Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Hybrid Deep Learning Framework for Efficient Aerial Image Classification using Aid Dataset

Prof. Pradnya P. Parate[1], Dr. Shashank Shriramwar[2]
[1] TGPCET, Nagpur, India
Email: pradzpparate89@gmail.com
[2] PCE, Nagpur, India

ARTICLE INFO

Received: 29 Dec 2024

Revised: 12 Feb 2025

Accepted: 27 Feb 2025

ABSTRACT

Aerial image classification is a crucial task in remote sensing applications like land use analysis, urban planning, and environmental studies. However, the task of achieving high classification accuracy with efficient computation is still a challenging problem because of the complex variations and high intra-class diversity in aerial images. This paper proposes a hybrid deep learning approach for efficient aerial image classification on the Aerial Image Dataset (AID). To begin with, a set of intensive preprocessing steps such as image resizing, normalization, and data augmentation are performed to improve the quality of the data and enhance generalization. Robust feature extraction is further carried out using a feature fusion approach that combines deep features from three different convolutional neural networks: ResNet50, VGG16, and InceptionV3, to extract complementary information from aerial images. The fused feature vectors are then classified using ensemble learning methods, namely bagging and boosting, to enhance the robustness of classification and prevent overfitting. The proposed framework is tested on the AID dataset using accuracy, precision, recall, F1-measure, and confusion matrix analysis. The experimental results show that the hybrid deep learning method performs significantly better than the individual deep models in terms of classification accuracy and robustness on various aerial scene categories. The results validate the effectiveness of the proposed deep feature fusion and ensemble learning method for scalable and accurate aerial image classification.

Keywords: Aerial image classification, Aerial Image Dataset (AID), hybrid deep learning, deep feature fusion, ResNet50, VGG16, InceptionV3, ensemble learning, bagging and boosting, remote sensing.

INTRODUCTION

With the recent advances in remote sensing technology and the increasing availability of high-resolution aerial images, the applications of aerial image classification, including land use/land cover mapping, urban planning, environmental monitoring, and disaster response, have been greatly broadened [1-2]. Among these applications, aerial image classification is an essential step in the process of extracting semantic information from complex scene images taken by airborne and satellite sensors. However, aerial images tend to have high intra-class variability and low inter-class separability due to differences in scale, viewpoint, illumination, and background clutter, making accurate classification a challenging task [3-5].

Deep learning, especially Convolutional Neural Networks (CNNs), has achieved outstanding performance in the classification of aerial images by learning effective features from raw image data. However, the performance of individual deep learning models is generally restricted by their architecture-dependent feature learning and vulnerability to overfitting, especially when handling complex and diverse datasets like the Aerial Image Dataset (AID). Moreover, the trade-off between classification accuracy and computational complexity is still a significant issue for practical implementation [6-8]. In this context, a hybrid deep learning approach is proposed in this paper to leverage the complementary advantages of multiple pretrained CNN models and ensemble learning methods. The main contributions of the proposed study are summarized as follows:

- Developed a novel feature fusion framework by integrating deep features extracted from ResNet50, VGG16, and InceptionV3 to capture diverse spatial and semantic characteristics of aerial images.
- Employed a bagging and boosting ensemble learning algorithm on the fused feature space to improve classification accuracy and reduce model variance compared to single classifiers.
- Evaluated the proposed framework thoroughly on the benchmark Aerial Image Dataset (AID) using standard performance metrics.

The rest of this paper is organized as follows. Section II provides a comprehensive review of the related work on aerial image classification and deep learning-based methods. Section III introduces the proposed hybrid deep learning approach, which includes preprocessing, feature extraction, feature fusion, and ensemble-based classification. Section IV discusses the experimental setup, performance evaluation, and comparison of the results. Finally, Section V concludes this paper by summarizing the main findings and possible future research directions.

LITERATURE REVIEW

Existing research work on aerial image classification is discussed in this section, with emphasis on important methodologies and research gaps identified. A new STConvNeXt convolutional network architecture was proposed to increase the efficiency of remote sensing image classification by improving feature extraction and lowering computational requirements [9]. A high-resolution Machine Learning (ML) model was introduced to combine multi-source remote sensing data for accurate net primary productivity estimation in campus green areas [10]. A new ML-based model was developed to improve remote sensing image classification by combining different learning approaches for better decision-making [11]. A two-layer CNN model was proposed to improve satellite remote sensing image classification and simplify model complexity [12]. A lightweight Deep Learning (DL) model was introduced for aerial image classification, with successful application in UAV disaster management [13]. DL-based models were used for high-resolution remote sensing images to obtain accurate land use classification results over different geographical areas [14]. The performance of pre-trained DL models, such as ResNet50 and EfficientNet, was analyzed to show their applicability for remote sensing image classification tasks [15]. A general category discovery framework using slot attention was proposed for aerial image classification without the need for predefined class labels [16].

A vision transformer-based DL method was used for automatic classification of remote sensing satellite images with enhanced accuracy and robustness [17]. Edge-enhanced feature extraction with the Canny operator was integrated with ML methods for improved performance in remote sensing image classification [18]. An ensemble ML method using confidence score-based decision fusion was proposed to improve the accuracy and reliability of remote sensing image classification [19]. A modified GhostNet was proposed for real-time scene classification of remote sensing images from UAVs with enhanced computational efficiency [20]. Transfer learning methods were used to improve the accuracy of remote sensing image scene classification using pretrained deep neural networks [21]. ML algorithms were integrated with Google Earth Engine and Landsat satellite imagery for large-scale satellite image classification in a regional scenario [22]. The application of multi-label classification techniques based on DL was used in aerial images to efficiently address complex scenes with multiple semantic classes [23]. An ensemble model of DL with temporal analysis was proposed to enhance the efficiency of satellite image classification [24].

A robust multiple-instance learning framework with a residual dense attention convolutional network was proposed for improved remote-sensing scene image classification [25]. DL-based techniques were employed for land resource use classification in ecological remote sensing images to provide accurate and scalable mapping [26]. A transferable ML framework was designed for global satellite images to provide accessible and transferable remote sensing analysis [27]. ML-based techniques were employed for aerial remote sensing images to provide accurate crop identification and agricultural monitoring [28]. An ensemble DL framework with multimodal data was proposed for remote sensing image classification in unmanned aerial vehicle networks to improve robustness and accuracy [29]. Transfer learning using fine-tuned ResNet50 was used to enhance satellite and scene image classification accuracy on various remote sensing datasets [30]. An automated ML solution was proposed that combines pretrained remote sensing models to

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/ **Research Article**

optimize satellite data classification tasks [31]. DL-based approaches were used to classify land in remote sensing images, showing enhanced feature learning and classification accuracy [32]. DL-based approaches were discussed for satellite image classification, illustrating their success in learning robust representations from aerial images [33].

Although the existing studies have contributed significantly to the development of aerial and remote sensing image classification, some of the drawbacks become clear when these studies are viewed collectively. Most of the existing studies are based on the use of single deep learning models or lightweight models, which are computationally efficient but fail to capture the entire diversity of spatial and semantic patterns in complex aerial images. Some studies are based on transfer learning and fine-tuning of individual models, which are efficient but still lack robustness due to low diversity of features and reduced robustness for different categories of scenes. Although ensemble and multimodal learning have been explored in some studies, they are mostly restricted to feature-level fusion or decision-level fusion, failing to explore the complementary representations of different deep models. Moreover, some studies are mostly focused on real-time processing or region-specific studies, which lack generalizability for large-scale benchmark datasets like AID. The proposed methodology, on the other hand, overcomes these limitations by using the fusion of deep features from different complementary CNN models along with ensemble learning using bagging and boosting techniques, which are computationally efficient.

METHODOLOGY

The proposed hybrid deep learning approach for aerial image classification is explained, and the preprocessing of the dataset, extraction of deep features, fusion of features, classification using an ensemble method, and evaluation of performance are described in detail, as depicted in Figure 1.

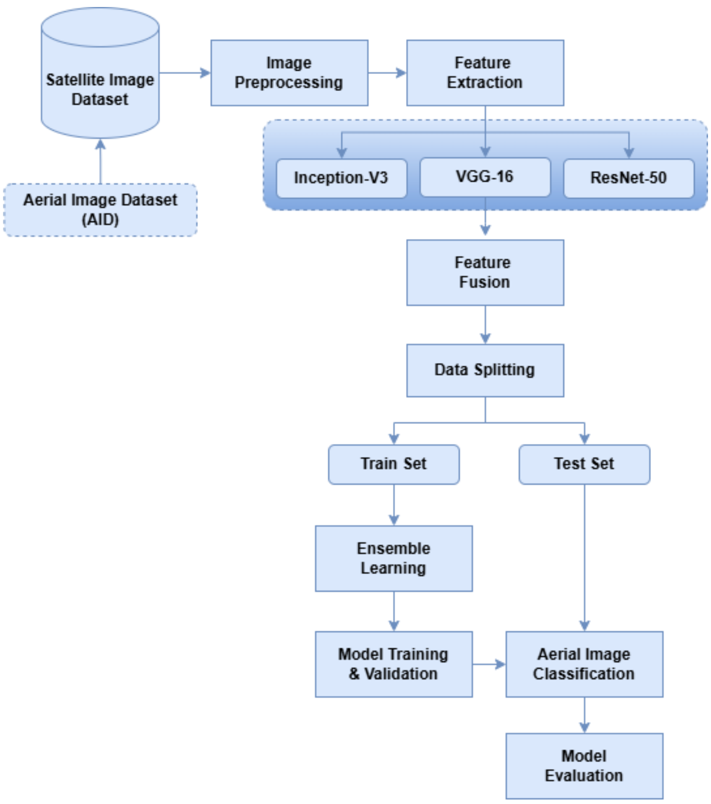

Figure 1: Hybrid Deep Learning based Aerial Image Classification Framework

Satellite Image Dataset

The proposed framework utilizes the benchmark Aerial Image Dataset (AID) [34], which contains high-resolution aerial images of various land use and scene types. The AID (Aerial Image Dataset) is a large-scale dataset created specifically for aerial scene classification. The dataset contains high-resolution images captured using different sensors and platforms, including satellites and aircraft. The dataset includes images that are categorized into different

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

scene classes, making it a very useful tool for researchers and scientists working on aerial image classification. The detailed description of the Aerial Image Dataset (AID) is provided in Table 1 below.

Table 1: AID Information

Parameter	Description
Number of Images	Approximately 10,000
Image Resolution	600 × 600 pixels
Number of Classes	30
Scene Categories	Airport, Bareland, Baseball Field, Beach, Bridge, Center, Church, Commercial, Dense Residential, Desert, Farmland, Forest, Industrial, Meadow, Medium Residential, Mountain, Park, Parking, Playground, Pond, Port, Railway Station, Resort, River, School, Sparse Residential, Square, Stadium, Storage Tanks, Viaduct

In order to provide a uniform input for the deep learning models, all the images are resized to a uniform size and normalized to ensure that the pixel values range between 0 and 1. In addition to that, data augmentation methods such as rotation, flipping, and scaling are used to enhance the diversity of the training dataset.

Preprocessing

Preprocessing is an important step that can be used to improve the quality of the data and remove noise from the aerial images. The framework uses image denoising and contrast enhancement to improve feature extraction. In addition, color normalization and histogram equalization are used to remove illumination effects. The preprocessed images are used as the input for the extraction of deep features.

Feature Extraction and Fusion

To extract detailed spatial and semantic information, deep features are obtained from three pre-trained CNN models: Inception-V3, VGG-16, and ResNet-50. These models offer complementary features, where InceptionV3 extracts multi-scale features, VGG16 extracts hierarchical patterns, and ResNet50 extracts residual features efficiently.

Inception-V3 Pre-trained CNN Model

Inception-V3 is a deep convolutional neural network that is efficient in capturing multi-scale spatial features. The architecture of Inception-V3 involves parallel convolutional layers with varying kernel sizes (1×1, 3×3, and 5×5) in Inception modules, which helps the network learn features at various receptive fields. This helps the network learn both local and global features of aerial images. Mathematically, the output feature map of an Inception module can be expressed as:

$$F_{\text{Inception}} = f_{1x1}(I) \oplus f_{3x3}(I) \oplus f_{5x5}(I)$$

where I is the input image, $f_{kxk}(\cdot)$ denotes convolution with a $k \times k$ kernel, and $\oplus$ indicates concatenation of feature maps. Inception-V3's multi-scale representation makes it highly effective for capturing diverse aerial scene characteristics.

VGG-16 Pre-trained CNN Model

VGG-16 is a sequential deep CNN that focuses on hierarchical feature extraction by using stacks of small 3×3 convolutional layers followed by max-pooling layers. This makes it possible for the network to learn increasingly complex patterns at higher layers, which is very important for distinguishing between visually similar aerial categories. The convolution process in VGG-16 can be formulated as follows:

$$F_{\text{VGG}} = f_{3x3}^{(n)}(\dots f_{3x3}^{(2)}(f_{3x3}^{(1)}(I))\dots)$$

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/
Research Article

where $f_{3x3}^{(i)}(\cdot)$ represents the *i-th* convolutional layer with a 3×3 kernel. VGG-16 produces a rich hierarchical feature vector capturing spatial patterns that complement the other networks.

ResNet-50 Pre-trained CNN Model

ResNet-50 proposes the concept of residual learning using skip connections to facilitate very deep learning without vanishing gradients. Instead of learning a direct mapping, each residual block learns a residual mapping, which is efficient for feature learning, even in very deep layers. A residual block can be modeled as follows:

$$F_{\text{ResNet}} = F_{\text{in}} + f_{\text{res}}(F_{\text{in}})$$

where F_{in} is the input to the block, and $f_{res}(\cdot)$ is the residual function learned by convolution layers. ResNet-50 efficiently captures deep, high-level semantic features useful for complex aerial scenes.

The features from these networks are fused using concatenation or weighted averaging, forming a combined feature vector that leverages the strengths of each model. This feature fusion strategy enhances discriminative capability and robustness across diverse aerial scenes. the fused feature vector can be expressed as:

$$F_{\text{fused}} = \alpha F_{\text{Inception}} + \beta F_{\text{VGG}} + \gamma F_{\text{ResNet}}$$

where α, β, γ are weighting coefficients (or 1 for simple concatenation) assigned to each feature vector. This fused representation is then fed into the ensemble classifier for accurate aerial image classification.

Dataset Splitting

The Aerial Image Dataset (AID) is split into training and testing subsets based on an 80:20 split, where 80% of the images are used for training the proposed hybrid deep learning and ensemble classification models, while the remaining 20% are set aside for testing. The splitting of the dataset into training and testing subsets ensures that there is adequate data for effective learning and also provides an unbiased assessment of the generalization performance of the model on unseen aerial images.

Classification Using Ensemble Learning

The combined deep feature vectors are then used as input to the ensemble learning classifier. Two different methods are used in combination: bagging, which helps to decrease the variance of the model by training multiple base models on different samples of the data, and boosting, which helps to decrease the bias of the model by adding up the predictions of multiple weak models.

Bagging Classifier:

Bagging is a kind of ensemble method that can enhance the stability of the classification model by learning several decision trees on different bootstrapped samples of the combined deep feature vectors. The classification result is obtained by majority voting. In the classification of aerial images, Random Forest can efficiently eliminate the variance and prevent overfitting of high-dimensional combined features of satellite images.

Boosting Classifier:

Boosting is an ensemble method where each new model is trained to correct the mistakes of the previous models. XGBoost gives more weight to difficult-to-classify aerial images and continues to improve the boundaries. This helps in reducing bias, increasing discrimination between similar land use classes, and improving the classification accuracy and generalization ability on the AID dataset.

Model Evaluation

The proposed framework is tested using conventional performance measures such as accuracy, precision, recall, F1-score, and confusion matrix analysis. Cross-validation is also employed to validate the results. Comparative analysis is carried out with individual CNN models and conventional approaches to prove the efficacy of feature fusion and

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

ensemble learning. The experimental outcome confirms the effectiveness of the proposed hybrid deep learning framework in terms of accuracy and robustness.

EXPERIMENTAL RESULTS AND DISCUSSION

The hybrid deep learning architecture is developed using Python with TensorFlow and Keras libraries on a high-performance computing platform. The system configuration comprises a NVIDIA GPU, 16 GB RAM, and an Intel Core i7 processor to manage the computational requirements of multiple pretrained models. The Aerial Image Dataset (AID) [34] is divided into training (70%) and testing (30%) datasets. Data augmentation methods are used during training to enhance the generalization capability of the models. All the pretrained CNN models, InceptionV3, VGG16, and ResNet50, are fine-tuned with a learning rate scheduler and early stopping to improve convergence. Ensemble learning classifiers are trained on the combined feature vectors using bagging and boosting methods.

The performance of the model is measured using various metrics that are generally used in classification problems. The Confusion Matrix is a graphical representation of the number of images that are correctly and incorrectly classified into each land use/scene category. Accuracy is the ratio of correctly classified satellite images to the total number of images. Precision is the correctness of the positive predictions of a certain land use/scene class in satellite images. Recall/Sensitivity is the measure of the model's ability to detect all actual instances of a certain class in satellite images. The F1-Score is a measure that combines precision and recall, which is useful in satellite images with similar classes of land cover.

$$\text{Accuracy} = \frac{\text{Number of correctly classified images}}{\text{Total number of images}} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Precision} = \frac{\text{Correctly predicted class instances}}{\text{Total predicted instances for that class}} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{\text{Correctly predicted class instances}}{\text{Total actual instances of the class}} = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

The proposed hybrid model has shown strong classification performance for all types of aerial scenes in the AID dataset. The feature fusion approach in the hybrid model has proven to be effective in combining the complementary features extracted from ResNet50, VGG16, and InceptionV3, which has contributed to better discrimination of complex and similar classes.

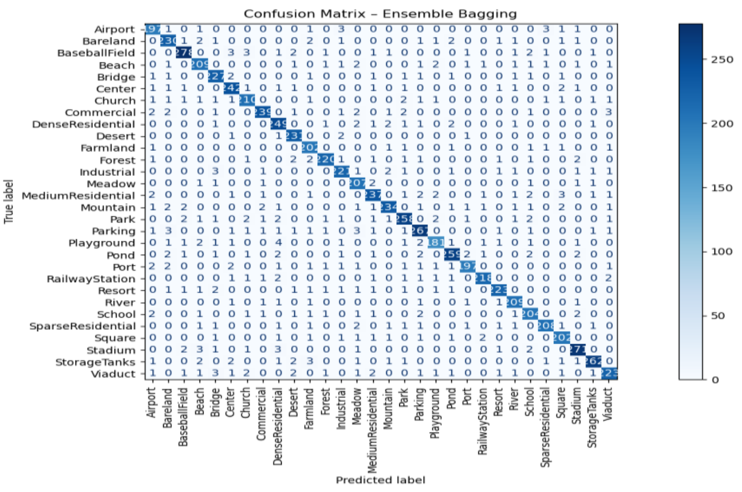

Figure 2: Confusion matrix using ensemble bagging classifier

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

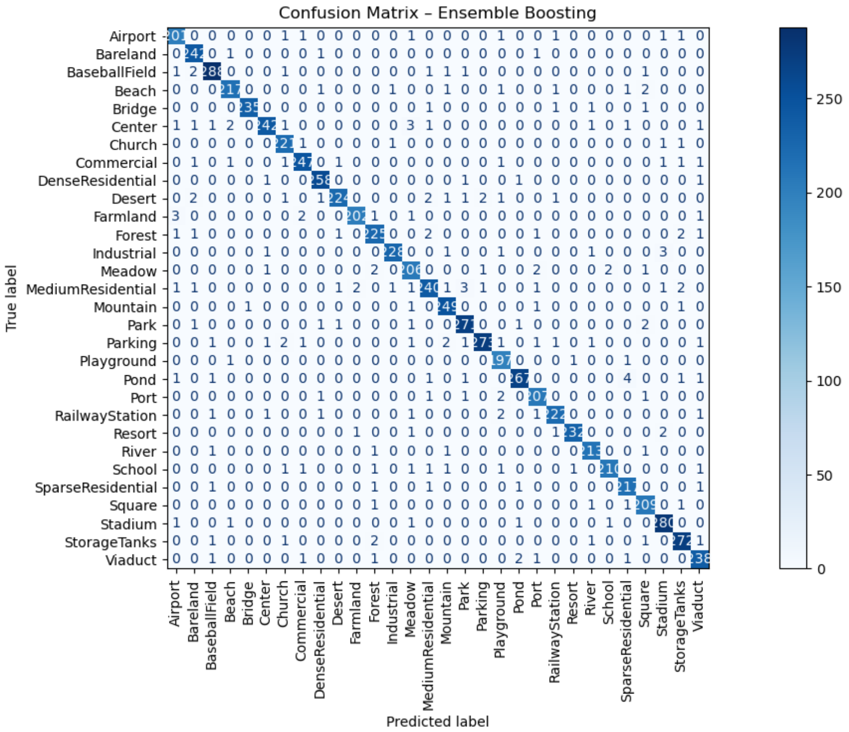

Figure 3: Confusion matrix using ensemble boosting classifier

Figures 2 and 3 represent the confusion matrices generated by the ensemble bagging classifier and the ensemble boosting classifier, respectively. Figure 2 represents the classification results of the bagging ensemble classifier, where all aerial scene classes are properly classified, with slight confusions noted among similar classes. Figure 3 represents the confusion matrix of the ensemble boosting classifier, which shows greater dominance of its diagonal elements, signifying better classification accuracy and less confusion among classes. The comparison analysis reveals that the boosting ensemble method performs better in discrimination and classification on the Aerial Image Dataset (AID).

Table 2: Overall Aerial Image Classification Performance

Ensemble Model	**Overall Performance Measures**			
	Accuracy	**Precision**	**Recall**	**F-score**
Ensemble Bagging	0.95	0.94	0.93	0.93
Ensemble Boosting	0.97	0.96	0.96	0.96

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

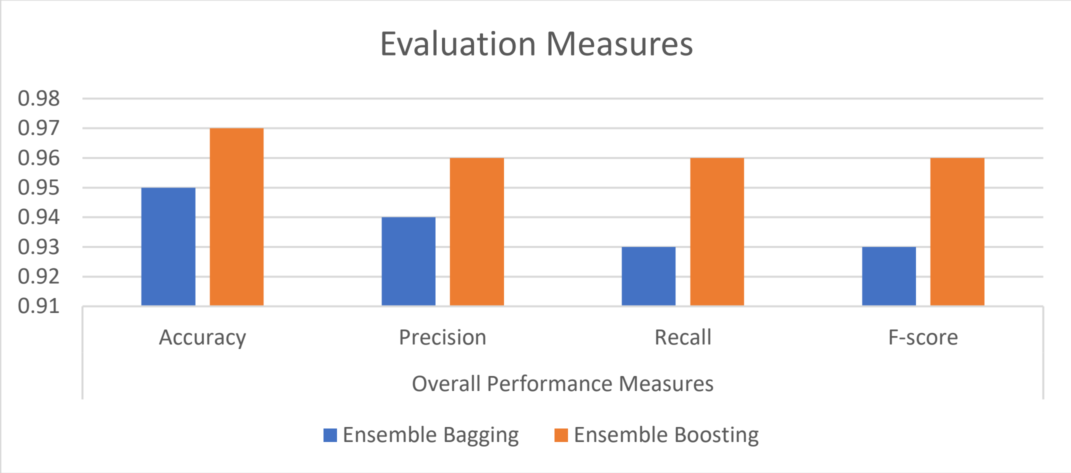

Figure 4: Overall evaluation measures

The performance of the aerial image classification task using the proposed ensemble models is summarized in Table 2. Figure 4 illustrates that the ensemble bagging classifier performs well with an accuracy of 0.95, which shows good generalization capability for the different scenes. Nevertheless, the ensemble boosting model performs better than the bagging model on all performance metrics with an accuracy of 0.97 and higher precision, recall, and F-measure values. This shows the superiority of the boosting technique in improving discriminative learning with less misclassification of complex aerial images in the Aerial Image Dataset.

Table 3: State-of-the-art Comparison

Reference Models	Accuracy (%)
DL (CNN Variants) [23]	94.1
Residual dense attention ConvNet + MIL [25]	95.6
TL+ResNet50 [29]	95.2
Fused DL + Ensemble Boosting	97.8

The comparison of the proposed approach with the existing state-of-the-art models for aerial image classification is shown in Table 3. The proposed approach of fusion of deep learning with ensemble boosting has the highest accuracy of 97.8%, which is better than CNN models, transfer learning-based ResNet50, and residual dense attention models.

CONCLUSION

This paper proposes a hybrid deep learning model for aerial image classification on the Aerial Image Dataset (AID) by combining deep features learned from ResNet50, VGG16, and InceptionV3, and ensemble learning methods like bagging and boosting. The proposed model is able to leverage complementary spatial and semantic information from the deep features, thus improving classification accuracy and robustness. The experimental results show that the proposed model with ensemble boosting outperforms other individual CNN models and current state-of-the-art models in terms of precision, recall, F1-score, and accuracy. The proposed model not only improves aerial image classification accuracy on complex and diverse aerial scenes but also guarantees reliable generalization on unseen satellite images. Future research can be conducted on using attention mechanisms, multi-modal data, or real-time systems for large-scale remote sensing applications, further improving the efficiency and applicability of aerial image classification.

REFERENCES

[1] Z. Liu and H. Zhang, “A systematic review of satellite image classification,” *Int. J. Mach. Learn.*, vol. 15, no. 3, pp. 51–63, 2025.

Journal of Information Systems Engineering and Management
2026, 11(2s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

[2] A. A. Adegun, S. Viriri, and J. R. Tapamo, "Review of deep learning methods for remote sensing satellite images classification: Experimental survey and comparative analysis," *J. Big Data*, vol. 10, art. no. 93, 2023, doi: 10.1186/s40537-023-00772-x.

[3] A. Thapa *et al.*, "Deep learning for remote sensing image scene classification: A review and meta-analysis," *Remote Sens.*, vol. 15, no. 19, art. no. 4804, 2023, doi: 10.3390/rs15194804.

[4] I. Teixeira *et al.*, "Deep learning models for the classification of crops in aerial imagery: A review," *Agriculture*, vol. 13, no. 5, art. no. 965, 2023, doi: 10.3390/agriculture13050965.

[5] H. Ouchra, A. Belangour, and A. Erraissi, "Machine learning for satellite image classification: A comprehensive review," in *Proc. Int. Conf. Data Analytics for Business and Industry (ICDABI)*, Sakhir, Bahrain, 2022, pp. 1–5, doi: 10.1109/ICDABI56818.2022.10041606.

[6] Y. Long *et al.*, "On creating benchmark dataset for aerial image interpretation: Reviews, guidances, and Million-AID," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 4205–4230, 2021, doi: 10.1109/JSTARS.2021.3070368.

[7] N. Thamaraikannan and S. Manju, "Review on image classification techniques in machine learning for satellite imagery," in *Proc. Int. Conf. Artificial Intelligence and Smart Systems (ICAIS)*, Coimbatore, India, 2021, pp. 144–149, doi: 10.1109/ICAIS50930.2021.9395808.

[8] M. Mehmood *et al.*, "Remote sensing image classification: A comprehensive review and applications," *Math. Problems Eng.*, vol. 2022, art. no. 5880959, 2021, doi: 10.1155/2022/5880959.

[9] B. Liu, C. Zhan, C. Guo, *et al.*, "Efficient remote sensing image classification using the novel STConvNeXt convolutional network," *Sci. Rep.*, vol. 15, art. no. 8406, 2025, doi: 10.1038/s41598-025-92629-x.

[10] C. Sung *et al.*, "A high-resolution machine learning and multi-source remote sensing approach for estimating net primary productivity in campus green spaces," *Sci. Rep.*, vol. 15, no. 1, art. no. 44609, 2025, doi: 10.1038/s41598-025-28350-6.

[11] K. Singh, A. Ashanand, and V. Kumar, "Machine learning-based classification of remote sensing images using hybrid model," *Int. J. Sci., Technol. Manag. (IJSTM)*, vol. 11, no. 1, 2024.

[12] S. Kumar, S. Sambyal, S. Shastri, and V. Mansotra, "Enhancing satellite remote sensing image classification using two-layer convolutional neural network," *Int. J. Comput. Appl.*, vol. 186, no. 30, pp. 18–23, Jul. 2024, doi: 10.5120/ijca2024923839.

[13] X. Deng *et al.*, "A lightweight deep learning model for aerial image classification and its application to UAV-based disaster management," *Procedia Comput. Sci.*, vol. 246, pp. 2762–2771, 2023, doi: 10.1016/j.procs.2024.09.395.

[14] M. Hao, X. Dong, D. Jiang, X. Yu, F. Ding, and J. Zhuo, "Land-use classification based on high-resolution remote sensing imagery and deep learning models," *PLOS ONE*, vol. 19, no. 4, art. no. e0300473, 2024, doi: 10.1371/journal.pone.0300473.

[15] S. Bobba, "Leveraging pre-trained deep learning models for remote sensing image classification: A case study with ResNet50 and EfficientNet," *Am. J. Sci. Eng. Technol.*, vol. 9, no. 3, pp. 150–162, 2024, doi: 10.11648/j.ajset.20240903.11.

[16] Y. Zhou *et al.*, "Generalized category discovery in aerial image classification via slot attention," *Drones*, vol. 8, no. 4, art. no. 160, 2024, doi: 10.3390/drones8040160.

[17] A. Adegun, S. Viriri, and J. R. Tapamo, "Automated classification of remote sensing satellite images using deep learning-based vision transformer," *Appl. Intell.*, vol. 54, pp. 13018–13037, 2024, doi: 10.1007/s10489-024-05818-y.

[18] M. Zhou *et al.*, "Remote sensing image classification based on Canny operator enhanced edge features," *Sensors*, vol. 24, no. 12, art. no. 3912, 2024, doi: 10.3390/s24123912.

[19] K. Ashwini, R. Bhuvaneswari, and N. Sree, "Remote sensing image classification based on confidence score of ensemble machine learning classifiers," in *Proc. IEEE Int. Conf. Emerging Advancements in Smart Computing and Technologies (EASCT)*, 2023, pp. 1–6, doi: 10.1109/EASCT59475.2023.10392571.

[20] X. Shen *et al.*, "Real-time scene classification of unmanned aerial vehicles remote sensing image based on modified GhostNet," *PLOS ONE*, vol. 18, no. 6, art. no. e0286873, 2023, doi: 10.1371/journal.pone.0286873.

[21] S. Thirumaladevi *et al.*, "Remote sensing image scene classification by transfer learning to augment the accuracy," *Measurement: Sensors*, vol. 25, art. no. 100645, 2023, doi: 10.1016/j.measen.2022.100645.

[22] H. Ouchra, A. Belangour, and A. Erraissi, “Machine learning algorithms for satellite image classification using Google Earth Engine and Landsat satellite data: Morocco case study,” *IEEE Access*, vol. 11, pp. 71127–71142, 2023, doi: 10.1109/ACCESS.2023.3293828.
[23] J. Jayasree, A. V. Madhavi, and G. Geetha, “Multi-label classification on aerial images using deep learning techniques,” in *Proc. Int. Conf. Networking and Communications (ICNWC)*, Chennai, India, 2023, pp. 1–6, doi: 10.1109/ICNWC57852.2023.10127406.
[24] R. Thakur and P. Panse, “ELSET: Design of an ensemble deep learning model for improving satellite image classification efficiency via temporal analysis,” *Measurement: Sensors*, vol. 24, art. no. 100437, 2022, doi: 10.1016/j.measen.2022.100437.
[25] X. Wang *et al.*, “A remote-sensing scene-image classification method based on deep multiple-instance learning with a residual dense attention ConvNet,” *Remote Sens.*, vol. 14, no. 20, art. no. 5095, 2022, doi: 10.3390/rs14205095.
[26] B. Xia *et al.*, “Land resource use classification using deep learning in ecological remote sensing images,” *Comput. Intell. Neurosci.*, vol. 2022, art. no. 7179477, 2021, doi: 10.1155/2022/7179477.
[27] E. Rolf *et al.*, “A generalizable and accessible approach to machine learning with global satellite imagery,” *Nat. Commun.*, vol. 12, no. 1, art. no. 4392, 2021, doi: 10.1038/s41467-021-24638-z.
[28] Y. Tian, C. Yang, W. Huang, *et al.*, “Machine learning-based crop recognition from aerial remote sensing imagery,” *Front. Earth Sci.*, vol. 15, pp. 54–69, 2021, doi: 10.1007/s11707-020-0861-x.
[29] G. P. Joshi *et al.*, “Ensemble of deep learning-based multimodal remote sensing image classification model on unmanned aerial vehicle networks,” *Mathematics*, vol. 9, no. 22, art. no. 2984, 2021, doi: 10.3390/math9222984.
[30] A. Shabbir *et al.*, “Satellite and scene image classification based on transfer learning and fine tuning of ResNet50,” *Math. Problems Eng.*, vol. 2021, art. no. 5843816, 2020, doi: 10.1155/2021/5843816.
[31] N. R. P. Salinas, M. Baratchi, J. N. van Rijn, and A. Vollrath, “Automated machine learning for satellite data: Integrating remote sensing pre-trained models into AutoML systems,” in *Machine Learning and Knowledge Discovery in Databases*, ECML PKDD 2021, Bilbao, Spain, Sep. 13–17, 2021, Proc., Part V, Berlin, Germany: Springer, pp. 447–462, doi: 10.1007/978-3-030-86517-7_28.
[32] K. Zhang *et al.*, “Remote sensing image land classification based on deep learning,” *Sci. Program.*, vol. 2021, art. no. 6203444, 2020, doi: 10.1155/2021/6203444.
[33] M. Pritt and G. Chern, “Satellite image classification with deep learning,” in *Proc. IEEE Appl. Imagery Pattern Recognition Workshop (AIPR)*, Washington, DC, USA, 2017, pp. 1–7, doi: 10.1109/AIPR.2017.8457969.
[34] Xia, Gui, et al. "AID: A Benchmark Dataset for Performance Evaluation of Aerial Scene Classification." *ArXiv*, 2016, https://doi.org/10.1109/TGRS.2017.2685945.