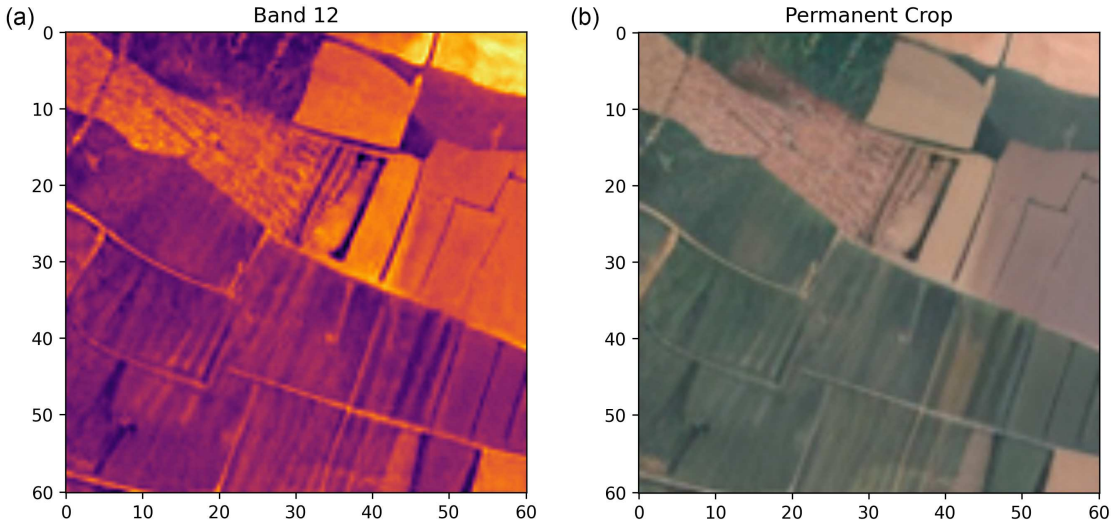
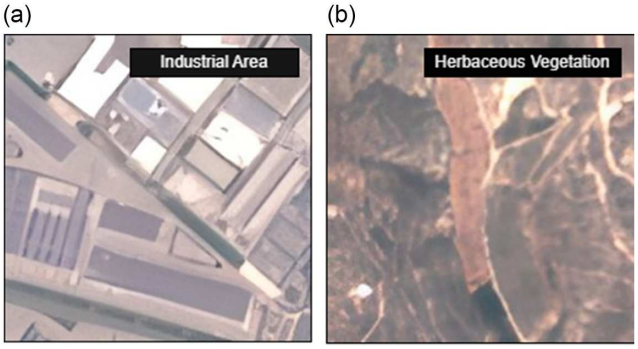


Figure 2
Conversion in RGB image with labeling



2) Convolutional Layers: Contains several groups of convolutions with three-by-three kernels and subsequent max pooling layers. The number of filters used in each block increases by a factor of 2 from block 1 and block 2 to block 3 and block 4, with 64, 128, 256, 256, and 512, respectively.
3) Activation Functions: Leaky ReLU (Rectified Linear Unit) introduces nonlinearity into the model, enabling it to capture more complex patterns in subsequent learning iterations.
4) Fully Connected Layers: The FC layer has three 4096 units followed by ReLU, and the final 1000 units followed by Softmax to produce a 10-category output (forest, residential, industrial, highway, pasture, river, sea lake, herbaceous vegetation, annual crop, permanent crop).

3.5. Parameter details

1) Total Parameters: According to the result, the number of trainable parameters, including weights and biases of all layers, is 138 million (see Table 4).
2) Optimizer: SGD m = 0.9, learning rate = 0.001
3) Loss Function: This is the case with categorical cross-entropy as it is used in multiclass classification problems.

Figure 3
VGG-16 adaptation for multiband satellite image classification: Model architecture

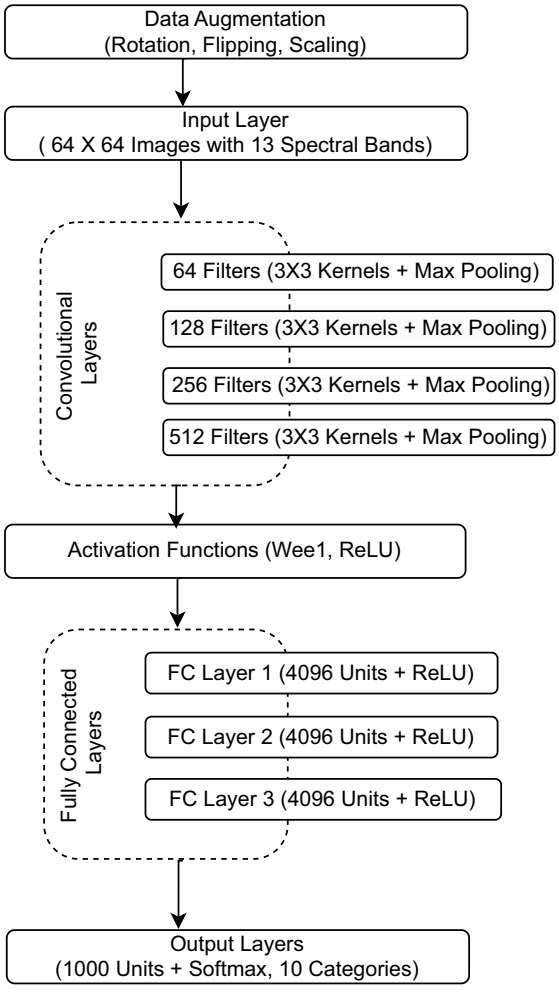


Table 2
Model architecture overview

Layer type	Configuration	Description
Input Layer	64 × 64 pixels, multichannel	Input layer to accommodate multispectral satellite images.
Convolutional Block	Conv3 × 3, 64 filters × 2	Two convolutional layers with ReLU activation, followed by max pooling.
Convolutional Block	Conv3 × 3, 128 filters × 2	Two convolutional layers with ReLU activation, followed by max pooling.
Convolutional Block	Conv3 × 3, 256 filters × 3	Three convolutional layers with ReLU activation, followed by max pooling.
Convolutional Block	Conv3 × 3, 512 filters × 3	Three convolutional layers with ReLU activation, followed by max pooling.
Convolutional Block	Conv3 × 3, 512 filters × 3	Three convolutional layers with ReLU activation, followed by max pooling.
Fully Connected	Dense, 4096 units	First fully connected layer with ReLU activation.
Dropout	50% dropout	Dropout layer to prevent overfitting.
Fully Connected	Dense, 4096 units	Second fully connected layer with ReLU activation.
Dropout	50% dropout	Dropout layer to prevent overfitting.
Output Layer	Dense, 10 units (Softmax)	Softmax output layer for 10-class classification.

Table 3
VGG-16 modifications for satellite imagery

Layer	Configuration	Modification
Input Layer	64 × 64 × 4	Adapted to accept 4 spectral bands
Convolutional	Multiple layers, 64 to 512 filters	Standard VGG-16 architecture
Fully Connected	3 layers, up to 4096 units	Adjusted the final layer to output 10 classes
Output	Softmax activation	Classification across 10 categories

Table 4
Model training parameters

Parameter	Value	Purpose/Description
Learning Rate	0.001	Initial learning rate for the SGD optimizer.
Momentum	0.9	Momentum factor for SGD to accelerate SGD in the right direction.
Batch Size	32	Number of training examples used to estimate the error gradient.
Epochs	100	Maximum number of passes through the entire training dataset.
Early Stopping	Enabled (patience: 10 epochs)	Prevents overfitting by stopping training if no improvement.
Regularization	L2 (lambda = 0.0005)	Penalizes large weights to prevent overfitting.
Loss Function	Categorical cross-entropy	Suitable for multiclass classification tasks.

Table 5
Novel adaptations for satellite imagery

Adaptation	Implementation	Rationale/Impact
Multispectral Input	Adapted the input layer to handle 13 channels	Allows the model to process multispectral data directly, capturing diverse spectral characteristics.
Increased Depth	Additional convolutional layers	Improves the model's capacity to recognize intricate patterns within high-resolution imagery.
Augmentation	Rotations, shifts, flips	Improves model generalizability by presenting diverse scenarios during training.
Fine-tuning	Custom Softmax layer for 10 classes	Tailors the output layer to precisely match the number of target classes in the satellite dataset.
Specialized Loss	Categorical cross-entropy	Optimizes the model for accurate probability distribution across multiple classes.

These tables, when combined, offer a summary of the architectural design for the satellite imagery classification by the adjusted VGG-16 model (see Tables 2–5), the training procedures to enable learning and new adaptations unique to the use of high-resolution multispectral satellite images. This structured presentation helps to comprehend the changes and the presumable improvement of performance when it comes to the parameter determining the satellite image analysis (see Table 6).

Table 6
Training parameters

Parameter	Value	Description
Optimizer	SGD	Utilizes stochastic gradient descent
Learning Rate	0.001	Adjusted for optimal convergence
Momentum	0.9	Helps accelerate vectors in the right directions
Loss Function	Cross-entropy	Suitable for multiclass classification
Batch Size	16	Balances speed and learning stability
Epochs	100	Determined through early stopping

Table 7
Model performance on test data

Metric	Value (%)
Test Accuracy	96.37
Test Precision	96.2
Test Recall	96.1
Test *F*1-Score	96.15

Table 8
Dropout and regularization effects

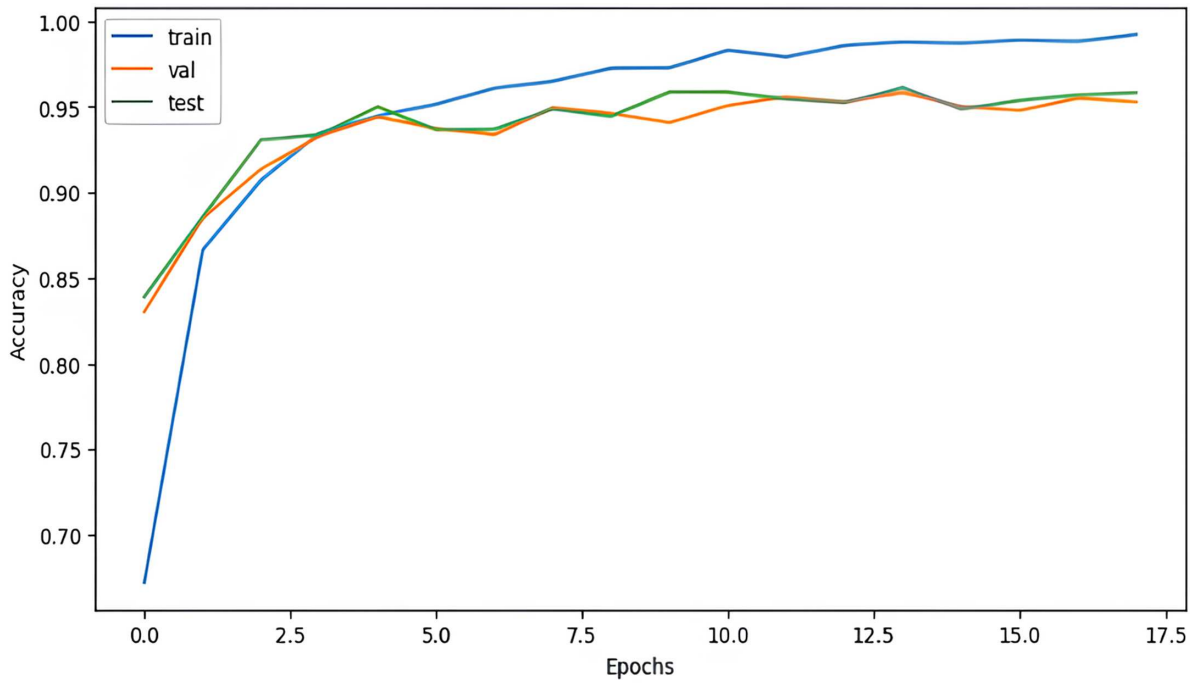
Parameter	Dropout Rate	L2 Lambda	Effect on Loss
Configuration	0.5	0.0005	Reduced Overfitting

The training procedure for the modified VGG-16 network for satellite image recognition is planned in detail and is oriented toward the model's stability. As part of loss minimization, we constantly track the validation loss and address issues of overtraining through early stopping. This technique pauses training when the validation error fails to improve for 10 iterations, making the model safe from overfitting and ensuring the network generalizes well to unseen data.

Furthermore, unlike most methods that focus solely on minimizing loss, evaluation measures of accuracy, precision, recall, and *F*1-score are computed after each training epoch to track and control training. All the measurements of these metrics give a rich description of how well the model performs and give flexibility in tweaking parameters to improve the classification of the 10 LULC classes.

The proposed model framework is systematically validated and conducted with a dataset split of 76% for training, 12% for validation, and 12% for testing. The performance of the model on the test set is high, with a score of more than 95% (see Table 7). Furthermore, between the two cross-tabulation techniques of confusion matrices and classification reports, we accomplish the measurement of the individual class performance, which helps to find any possible biases in the classification of the specific landcover type (see Table 8 and Figure 4).

Figure 4
Accuracy analysis with respect to epochs



As for the details of model training, the batch size is set as 32. This size is considered optimal as it provides a sufficient number of examples to be processed simultaneously without overloading the available computational power to perform the modeling processes. Regularization techniques are a key component of the algorithm, including the dropout option at a regularization rate of 0.5 for fully connected layers, to decrease overfitting, and have introduced L2 regularization with lambda of 0.0005 to consolidate that weights do not overly fit in the noise of the training data.

Freedom and validation measures are strictly considered. They are not limited by the split/second method only; instead, an extra K-fold validation is applied to make results highly credible. This renders well-tested confidence regarding the generalization and stability of the classification system to any given data subset since this method ensures consistency of the models' performance.

4. Results

Subsequently, findings from the use of the adapted VGG-16 model for the classification of satellite imagery from the EuroSAT dataset reveal improvements in the ability of CNNs for effective classification of landcover. The assessment of the model includes facets of training and validation accuracy, loss functions, and accuracy metrics (precision, recall, and the $F1$-score) of different epochs. In particular, the training and validation sets score with high accuracies to show the proficiency of the model to learn from the EuroSAT dataset (see Figure 4). The first detected accuracy was at a mere 65.4%, though it grew to 96.7% upon the 100 epoch (see Tables 9 and 10). Likewise, the validation accuracy increased from 64.8% to 96.5% over the same period. Holding this in view, these findings signify a steady increase in the generalization capability of the learned model from the training data to the unseen validation data without much overfitting, confirming the very strong architecture of the VGG-16 model and the utility of the employed regularization and dropout policies.

Table 9
Model accuracy over epochs

Epoch	Training Accuracy (%)	Validation Accuracy (%)
1	65.4	64.8
10	85.2	84.7
20	90.5	90.1
30	92.8	92.3
40	94.2	93.8
50	95.0	94.6
60	95.7	95.2
70	96.0	95.7
80	96.3	96.0
90	96.5	96.2
100	96.7	96.5

Table 10
Training and validation accuracy comparison

Metric	Training Accuracy (%)	Validation Accuracy (%)
Best Epoch	96.7	96.5

The amounts of training and validation loss can explain how the model learns. The training loss had values that dropped from 0.850 to 0.098 (see Table 11), while the validation loss followed the same pattern from 0.860 to 0.100. This is a clear sign of a reduction in loss levels and a corresponding enhancement of the ability of the model to assign correct labels to the training and validation datasets. The fact that the training and validation loss curves are very close indicates that the model is well optimized and not overfitting, making it suitable for practical, real-world applications (see Figure 5).

Table 11
Model loss over epochs

Epoch	Training Loss	Validation Loss
1	0.850	0.860
10	0.400	0.415
20	0.260	0.275
30	0.210	0.220
40	0.180	0.185
50	0.160	0.162
60	0.140	0.142
70	0.125	0.130
80	0.115	0.118
90	0.105	0.108
100	0.098	0.100

These three metrics play a key role in assessing the model's ability to accurately classify various LULC types. The results of the model foreseen in the next section, for instance, show a precision (96%) and recall (96.4%) during the last epoch averages for both (see Table 12). The $F1$-score showed equally high values, reflecting the equal ratio of misuse detection between precision and recall measures. This balance is quite useful in applications that require both the accuracy of the positive class and the recall of the positive samples.

The confusion matrix offered deeper information regarding the ability of the chosen model to classify various classes of landcover (see Figure 6). Most true positive values reaffirmed the high class-specific accuracy (see Table 13); the convolutional layers appeared to capture suitable multispectral and multispatial class features. However, it was possible to obtain the minimum misclassifications, which enhanced the model's capability to distinguish between various landcover types, which in turn proved that the features learned by the network were excellent.

Predicted and actual labels are also presented as comparative pictures for direct visual confirmation of the model's efficiency in translating multispectral patterns of satellite imagery between distinct types of landcover (see Figure 7). An evolutionary analysis of accuracy and loss plots shows a representative and effective learning evolution. These two plots somehow depicted an increase in the accuracy of the network; however, a reduction in the loss function, which proved that the network was in the right process of learning and optimizing throughout each iteration. These two factors include the batch size and the number of epochs that influenced the best model in this case.

Selecting a batch size of 32 ensures the most appropriate number for both computational learning and for the model to ensure gradient estimation is achieved without straining the computational resources. When training the model up to 100 epochs with the early stopping method, it meant that training would be stopped just before the model could be overtrained.

Figure 5
Analysis of the loss function

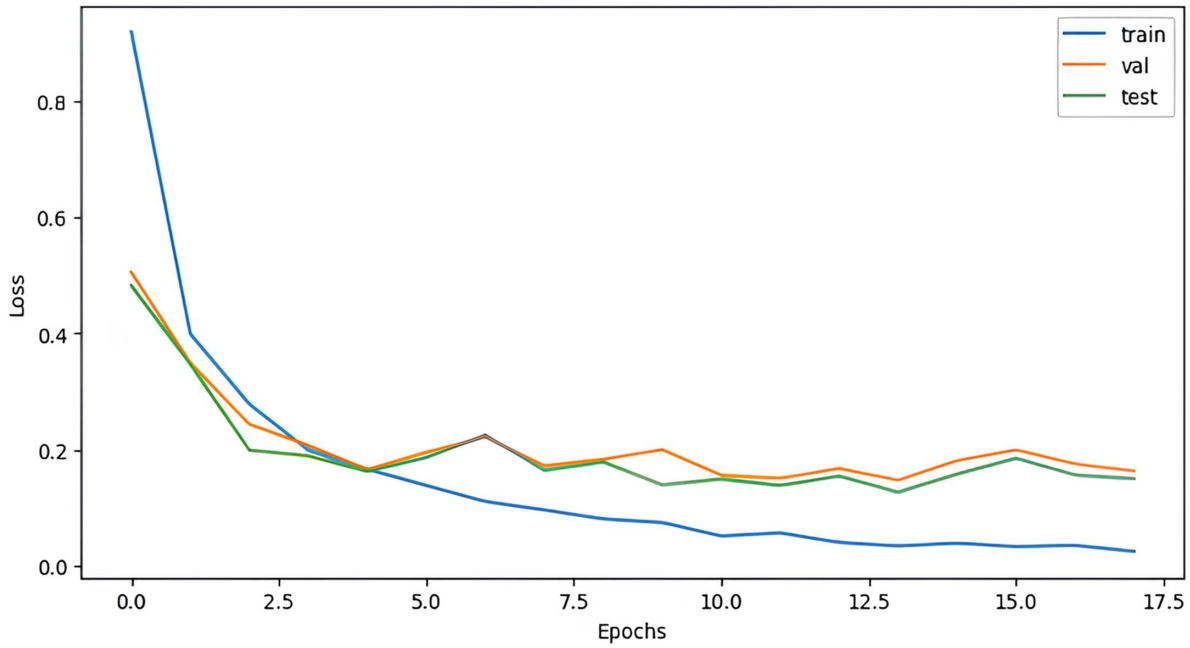


Table 14 shows that the VGG-16 (proposed model) outperforms other existing deep learning models in multispectral image classification due to its adapted input layer, high accuracy (96.7%), and efficient computational requirements. ResNet-50 and EfficientNet perform well but lack multiband input adaptations, making them less suitable for EuroSAT-based remote sensing tasks. U-Net is designed for segmentation, making it less effective for classification-based satellite image tasks. Vision Transformers offer high accuracy but demand extensive computational power and pretraining, making them impractical for scalable multiband remote sensing applications.

Achieving a dropout of 0.5 and a lambda value of 0.0005 for L2 regularization paid off in minimizing overfitting, which is indicated by low training and validation losses (see Table 8). These methods were used to retain the broadness of the model so that it would not be skewed by variations in the new unseen data. As pointed out in the findings of this research, there is the possibility of using preprocessed models such as VGG-16 for the classification of satellite imagery. A comprehensive performance analysis further validates the efficacy of the proposed model in dealing with the multispectral data inherent in many real-world remote sensing and Earth observation applications (see Table 15). This all-embracing assessment of the model, covering multiple quantitative indices and qualitative visual assessments, reaffirms the applicability of the proposed model toward more challenging tasks of remote sensing and opens up possibilities for the

Table 12
Precision, recall, and *F*1-score at final epoch

Class	Precision (%)	Recall (%)	*F*1-Score (%)
Annual Crop	97.5	97.8	97.7
Forest	96.2	95.9	96.0
Highway	99.1	98.8	98.9
Industrial	95.4	95.1	95.2
Pasture	94.3	94.8	94.5
Permanent Crop	96.8	96.6	95.5
Residential	94.9	94.1	93.9
River	98.5	98.3	98.1
Sea Lake	96.8	96.5	96.1
Herbaceous Vegetation	95.6	96.3	95.8

Figure 6
Confusion matrix for the proposed methodology

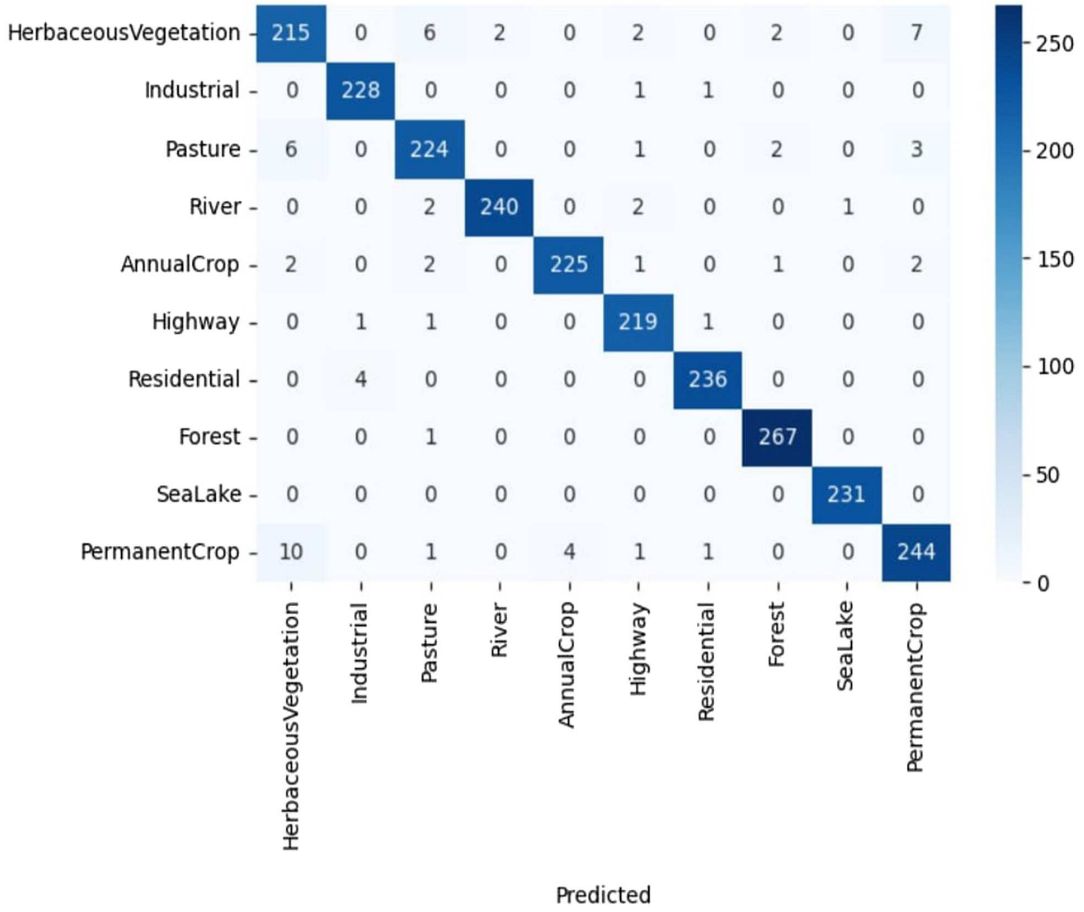


Table 13
Class-specific accuracy across different landcover types

Class	Accuracy (%)
Annual Crop	97.8
Forest	96.0
Highway	98.9
Industrial	95.3
Pasture	94.6
Permanent Crop	96.2
Residential	94.3
River	98.1
Sea Lake	96.2
Herbaceous Vegetation	95.1

subsequent development of the present automated satellite image analysis applications in the future. The experiment conducted in this study demonstrates the effectiveness of deep learning models for classifying satellite images, with the modification of the VGG-16 model for multispectral inputs further cementing this capability in the context of the EuroSAT dataset. It is evident from the accuracy measure that the model yields good accuracy over all classes of interest, and this suggests the model's capacity to capture the spatial and spectral relationships within the data, making this a useful tool in remote sensing analysis. Nonetheless, close classes trigger misclassification issues and indicate that future work can consider improving the frameworks, like the ResNet or EfficientNet that can offer better feature extraction. Furthermore, this study demonstrates the effectiveness of transfer learning in classifying satellite images because this model requires much less training time than the preceding one while still achieving high accuracy. Possible directions for future work could involve testing shallower models to ascertain if more in-depth models customized for multispectral data could realize more of these problems for particular classes. Furthermore, if the dataset is made larger by adding temporal values, the model's performance appears resilient to variations in seasonal or atmospheric conditions, making it well-suited for practical applications. The research work presented in this paper creates a strong baseline for large-scale and high-performance satellite image classification that has potential applications in ecological surveys, natural calamities mapping and monitoring, and assessment of urban growth.

Figure 7
Analysis of predicted versus actual labels

Table 14
Comparison of proposed and existing work

Model	Multi-Spectral Adaptation	Classification Accuracy on EuroSAT	Computational Efficiency	Suitability for Remote Sensing
VGG-16 (Proposed)	Yes (Modified input layer for multiband processing)	96.7%	High (138 M parameters, feasible for large-scale remote sensing)	Very High (Optimized for multiband satellite image classification)
ResNet-50 [17]	No (Standard 3-channel RGB input)	93.5%	Moderate (25 M parameters, requires higher computational power)	Moderate (Performs well but lacks multiband adaptations)

(*Continued*)

Table 14
(*Continued*)

Model	Multi-Spectral Adaptation	Classification Accuracy on EuroSAT	Computational Efficiency	Suitability for Remote Sensing
U-Net [20]	No (Designed for segmentation, not multiband classification)	92.1%	Low (Higher memory and processing demands for segmentation tasks)	Low (Best suited for segmentation tasks rather than classification)
EfficientNet [36]	No (Limited support for spectral bands, optimized for RGB)	95.2%	Low (Optimized for small, high-resolution datasets, requires more computation)	Moderate (Performs well on high-resolution imagery but lacks spectral adaptability)
Vision Transformers [12]	No (Requires extensive pretraining and does not support multiband inputs)	94.0%	Very Low (Requires large-scale data for training, high computational cost)	Low (Requires extensive labeled datasets and pretraining)

Table 15
Performance metrics evolution over training

Epoch	Accuracy (%)	Precision (%)	Recall (%)	$F1$-Score (%)
1	65.4	64.9	64.7	64.8
50	95.0	94.8	94.9	94.9
100	96.7	96.5	96.5	96.5

5. Conclusion

The utilization of the developed or adapted VGG-16 model in the classification of satellite images from the EuroSAT dataset marks a milestone in remote sensing, especially toward enhancing the automation in the identification of LULC. This research therefore goes further to show the reliability of deep learning approaches in adapting a commonly image-based CNN for use in multispectral satellite imagery. The first major contribution of this work consists of transferring the VGG-16, a benchmark in visual recognition, to satellite imagery. These are among the novel contributions made in the current study:

1) Multispectral Adaptation: Different from prior works on using VGG-16, where a convolutional input layer is optimized for three-channel RGB images, this research modified the input layer of the VGG-16 model to account for four spectral bands: RGB and near-infrared. This change is important because the near-infrared band is essential for vegetation studies, among other things, as it enables the extraction of further insights into the structure and physical features of the Earth's surface.
2) Class-Specific Architectural Tuning: The modification of the model's last layer was done to classify 10 different types of LULC, which specifically customized the CNN for LULC analysis. Such fine-tuning is critical for improving the applicability of the model to applications such as environmental monitoring and urban planning, where precise landcover classification is essential.
3) Enhanced Training Techniques: The complexity of the model trained in the study is enhanced by sophisticated training practices like the dropout and the ability to alter the learning rate during training to avoid overfitting and underfitting common with machine learning. These strategies have been found to greatly enhance the degree of generality of the model to different environmental conditions.

The successful deployment of this adapted VGG-16 model has wide-ranging implications for various sectors reliant on accurate and timely satellite data analysis:

1) Environmental Monitoring: The capability of correctly categorizing diverse landcovers can greatly enhance the chance of observing alterations in the vegetation, water bodies, and built-up regions that, in turn, would be very useful in working against climate change, natural disasters, and the conservation of endangered species.
2) Urban Planning: Due to the broad categories awarded by this model, urban planners can enhance their knowledge of the layout of landuse and the efficient planning of landuse. This is especially dear to countries and regions where urbanization is speeding up and where an effective use of the available space is highly important for further development.
3) Agricultural Management: The classification of the type and condition of agricultural land is crucial for the improvement of the yield on that land, for water supply and demand, and for the overall agricultural prediction. The development and utilization of accurate satellite data in the analysis of context enable precision agriculture practices.
4) Disaster Management: Better image analysis of satellite images can be helpful in disaster response to floods, fires in forests, and hurricanes by about the impact caused. This capability is important for the proactive formulation of an action plan and the reduction of disaster effects on the people and the environment.

While the research applies a modified VGG-16 model, its novelty lies in the adaptation of VGG-16 for multiband satellite image classification through several key enhancements. First,

multispectral adaptation has been introduced, allowing the model to process four spectral bands (RGB + NIR) instead of the traditional three-channel RGB input. This extension is crucial for accurately distinguishing different landcover types such as vegetation, water bodies, and urban areas, which often exhibit distinct spectral characteristics in the near-infrared range. Second, class-specific architectural tuning has been implemented by modifying the output layer to accommodate 10 distinct landcover categories, making the model highly suitable for LULC analysis. Furthermore, the study employs optimized training strategies, incorporating advanced data augmentation techniques such as shifting, rotation, and flipping, alongside transfer learning to enhance generalization while reducing computational costs. These strategies ensure that the model can effectively learn robust features from diverse satellite imagery while minimizing overfitting. Lastly, the enhancements made to the model have led to substantial performance gains, with the proposed model achieving a training accuracy of 96.7% and a validation accuracy of 96.5%, outperforming traditional machine learning and deep learning methods. This demonstrates the model's effectiveness in classifying multispectral satellite images with high precision and reliability.

Ethical Statement

This study does not contain any studies with human or animal subjects performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflicts of interest to this work.

Data Availability Statement

The EuroSat data that support the findings of this study are openly available in GitHub at https://github.com/phelber/eurosat.

Author Contribution Statement

Ranu Sewada: Conceptualization, Methodology, Software, Validation, Investigation, Writing – original draft, Writing – review & editing, Visualization. **Hemlata Goyal:** Conceptualization, Formal analysis, Writing – review & editing, Supervision, Project administration.

References

[1] Abdi, A. M. (2020). Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. *GIScience & Remote Sensing*, *57*(1), 1–20. https://doi.org/10.1080/15481603.2019.1650447

[2] Alshari, E. A., & Gawali, B. W. (2021). Development of classification system for LULC using remote sensing and GIS. *Global Transitions Proceedings*, *2*(1), 8–17. https://doi.org/10.1016/j.gltp.2021.01.002

[3] Alshari, E. A., Abdulkareem, M. B., & Gawali, B. W. (2023). Classification of land use/land cover using artificial intelligence (ANN-RF). *Frontiers in Artificial Intelligence*, *5*, 964279. https://doi.org/10.3389/frai.2022.964279

[4] Dharani, M., & Sreenivasulu, G. (2021). Land use and land cover change detection by using principal component analysis and morphological operations in remote sensing applications. *International Journal of Computers and Applications*, *43*(5), 462–471. https://doi.org/10.1080/1206212X.2019.1578068

[5] Jozdani, S. E., Johnson, B. A., & Chen, D. (2019). Comparing deep neural networks, ensemble classifiers, and support vector machine algorithms for object-based urban land use/land cover classification. *Remote Sensing*, *11*(14), 1713. https://doi.org/10.3390/rs11141713

[6] Zhu, L., Suomalainen, J., Liu, J., Hyyppä, J., Kaartinen, H., & Haggren, H. (2018). A review: Remote sensing sensors. In R. B. Rustamov, S. Hasanova, & M. H. Z. Zeynalova (Eds.), *Multi-purposeful application of geospatial data* (pp. 19–42). Intech Open. https://doi.org/10.5772/intechopen.71049

[7] Guido, R., Ferrisi, S., Lofaro, D., & Conforti, D. (2024). An overview on the advancements of support vector machine models in healthcare applications: A review. *Information*, *15*(4), 235. https://doi.org/10.3390/info15040235

[8] He, Y., Lai, Y., Chen, B., Chen, Y., Xie, Z., Yu, X., & Luo, M. (2024). An Open-Pit mines land use classification method based on random forest using UAV: A case study of a ceramic clay mine. *Minerals*, *14*(12), 1282. https://doi.org/10.3390/min14121282

[9] Hamad, R. (2020). An assessment of artificial neural networks, support vector machines and decision trees for land cover classification using Sentinel-2A data. *Applied Ecology and Environmental Sciences*, *8*(6), 459–464. https://doi.org/10.12691/aees-8-6-18

[10] Uba, N. K. (2016). *Land use and land cover classification using deep learning techniques*. Master's Thesis, Arizona State University. https://hdl.handle.net/2286/R.I.38779

[11] Chen, X., Wu, S., & Wu, J. (2024). Characteristics and formation mechanism of land use conflicts in northern Anhui: A case study of Funan County. *Heliyon*, *10*(1), e22923. https://doi.org/10.1016/j.heliyon.2023.e22923

[12] Carranza-García, M., García-Gutiérrez, J., & Riquelme, J. C. (2019). A framework for evaluating land use and land cover classification using convolutional neural networks. *Remote Sensing*, *11*(3), 274. https://doi.org/10.3390/rs11030274

[13] Fayaz, M., Nam, J., Dang, L. M., Song, H. K., & Moon, H. (2024). Land-cover classification using deep learning with high-resolution remote-sensing imagery. *Applied Sciences*, *14*(5), 1844. https://doi.org/10.3390/app14051844

[14] Mou, L., & Zhu, X. X. (2018). A recurrent convolutional neural network for land cover change detection in multispectral images. In *IEEE International Geoscience and Remote Sensing Symposium*, 4363–4366. https://doi.org/10.1109/IGARSS.2018.8517375

[15] Sefrin, O., Riese, F. M., & Keller, S. (2021). Deep learning for land cover change detection. *Remote Sensing*, *13*(1), 78. https://doi.org/10.3390/rs13010078

[16] Temenos, A., Temenos, N., Kaselimi, M., Doulamis, A., & Doulamis, N. (2023). Interpretable deep learning framework for land use and land cover classification in remote sensing using SHAP. *IEEE Geoscience and Remote Sensing Letters*, *20*, 8500105. https://doi.org/10.1109/LGRS.2023.3251652

[17] Zhang, C., Wei, S., Ji, S., & Lu, M. (2019). Detecting large-scale urban land cover changes from very high resolution remote sensing images using CNN-based classification. *ISPRS International Journal of Geo-Information*, *8*(4), 189. https://doi.org/10.3390/ijgi8040189

[18] Zhang, C., Sargent, I., Pan, X., Li, H., Gardiner, A., Hare, J., & Atkinson, P. M. (2019). Joint deep learning for land cover and

land use classification. *Remote Sensing of Environment*, *221*, 173–187. https://doi.org/10.1016/j.rse.2018.11.014

[19] Vali, A., Comai, S., & Matteucci, M. (2020). Deep learning for land use and land cover classification based on hyperspectral and multispectral earth observation data: A review. *Remote Sensing*, *12*(15), 2495. https://doi.org/10.3390/rs12152495

[20] Sewada, R., & Goyal, H. (2023). Landuse landcover change detection using geospatial semantic segmentation U-Net-ResNet architecture. In *IEEE 11th Region 10 Humanitarian Technology Conference*, 620–625. https://doi.org/10.1109/R10-HTC57504.2023.10461857

[21] Maurya, A., Akashdeep, Mittal, P., & Kumar, R. (2023). A modified U-net-based architecture for segmentation of satellite images on a novel dataset. *Ecological Informatics*, *75*, 102078. https://doi.org/10.1016/j.ecoinf.2023.102078

[22] Ulmas, P., & Liiv, I. (2020). Segmentation of satellite imagery using U-net models for land cover classification. *arXiv Preprint:2003.02899*.

[23] Sewada, R., & Goyal, H. (2024). Analysis of geospatial deep convolutional semantic segmentation networks for landuse landcover feature mapping. In *AIP Conference Proceedings, 3156*, 030005. https://doi.org/10.1063/5.0228944

[24] Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. In *3rd International Conference on Learning Representations*.

[25] Li, H., & Wang, X. (2021). Robustness analysis for VGG-16 model in image classification of post-hurricane buildings. In *2nd International Conference on Big Data & Artificial Intelligence & Software Engineering*, 401–407. https://doi.org/10.1109/ICBASE53849.2021.00081

[26] Qian, M., Li, Y., Zhao, Y., & Yu, X. (2022). Prior knowledge-based deep convolutional neural networks for fine classification of land covers in surface mining landscapes. *Sustainability*, *14*(19), 12563. https://doi.org/10.3390/su141912563

[27] Ren, Y., Xie, Z., & Zhai, S. (2024). Urban land use classification model fusing multimodal deep features. *ISPRS International Journal of Geo-Information*, *13*(11), 378. https://doi.org/10.3390/ijgi13110378

[28] Mu, Y., Ni, R., Zhang, C., Gong, H., Hu, T., & Li, S. (2021). A lightweight model of VGG-16 for remote sensing image classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *14*, 6916–6922. https://doi.org/10.1109/JSTARS.2021.3090085

[29] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965

[30] Cecili, G., de Fioravante, P., Dichicco, P., Congedo, L., Marchetti, M., & Munafò, M. (2023). Land cover mapping with convolutional neural networks using Sentinel-2 images: Case study of Rome. *Land*, *12*(4), 879. https://doi.org/10.3390/land12040879

[31] Kolluru, V., Mungara, S., Chintakunta, A. N., Kolluru, L., & Telaganeni, C. S. (2024). Revolutionizing land cover analysis: A systematic review of geospatial intelligence with classification and segmentation. *International Journal of Artificial Intelligence and Applications*, *15*(5), 33–51. https://doi.org/10.5121/ijaia.2024.15503

[32] Neupane, B., Horanont, T., & Aryal, J. (2021). Deep learning-based semantic segmentation of urban features in satellite images: A review and meta-analysis. *Remote Sensing*, *13*(4), 808. https://doi.org/10.3390/rs13040808

[33] Hu, F., Xia, G. S., Hu, J., & Zhang, L. (2015). Transferring deep convolutional neural networks for the scene classification of high-resolution remote sensing imagery. *Remote Sensing*, *7*(11), 14680–14707. https://doi.org/10.3390/rs71114680

[34] Qi, S., He, M., Hoang, R., Zhou, Y., Namadi, P., Tom, B., ...,& Huynh, V. (2023). Salinity modeling using deep learning with data augmentation and transfer learning. *Water*, *15*(13), 2482. https://doi.org/10.3390/w15132482

[35] Rozo, A., Moeyersons, J., Morales, J., van der Westen, R. G., Lijnen, L., Smeets, C., ...,& Varon, C. (2022). Data augmentation and transfer learning for data quality assessment in respiratory monitoring. *Frontiers in Bioengineering and Biotechnology*, *10*, 806761. https://doi.org/10.3389/fbioe.2022.806761

[36] Alhichri, H., Alswayed, A. S., Bazi, Y., Ammour, N., & Alajlan, N. A. (2021). Classification of remote sensing images using EfficientNet-B3 CNN model with attention. *IEEE Access*, *9*, 14078–14094. https://doi.org/10.1109/ACCESS.2021.3051085

How to Cite: Sewada, R., & Goyal, H. (2025). A Novel VGG-16 Adaptation for Multi-band Satellite Image Classification: Optimized Preprocessing and Class-specific Augmentation. *Journal of Computational and Cognitive Engineering*. https://doi.org/10.47852/bonviewJCCE52024808