J Inno Sci Eng, 2025, 9(1):1-14

DOI:https://doi.org/ 10.38088/jise.1527548

JISE

Research Article

Aircraft Recognition Based on CNN Using Satellite Images

Meriç GENÇ[1] ,Yıldıray YALMAN[2*]

[1] *Department of Electrical and Electronics Engineering, Piri Reis University, 34940, İstanbul, Türkiye*

[2] *Department of Computer Engineering, Piri Reis University, 34940, Istanbul, Türkiye*

ARTICLE INFO

Received Date: 3/08/2024
Accepted Date: 24/11/2024

Cite this paper as:
Genc M. and Yalman Y. (2025). *Aircraft Recognition Based on CNN Using Satellite Images.*Journal of Innovative Science and Engineering. 9(1): 1-14

*Corresponding author: Yıldıray Yalman
E-mail:yyalman@pirireis.edu.tr

Keywords:
VGG16
VGG19
Convolutional Neural Networks
Aircraft Recognition
Visual Recognition

ABSTRACT

This study investigates the use of Convolutional Neural Networks (CNN), particularly the VGG16 and VGG19 architectures, for aircraft recognition with satellite-derived image data. Deep learning, especially multi-layer neural networks, addresses significant limitations in artificial intelligence, allowing advanced models to achieve high accuracy in complex tasks like aircraft recognition. The MTARSI dataset is exclusively used to evaluate these models. Motivated by the importance of accurate aircraft recognition in civil aviation, military security, and emergency interventions, this study aims to develop a CNN-based aircraft recognition system. Experimental results show that VGG19 outperforms VGG16, achieving an accuracy of 89.29% compared to 82.67% for VGG16. These findings highlight VGG19's advantage over traditional methods and underscore its potential in future military aircraft recognition systems.

1. Introduction

In today's rapidly digitizing world, technological tools that were once exclusive to defense systems have found effective applications in military aircraft recognition, particularly using satellite imagery and deep learning algorithms. The advancement of technology, coupled with the efficient use of deep learning techniques, has enabled the use of images and videos for a range of tasks, such as computer vision studies, which address crucial issues including national security, traffic management, and agricultural health monitoring.

The development of artificial intelligence, especially in the domain of computer vision, has significantly increased the demand for more efficient and effective vision systems. As artificial intelligence evolves, it has led to the creation of more sophisticated tools for a variety of tasks, such as object recognition, behavioral analysis, and scene understanding. Research in AI and deep learning, particularly focusing on applying these technologies for the benefit of humanity and simplifying daily life, has been particularly successful in high-demand tasks such as surveillance and object detection, which require intense focus. These innovations have not only ensured more efficient use of human resources but also minimized human errors and negligence in critical tasks. In recent years, researchers have focused on enhancing the capabilities of deep learning algorithms. Notably, deep learning architectures such as VGG16 and VGG19 have demonstrated impressive success in image recognition and classification tasks. These models, with their deep architectures comprising multiple layers, are highly effective in learning complex image features and improving performance in challenging classification problems.

This study aims to evaluate the performance of VGG16 and VGG19 architectures for military aircraft classification using the MTARSI dataset. The MTARSI dataset contains high-resolution images of military aircraft taken from various angles, making it an ideal resource for identifying different aircraft types and models. In this research, object recognition and classification tasks were performed by using VGG16 and VGG19 models, starting from pre-trained weights and fine-tuned transfer learning techniques. Preprocessing steps, including resizing and normalization, were applied to the dataset images before training. On the other hand, the aim of this study is to investigate the effectiveness of the VGG16 and VGG19 architectures in military aircraft classification using satellite imagery. By analyzing the performance of these models, we aim to demonstrate the potential of deep learning methods in improving the accuracy and efficiency of military aircraft classification systems.

The application of artificial intelligence (AI) and deep learning algorithms in the defense industry has led to significant advancements, particularly in areas such as military aircraft recognition using satellite imagery. In recent years, the focus has shifted towards using satellite images and remote sensing data to improve aircraft classification systems. Traditional image recognition methods, such as SIFT (Scale-Invariant Feature Transform) combined with Bag of Words (BOW), and HOG (Histogram of Oriented Gradients) with Support Vector Machines (SVM), have been widely used for this purpose. While these methods have provided some success, they are often computationally expensive and are sensitive to changes in image resolution and viewing angles. These techniques also struggle with distinguishing between similar-looking objects, which makes the task of recognizing military aircraft particularly challenging.

With the rapid development of deep learning methods, particularly Convolutional Neural Networks (CNNs), new solutions have emerged to improve classification tasks. CNN architectures like VGG16 and VGG19, known for their ability to learn complex image features, have shown remarkable performance in various computer vision tasks, including object classification, scene recognition, and more. These deep learning models can process large volumes of data and automatically extract important features from images without requiring hand-crafted feature engineering.

Recent studies have demonstrated the effectiveness of CNNs, particularly VGG architectures, in the classification of military aircraft from satellite imagery. For instance, Chen et al. [1] proposed a Region Locating Network (RLN) to improve the Faster R-CNN framework. This RLN strategy is designed to identify regions, such as runway and apron areas, where aircraft are typically located. They used a dataset of 265 images from 12 different airports, totaling over 6,000 annotated aircraft, to train the model. Their approach achieved a detection accuracy of 53.64%. To enhance the training dataset, they applied data augmentation techniques, including flipping the images horizontally and vertically.

Luo and Shi [2] addressed the problem of efficient aircraft localization in remote sensing images using a simple yet effective object proposal method. Their approach involved generating a small set of bounding boxes likely to contain the objects of interest (in this case, aircraft) and applying HOG features with an SVM classifier for detection. The model achieved an 80% detection accuracy on test data from 20 airport images collected from Google Earth, with image sizes ranging from 1000 × 1000 to 2000 × 2000.

Another related study, by Zhang et al. [3], explored the integration of traditional computer vision methods with deep learning to improve the accuracy of aircraft recognition. The researchers combined SIFT features with a CNN classifier, resulting in a notable improvement in the accuracy of aircraft detection in satellite images. They also introduced a novel data augmentation method that generated realistic synthetic images to increase the variety of the training dataset, further boosting the model’s performance.

These advancements in deep learning and computer vision have significantly contributed to the improvement of military aircraft recognition systems, enhancing both the accuracy and efficiency of such tasks. By utilizing pre-trained deep learning models and transfer learning, these systems can be further fine-tuned to achieve better results, especially in highly specialized fields like defense and national security.

In conclusion, the integration of artificial intelligence and deep learning algorithms has revolutionized the defense industry and various other industrial applications. Projects that explore and implement such technologies are paving the way for future advancements, contributing to a safer and more efficient world. The success of models like VGG16 and VGG19 in military aircraft classification demonstrates the potential of deep learning methods in enhancing national security and defense systems, reducing human error, and improving operational efficiency.

2. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNN) are a type of Multi-Layer Perceptron (MLP), characterized by multi-layered detectors. Cells in the visual center are divided into subregions to encompass the entire image, with simple cells focusing on edge-like features, and complex cells having broader receptors concentrating on the entire visual field. The CNN algorithm, an advanced neural network, is inspired by the visual center of animals. The mathematical convolution operation here can be thought of as a neuron's response to stimuli within its receptive field [3,4,5]. CNN consists of one or more convolutional layers, a subsampling layer, and one or more fully connected layers, similar to a standard multi-layer neural network, following that [6]. The LeNet architecture, introduced by Yann LeCun in 1988 and illustrated in Figure 1, underwent continuous improvements until the late 1990s [7]. Within the LeNet network, the lower layers consist of consecutively placed convolution and maximum pooling layers, while the subsequent upper layers align with the structure of fully connected traditional Multi-Layer Perceptron (MLP). Convolutional Neural Network (CNN) algorithms have diverse applications across various domains, such as image and audio processing, natural language processing (NLP), and biomedical research. Notably, CNN has excelled in the field of image processing, achieving state-of-the-art results. In a study focusing on the MNIST dataset, Cireşan has significantly reduced the error rate by using CNN [8]. Another investigation by Cireşan and colleagues, involving the MNIST and NORB datasets, showcased the rapid learning capabilities of CNN, surpassing conventional methods [9].

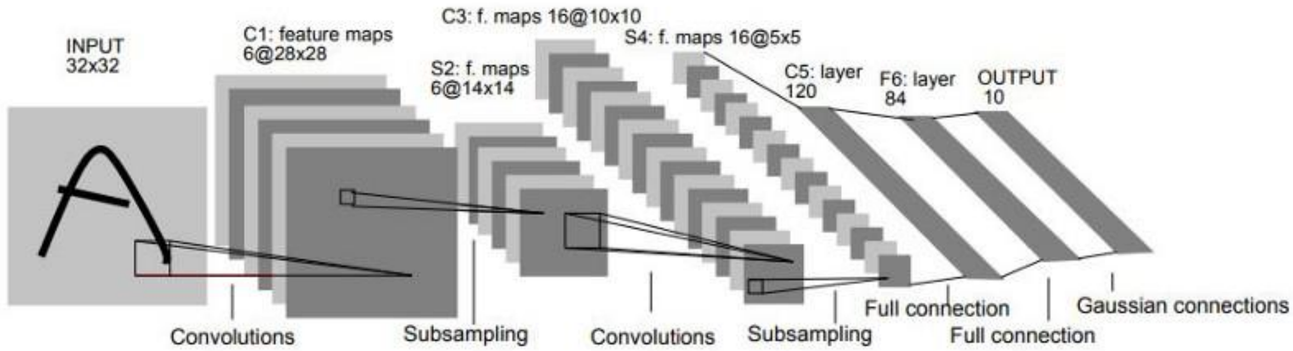


Figure 1: Demographics of the study participants.

In 2014, top-ranking teams in the ImageNet Competition, dealing with millions of images and hundreds of object classes in object classification and detection, employed modified versions of CNN algorithms. A subsequent study in 2015 highlighted CNN's ability to capture faces at wide angles, including upside-down faces. The network underwent training on a database comprising 200,000 images with faces at various angles and orientations, along with an additional 20 million images without faces [10].

The evolution and success of CNN algorithms, demonstrated through these studies, underscore their versatility and effectiveness across a spectrum of applications, contributing significantly to advancements in image processing and related fields.

CNN models have shown versatility beyond image processing and can be applied to various NLP problems. Exceptional results were achieved in semantic parsing [11], query generation [12], sentence modeling [13], classification [14], and prediction problems [13]. CNN algorithms have also been employed in drug discovery. AtomNet, developed by Atomwise in 2015, was the first deep neural network designed for drug design. Trained on 3D representations of chemical reactions, the system was used to discover new biomolecules for diseases such as Ebola and sclerosis [16].

2.1. Visual Geometry Group

The Visual Geometry Group (VGG) is a research group at the University of Oxford, and VGG represents a series of deep learning models developed by this group. VGG models, particularly the ones used for image classification tasks on large datasets like ImageNet, have achieved high accuracy. A distinctive feature of VGG models is their deep architecture in convolutional neural networks (CNN). For example, the VGG16 model has a 16-layer structure, consisting of 13 convolutional layers and 3 fully connected layers. This depth allows the model to effectively learn complex features.

VGG models use convolutional layers to identify different features in an image. These convolutional layers enable the model to understand the patterns in the image by creating feature maps. VGG models trained on extensive datasets often offer the ability to use their pre-trained weights for transfer learning or fine-tuning. This feature allows for a quick and effective start in new tasks. VGG models have shown significant success in various computer vision tasks, such as image classification, object detection, and localization. Their impressive performance in ImageNet competitions highlights the impact of VGG in both industry and research domains.

The VGG, or Visual Geometry Group, represents a typical design for a deep Convolutional Neural Network (CNN) characterized by numerous layers. The term “deep” refers to the significant layer count, with VGG-16 or VGG-19 having 16 or 19 convolutional layers, respectively. The VGG architecture serves as the foundation for constructing innovative models for object identification. As a deep neural network, VGGNet surpasses benchmarks across various tasks and datasets beyond ImageNet. It continues to be one of the most employed architectures for image recognition in use today. Figure 2 shows the VGG architecture structure.

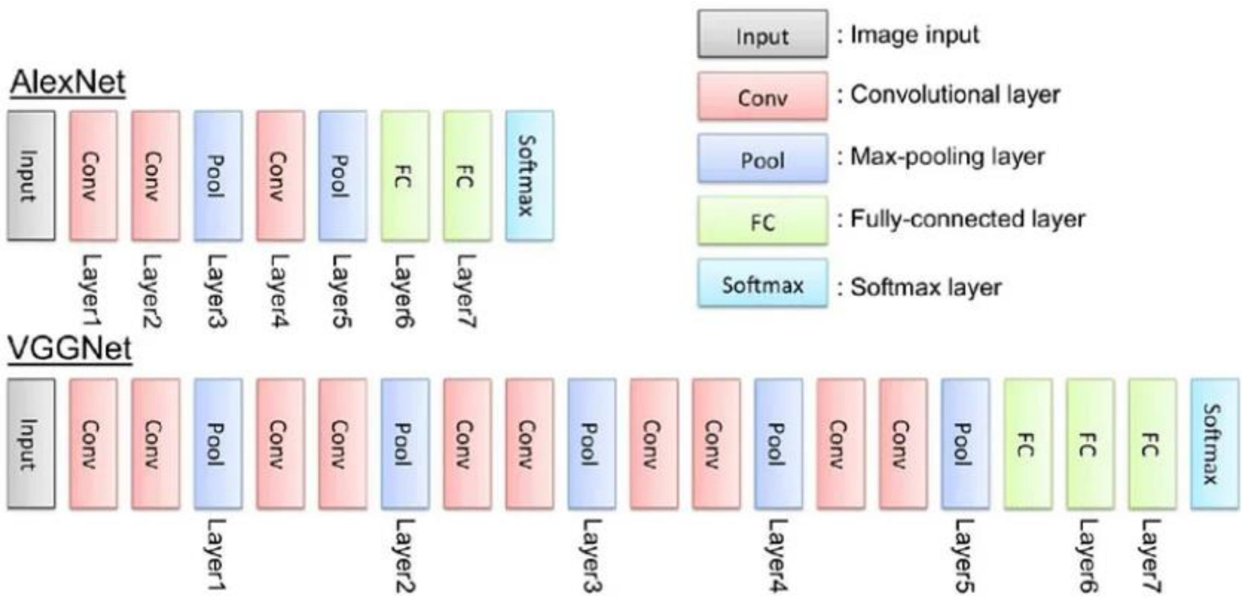


Figure 2: VGG architectures.

2.2.1. VGG 16

The convolutional neural network model known as VGG, or VGGNet, specifically with 16 layers, is commonly referred to as VGG16. This model was developed by A. Zisserman and K. Simonyan at the University of Oxford. The research paper titled "Very Deep Convolutional Networks for Large-Scale Image Recognition" presents the model released by these researchers.

The VGG16 model attains a top-5 test accuracy of approximately 92.7 percent on ImageNet. ImageNet, a dataset comprising over 14 million photos across nearly 1000 categories, served as the testing ground. VGG16 emerged as one of the most favored models during ILSVRC2014. Its superior performance over AlexNet is attributed to the substitution of multiple 3x3 kernel-sized filters for larger ones. The training process for the VGG16 model spanned several weeks, utilizing Nvidia Titan Black GPUs.

With 16 layers, VGGNet-16 excels in classifying images into 1000 distinct object categories, including items like keyboards, animals, pencils, mice, etc., as mentioned earlier. The model accommodates images with a resolution of 224 by 224 pixels. Figure 3 shows the VGG16 architecture structure.

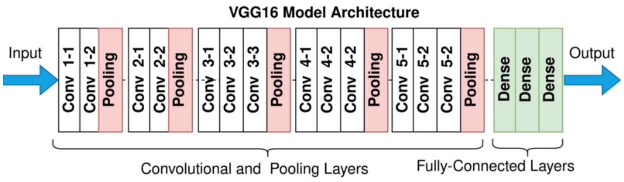


Figure 3: VGG16 architecture.

2.2.2. VGG 19

The VGG19 model, also known as VGGNet-19, follows the same fundamental concept as the VGG16 model, except it features 19 layers. The numbers "16" and "19" correspond to the model's weight layers, specifically the convolutional layers. In contrast to VGG16, VGG19 incorporates three additional convolutional layers.

3. Dataset Preparation

In this experimental study, the choice of utilizing the Multi-type Aircraft of Remote Sensing Images (MTARSI) dataset over traditional datasets like Pascal VOC, MS COCO, CIFAR, ImageNet, and MNIST stems from several critical considerations, elucidating the rationale behind the selection of specific aircraft types, namely "A-10", "B-52", "C-21", and "F-22". Additionally, example photos can be seen in Figures 4, 5, 6, and 7.

The MTARSI dataset offers a unique advantage due to its specialized focus on remote sensing images of various aircraft types, specifically curated from Google Earth satellite imagery. With a total of 9,385 meticulously collected images, spanning 20 distinct aircraft types across 36 airports, MTARSI represents a comprehensive and diverse collection that presents a rich landscape for training and validating object recognition and detection models.

The decision to narrow down the selection to "A-10", "B-52", "C-21", "F22" aircraft types is founded upon the availability of a substantial number of images for these specific categories within the MTARSI dataset. By focusing on these aircraft types, which are well-represented within the dataset, the study aims to ensure a robust and extensive training and validation process.

Figure 4: A-10 aircraft.

Furthermore, the careful curation and labeling of each image by seven specialists in the field of remote sensing image interpretation underscore the high quality and reliability of the MTARSI dataset. This meticulous labeling process enhances the credibility and accuracy of the dataset, mitigating potential ambiguities or inconsistencies in object annotations.

Figure 5: B-52 aircraft.

Moreover, the selection of these specific aircraft types aligns with the objectives of the study, which may prioritize certain types of aircraft based on their prevalence, significance, or relevance to specific applications or domains. For instance, “A-10”, “B-52”, “C-21”, and “F22” aircraft types may hold particular importance in military or aviation-related contexts, thereby justifying their inclusion in the experimental investigation.

By leveraging the MTARSI dataset and focusing on these specific aircraft types, the study aims to contribute valuable insights and advancements to the field of object recognition and detection, particularly in the domain of remote sensing imagery analysis. This targeted approach not only harnesses the richness and diversity of the dataset but also ensures the relevance and applicability of the findings to real-world scenarios and applications.

The table provided outlines the distribution of data samples across the “Train”, “Validation” and “Test” sets for each of the selected aircraft types: “A-10”, “B-52”, “C-21”, “F-22”. These values play a crucial role in ensuring the efficacy and reliability of the experimental study conducted by using the MTARSI dataset.

Figure 6: C-21 aircraft.

The Train set comprises the largest portion of data samples and is utilized for training the deep learning models. The aircraft types and their numbers in the “training” set are as follows: 324 for “A-10”, 432 for “B-52”, 392 for “C-21”, 672 for “F-22”. This distribution ensures that the models have access to a substantial amount of training data, facilitating the learning process and enabling them to capture diverse patterns and characteristics associated with each aircraft type.

Figure 7: F-22 aircraft.

Following the training phase, the Validation set is employed to fine-tune model parameters and monitor performance during training iterations. The Validation set contains a smaller subset of data samples compared to the Train set, serving as an independent evaluation mechanism to gauge the generalization capabilities of the models. The aircraft types and their numbers in the “validation” set are as follows: 59 for “A-10”, 55 for “B-52”, 49 for “C-21”, 84 for “F-22”.

Finally, the Test set is utilized to assess the overall performance and generalization ability of the trained models on unseen data. The Test set is crucial for evaluating the models’ effectiveness in accurately recognizing and detecting aircraft types under real-world conditions. The aircraft types and their numbers in the “test” set are as follows: 59 for “A-10”, 55 for “B-52”, 49 for “C-21”, 84 for “F-22”.

Overall, this meticulous partitioning of data into Train, Validation, and Test sets ensures a rigorous and systematic evaluation of the deep learning models trained on the MTARSI dataset. By leveraging these carefully curated datasets, the study aims to achieve robust and reliable results in the domain of object recognition and detection for remote sensing images of various aircraft types.

4. The Proposed Aircraft Recognition Approach

In this study, a project was conducted for the recognition of military aircraft using VGG16 and VGG19 models. The MTARSI dataset was utilized throughout the project, following several stages as depicted in Figure 8. These stages comprise two main phases: “data preparation”, “training and testing”.

The initial stage, data preparation, involves several steps. First, the MTARSI dataset was thoroughly examined to ensure its suitability for training Convolutional Neural Networks (CNNs) for aircraft recognition. Relevant data were extracted from the dataset, focusing on obtaining a representative sample of images for various aircraft classes. The extracted data were then categorized into different classes of aircraft to facilitate supervised learning, which is crucial for training the CNN models to differentiate between various types of aircraft. Finally, the data were divided into three subsets: training data (80%), validation data (10%), and test data (10%). Training data was used to train the CNN models, validation data to validate the model during training and adjust hyperparameters, and test data to evaluate the final performance of the trained model.

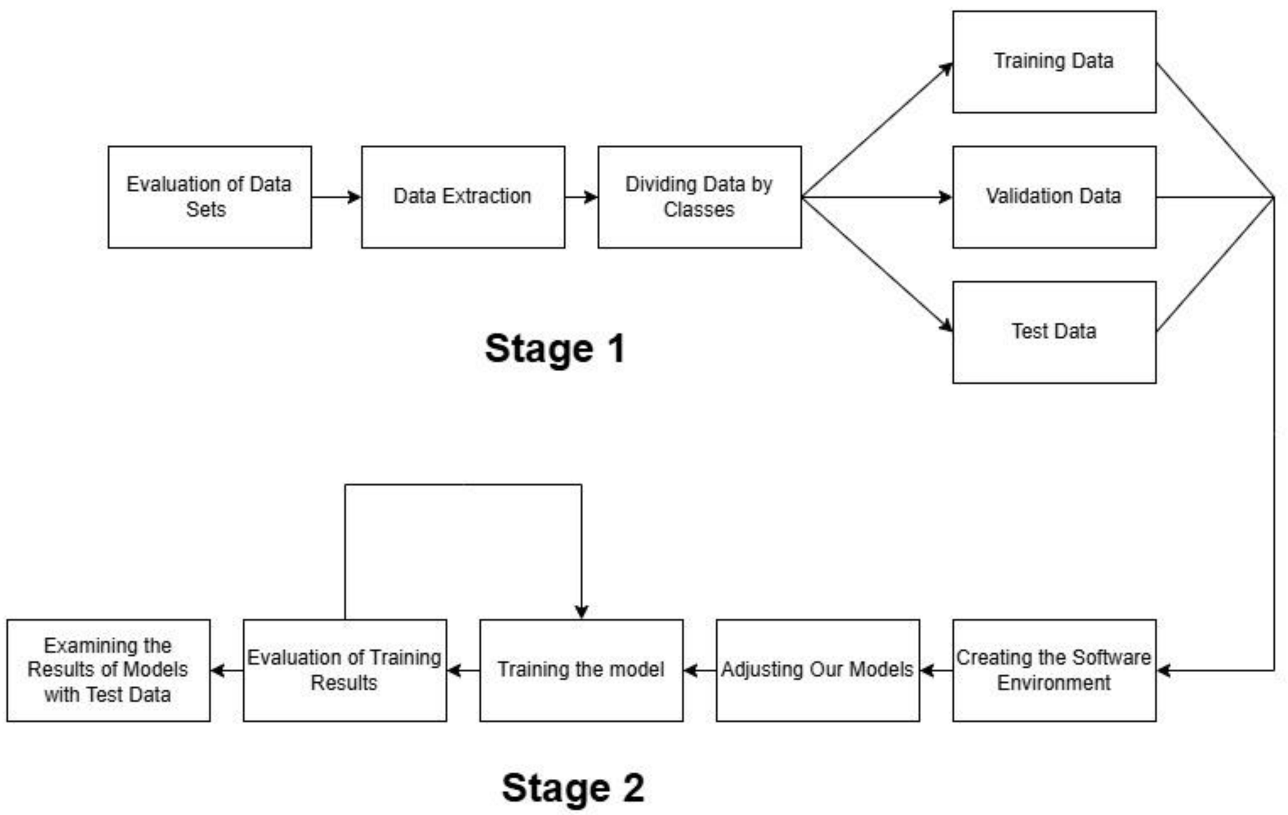


Figure 8: Flowchart of the proposed aircraft recognition approach.

The second stage, model training and testing, encompasses several sub steps. A suitable software environment was set up, including the installation of necessary libraries and frameworks such as TensorFlow and Keras. Initial parameters for the VGG16 and VGG19 models were set, and to prevent overfitting, the maximum epoch value was set to 10. The models were then trained by using the training data, which involved feeding the images into the models and optimizing the parameters through backpropagation. The performance of the models was evaluated by using the validation data, monitoring metrics such as accuracy and loss, and necessary adjustments were made to the models. Based on the evaluation results, parameters were fine-tuned, and the models were re-trained to improve the performance. Finally, the models were tested by using the test data, generating a confusion matrix and calculating test accuracy to assess the overall performance of the models.

The final trained models were evaluated by using the test data. The confusion matrix provided insights into the classification performance across different aircraft classes, and the test accuracy metric helped in quantifying the overall effectiveness of the models. The results indicated that both VGG16 and VGG19 models were successful in recognizing military aircraft with high accuracy, demonstrating the potential of CNNs in satellite image analysis for aircraft recognition.

In conclusion, the presented method successfully implemented VGG16 and VGG19 models for the recognition of military aircraft using satellite images from the MTARSI dataset. Through a structured approach involving data preparation, model training, and testing the models were optimized, and they demonstrated robust performance. Future work could explore the integration of additional data augmentation techniques and the use of other advanced CNN architectures to further enhance recognition accuracy.

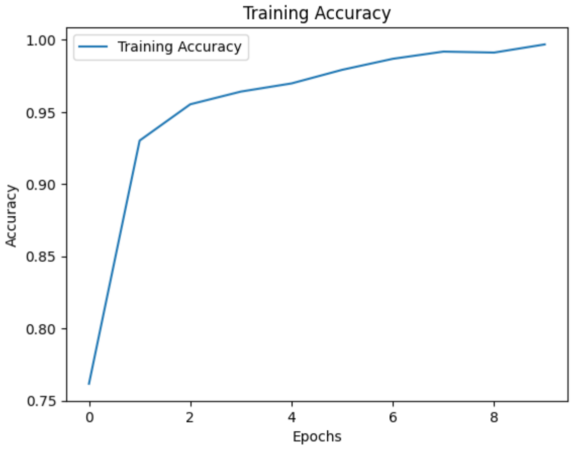
5. Experimental Results and Discussions

In this study, deep learning models, specifically VGG-16 and VGG-19, were trained on the MTARSI dataset to classify aircraft images. The VGG-16 model demonstrated excellent performance, achieving an accuracy of 99.31% on the training set and 96.35% on the validation set (Figure 9). These results reflect both a strong fit to the training data and a robust generalization capability, indicating that the VGG-16 model successfully learned to classify images from the MTARSI dataset. However, given the widespread use of these architectures in similar studies, it is essential to explore novel modifications or techniques that could further differentiate the approach and improve its performance beyond the existing methods.

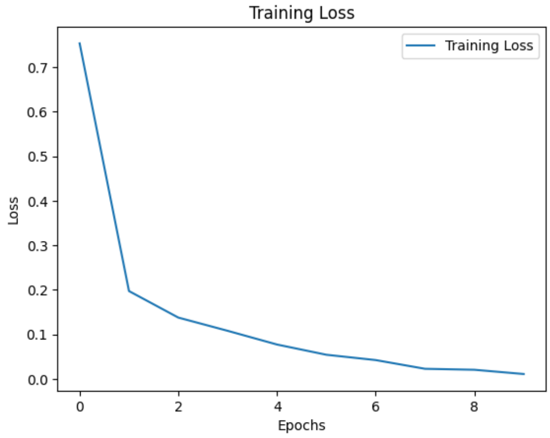
In comparison, the VGG-19 model, with its deeper architecture, also demonstrates excellent performance. Achieving a training accuracy of 99.62% and a validation accuracy of 98.96% (Figure 9), VGG-19 shows a strong fit to the training data and robust generalization to unseen data. While the VGG-19 model provides slightly higher accuracy than VGG-16, the marginal performance increase raises questions about the trade-off between model complexity and computational resources. The added depth of VGG-19 allows for more complex feature extraction, but future studies should investigate whether this increase in accuracy is substantial enough to justify the higher computational cost associated with deeper architectures.

These findings demonstrate that both VGG-16 and VGG-19 architectures can be effectively utilized for aircraft classification on the MTARSI dataset, each offering distinct advantages. While VGG-16 is more computationally efficient, with fewer layers and faster training times, VGG-19's deeper architecture enables more nuanced feature extraction, which could improve classification accuracy in more complex tasks. However, given the extensive use of these architectures in similar studies, future work could explore modifications, such as incorporating advanced regularization techniques or hybrid models, to further enhance the performance of these networks, particularly in specialized image datasets like MTARSI.

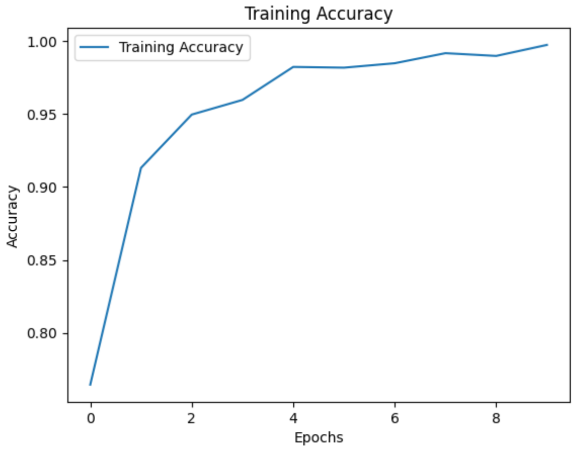
In recent years, deep learning models have revolutionized the field of computer vision, achieving unprecedented performance across a wide array of tasks. The Visual Geometry Group (VGG) architecture family has become one of the most influential approaches for image classification, delivering remarkable results on benchmark datasets such as ImageNet. Despite the success of VGG models in standard applications, there is an ongoing need to innovate and optimize these architectures for specific domains, such as satellite imagery or specialized datasets like MTARSI. Exploring modifications, such as layer-wise fine-tuning or novel data augmentation strategies, could further enhance their performance and applicability.



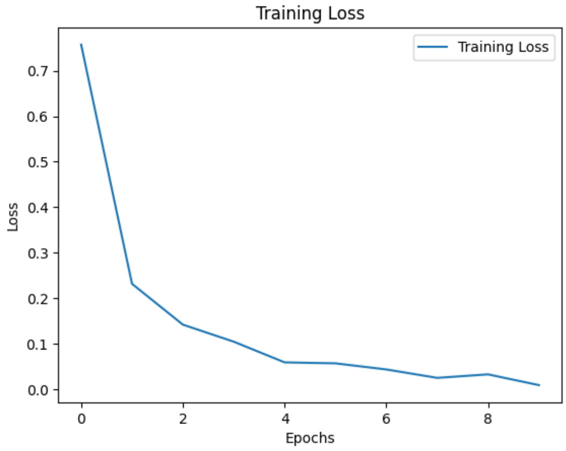
(a) VGG-19



(c) VGG-19



(b) VGG-16



(d) VGG-16

Figure 9: Training accuracy and Training loss for VGG-16 and VGG-19.

On a separate test dataset, VGG-19 achieved an accuracy of 85.03%, slightly surpassing VGG-16, which reached an accuracy of 84.71%. This result underscores the advantage of deeper architectures like VGG-19, which are capable of capturing more complex features from the data. However, the small difference in performance between the two models raises important questions regarding the trade-off between network depth and computational efficiency. Future research could explore alternative methods, such as model pruning or the use of lighter architectures, to maintain or even surpass these accuracy levels while reducing training time and resource usage.

In conclusion, our extensive experiments demonstrate that deep learning models, particularly VGG-19, significantly outperform traditional image processing methods across all evaluation metrics. The superior performance of these models, achieving accuracy rates of up to 85.03%, reinforces the paradigm shift towards deep learning approaches in modern image classification tasks. These findings have important implications for both theoretical understanding of deep architectures and practical applications in real-world scenarios.

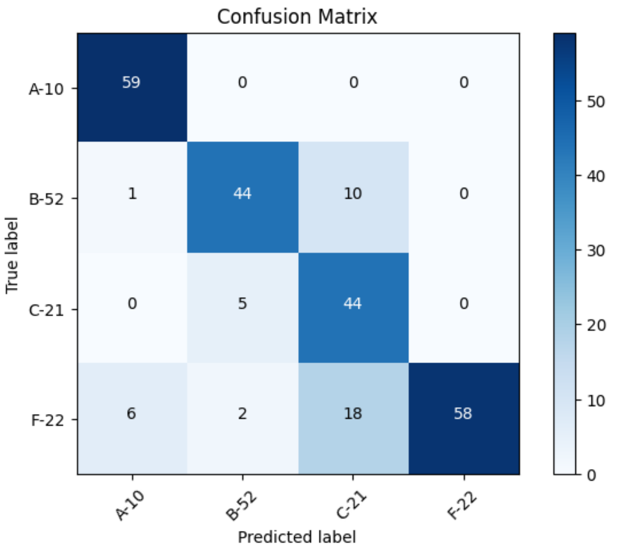


Figure 10: Confusion matrix for VGG-16.

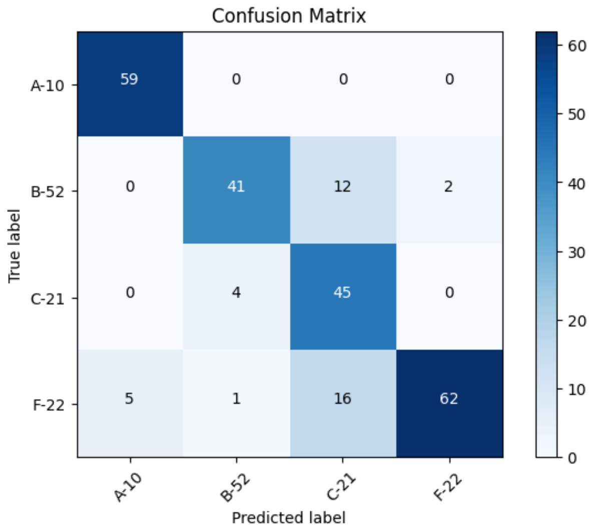


Figure 11: Confusion matrix for VGG-19.

The performance results of two different deep learning models (VGG-16 and VGG-19) are shown in the confusion matrices. Figure 10 presents the results of the VGG-16 model. In this model, there were 59 correct predictions for the A-10 class, 44 for the B-52 class, 44 for the C-21 class, and 58 for the F-22 class. Overall, the model demonstrated good classification performance although some classes exhibited confusion. Figure 11 displays the results of the VGG-19 model. In this model, there were 59 correct predictions for the A-10 class, 41 for the B-52 class, 45 for the C-21 class, and 62 for the F-22 class. The VGG-19 model performed better than VGG-16, particularly for the F-22 class, but it achieved slightly lower success for the B-52 class. Both models managed to predict the A-10 class exceptionally well.

Table 1 presents the accuracy rates of various machine learning (ML) and deep learning (DL) models across four distinct classes: A-10, B-52, C-21, and F-22. As highlighted in recent studies, deep learning models consistently outperform traditional machine learning approaches in classification accuracy. Specifically, convolutional neural networks (CNNs), such as VGG-16 and VGG-19, show exceptional performance across nearly all classes when compared to traditional methods like SIFT + BOWN, HOG + SVM, ScSPM, and LLC. These findings are consistent with recent advancements in deep learning-based image classification, which have demonstrated improved accuracy in complex tasks,

Table 1: Algorithms' comparison results.

Method	A-10	B-52	C-21	F-22
SIFT [17] + BOWN	0.49	0.56	0.60	0.48
HOG[18] + SVM	0.56	0.60	0.58	0.53
ScSPM [19]	0.61	0.61	0.60	0.51
LLC [20]	0.59	0.63	0.63	0.59
VGG-16 (The proposed model)	**1.00**	**0.80**	0.89	0.69
VGG-19 (The proposed model)	**1.00**	0.74	**0.91**	**0.73**

Several traditional machine learning techniques were evaluated for image classification across the four classes. The SIFT [17] + BOWN method, which combines Scale-Invariant Feature Transform (SIFT) for feature extraction and Bag-of-Words (BOWN) for classification, achieves accuracy rates of 49% for A-10, 56% for B-52, 60% for C-21, and 48% for F-22. The Histogram of Oriented Gradients (HOG) [18] combined with Support Vector Machines (SVM) achieves 56% for A-10, 60% for B-52, 58% for C-21, and 53% for F-22. The ScSPM [19] method, using spatial pyramid matching with sparse coding, reaches 61% for A-10, 61% for B-52, 60% for C-21, and 51% for F-22. Lastly, the LLC [20] method, which employs locality-constrained linear coding, achieves accuracy rates of 59% for A-10, 63% for B-52, 63% for C-21, and 59% for F-22. The hyperparameters, including the number of clusters for SIFT and kernel choice for SVM, are discussed in the experimental results section.

In terms of deep learning models, both VGG-16 and VGG-19 achieve significantly higher accuracy across all classes compared to traditional methods. The VGG-16 model achieves a perfect accuracy rate of 100% for A-10, 80% for B-52, 89% for C-21, and 69% for F-22. Similarly, the VGG-19 model reaches 100% accuracy for A-10, 74% for B-52, 91% for C-21, and 73% for F-22. These results align with the findings in recent literature, where VGG architectures were demonstrated to outperform other deep learning models in similar classification tasks. The choice of hyperparameters, such as learning rate, batch size, and the number of layers, plays a crucial role in model performance and is further detailed in the experimental results section.

Performance analysis by class reveals that the VGG-16 and VGG-19 models achieve the highest accuracy in the A-10 class, both attaining 100%, while the traditional SIFT + BOWN method demonstrates the lowest accuracy of 49%. In the B-52 class, the VGG-16 model yields the highest accuracy of 80%, whereas the SIFT + BOWN method achieves the lowest at 56%. The C-21 class sees the VGG-19 model lead with an accuracy of 91%, while the HOG + SVM method results in the lowest accuracy of 58%. Lastly, in the F-22 class, the VGG-19 model achieves the highest accuracy of 73%, with the lowest accuracy of 48% again observed for the SIFT + BOWN method. These results highlight the clear advantages of deep learning models, particularly VGG-19, in achieving higher accuracy across all classes.

In conclusion, deep learning models, especially the VGG-16 and VGG-19 architectures, significantly outperform traditional machine learning methods across all four classes. The VGG-19 model stands out, achieving the highest accuracy across all classes, including a perfect 100% accuracy for the A-10 class. The superior performance of these models, particularly in complex image classification tasks, emphasizes the potential of deep learning in overcoming the limitations of traditional methods. Unlike previous studies that primarily focused on individual models, this study demonstrates the clear advantages of VGG-16 and VGG-19 in a multi-class classification scenario, positioning them as preferred choices for machine learning applications in similar domains.

Table 2 presents a comparison of the accuracy rates achieved by various image classification methods. Traditional image processing and machine learning techniques, such as SIFT [17] combined with Bag-of-Words of Visual Words (BOVW) and Histogram of Oriented Gradients (HOG) [18] with Support Vector Machines (SVM), yield relatively lower accuracy rates of 53.25% and 56.75%, respectively. Other conventional methods, such as ScSPM [19] and LLC [20], demonstrate slightly higher performance, achieving accuracy rates of 58.25% and 61.00%, respectively. These results reflect the limitations of traditional approaches in handling more complex image classification tasks.

Table 2: Accuracy comparison.

Method	Accuracy
SIFT [17] + BOWN	53.25%
HOG[18] + SVM	56.75%
ScSPM [19]	58.25%
LLC [20]	61.00%
VGG-16 (The proposed model**)**	**84.71%**
VGG-19 (The proposed model**)**	**85.03%**

In contrast, deep learning models demonstrate a marked improvement in performance. Specifically, our experiments with the VGG-16 and VGG-19 models yield accuracy rates of 84.71% and 85.03%, respectively. These results underscore the superior capability of deep learning approaches in image classification tasks when compared to traditional methods. The higher accuracy achieved by these models is consistent with recent research highlighting the advantages of deep learning in complex classification problems. The success of deep learning models can be largely attributed to their deep, layered architectures and the large-scale datasets they are trained on, which enable these models to learn and distinguish complex patterns in images with greater efficiency. In particular, the VGG-19 model, with its deeper architecture, exhibits a slight performance advantage over VGG-16 in terms of accuracy. This reflects the potential of deeper models to capture more intricate features and improve classification outcomes.

In conclusion, deep learning models, particularly the VGG-19 architecture, have demonstrated superior accuracy rates compared to traditional image processing methods. This reinforces the growing evidence supporting the effectiveness of deep learning techniques in image classification tasks, especially in handling complex data and achieving high-performance results. The findings of this study contribute to the increasing body of research advocating for the use of deep learning in practical classification applications.

6. Conclusions

In this study, we investigated the application of VGG-16 and VGG-19 architectures for military aircraft classification using satellite imagery, achieving accuracies of 84.71% and 85.03% respectively. Additionally, our experiments showed that maximum efficiency was reached at 10 epochs, demonstrating efficient model convergence. While these results demonstrate improvement over traditional methods such as SIFT+BOWN (53.25%) and HOG+SVM (56.75%), it is important to acknowledge that our approach primarily utilized standard implementations of established architectures without significant novel modifications. The performance metrics across different scenarios varied considerably, with the highest accuracy achieved for A-10 classification (1.00) and lower performance for B-52 (0.74) and F-22 (0.73) classifications, suggesting room for improvement in handling certain aircraft types. Our comparative analysis reveals that while the results are promising, the improvements are incremental rather than transformative in nature.

The study's limitations highlight several opportunities for future research directions, including the development of custom architectural modifications specifically designed for satellite imagery analysis, integration of domain-specific preprocessing techniques, and exploration of hybrid approaches combining traditional computer vision methods with deep learning. Furthermore, the field would benefit from investigation into more recent architectural innovations and the incorporation of domain-specific knowledge into model design. Looking ahead, we believe that advancing this research area requires moving beyond standard implementations to develop novel methodological approaches that specifically address the unique challenges of satellite-based aircraft classification. This could include developing specialized layers or modules within the neural network architecture, implementing advanced data augmentation techniques specific to aerial imagery, and creating more robust feature extraction methods that better handle variations in aircraft orientation and atmospheric conditions. While our current results demonstrate the viability of deep learning approaches for military aircraft classification, they also underscore the need for more innovative solutions to push the boundaries of what is possible in this specialized domain.

Article Information Form

Financial Disclosure: The author (s) has no received any financial support for the research, authorship or publication of this study.

Authors' Contribution: The authors contributed equally to the study.

Conflict of Interest/Common Interest: No conflict of interest or common interest has been declared by the authors.

Ethics Committee Approval: This study does not require ethics committee permission or any special permission.

References

[1] Chen, H., Zhao, J., Gao, T., and Chen, W. (2018). Fast Airplane Detection with Hierarchical Structure in Large Scene Remote Sensing Images at High Spatial Resolution. *IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, July 22-27, 4846–4849.*

[2] Luo, Q., and Shi, Z. (2016). Airplane detection in remote sensing images based on Object Proposal. *International Geoscience and Remote Sensing Symposium (IGARSS), 1388–1391.*

[3] Fukushima, K. N. (1980). A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol. Cybern., 36(4): 193–202.*

[4] Hubel, D. H. and Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *J. Physiol., 195(1):215–243.*

[5] Le Cun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proc. IEEE, 8(11): 2278–2324.*

[6] Le Cun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature*, 521: 436–444.

[7] Le Cun, Y., (1989). Handwritten digit recognition: applications of neural network chips and automatic learning. *IEEE Commun. Mag., 27(11): 41–46.*

[8] Cirean, D., Meier, U., and Schmidhuber, J. (2012). Multi-column Deep Neural Networks for Image Classification., Feb. 2012.

[9] Cirean, D. C., Meier, U., Masci, J., and Gambardella, L. M. (2012). Flexible, High Performance Convolutional Neural Networks for Image Classification. *in Proceedings of the 22nd International Joint Conference on Artificial Intelligence, 1237–1242.*

[10] Farfade, S. S., Saberian, M., and Li, L.-J. (2015). Multiview Face Detection Using Deep Convolutional Neural Networks. *Proceedings of the 5th ACM on International Conference on Multimedia Retrieval. 643–650.*

[11] Grefenstette, E., Blunsom, P., de Freitas, N., and Hermann, K. M. (2014). A Deep Architecture for Semantic Parsing. *The Semantic Parsing Workshop, April.*

[12] Shen, Y., He, X., Gao, J., Deng, L., and Mesnil, G. (2014). Learning semantic representations using convolutional neural networks for web search. *23rd International Conference on World Wide Web - WWW '14 Companion, 373–374.*

[13] Kalchbrenner, N., Grefenstette, E., and Blunsom, P. (2014). A Convolutional Neural Network for Modelling Sentences. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics. 655-665.*

[14] Kim, Y. (2014). Convolutional neural networks for sentence classification. *Conference on Empirical Methods in Natural Language Processing (EMNLP). 1746-1751.*

[15] Collobert, R. and Weston, J. (2008). A unified architecture for natural language processing: Deep neural networks with multitask learning. *25th International conference on Machine learning. 20(1):160–167.*

[16] Wallach, I., Dzamba, M., and Heifets, A. (2015). AtomNet: A Deep Convolutional Neural Network for Bioactivity Prediction in Structure-based Drug Discovery.

[17] Lowe, D.G. (2004). Distinctive image features from scale-invariant keypoints, Int. *J. Comput. Vision. 60(2): 91–110.*

[18] Dalal, N., Triggs, B. (2005). Histograms of oriented gradients for human detection, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2005), San Diego, CA, USA. 886–893.*

[19] Yang, J., Yu, K., Gong, Y., Huang, T.S. (2009). Linear spatial pyramid matching using sparse coding for image classification, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009). 1794–1801.*

[20] Yu, K., Zhang, T. and Gong, Y. (2009) Nonlinear learning using local coordinate coding, *Advances in Neural Information Processing Systems, Vancouver, Canada, 2223–2231.*