Tikrit Journal of Engineering Sciences (2024); 31(3): 202-211
DOI: http://doi.org/10.25130/tjes.31.3.19

IRAQI Academic Scientific Journals

ISSN: 1813-162X (Print); 2312-7589 (Online)

Tikrit Journal of Engineering Sciences

available online at: http://www.tj-es.com

TJES Tikrit Journal of Engineering Sciences

Airplane Detection Using Deep Learning Based on VGG and SVM

Zainab A. Khalaf [a*], **Auday Al-Mayyahi** [b], **Ammar Aldair** [b], **Phil Birch** [c]

a College of Education for Pure Sciences, University of Basrah, Basrah, Iraq.
b College of Engineering, University of Basrah, Basrah, Iraq.
c School of Engineering and Informatics, University of Sussex, Brighton, UK.

Abstract: Object detection is widely utilized in many applications, such as airport surveillance, prevention of potential collisions, aid in airspace management, and enhancing overall aviation safety. This paper proposes an algorithm for airplane detection regardless of the airplane's model, type, or color variations. The main challenges in automatic airplane detection tasks could be the differences in scale, the orientation of the airplanes, and similarity with other objects. Therefore, an airplane detection system must be designed to achieve good discrimination without the influence of airplane rotation, pose, or resolution. Object detection can be performed by considering three major phases, i.e., feature extraction, detection of an airplane, and evaluation of the airplane. To extract the plane region, a deep feature extraction method is used with the VGG model. The plane is detected using the SVM. Two datasets were used to evaluate the designed system's effectiveness. The results achieved a 99% F score using the Caltech-101 dataset and 98% for the FGVC-aircraft dataset.

Keywords:
Airplane Detection; Deep Learning; Feature Extraction; VGG; SVM.

Highlights:

- *Airplane detection is done using the VGG and SVM fusion.*
- *The deep feature extraction method is effectively used in VGG.*
- *Caltech-101 and FGVC-Aircraft datasets were used to evaluate the designed system's effectiveness.*

ARTICLE INFO

Article history:

Received	19 July	2023
Received in revised form	18 Jan.	2024
Accepted	28 Jan.	2024
Final Proofreading	08 Aug.	2024
Available online	20 Aug.	2024

Citation: Khalaf ZA, Al-Mayyahi A, Aldair A, Birch P. **Airplane Detection Using Deep Learning Based on VGG and SVM**. *Tikrit Journal of Engineering Sciences* 2024; **31**(3): 202-211.
http://doi.org/10.25130/tjes.31.3.19

Corresponding author:

Zainab A. Khalaf
College of Education for Pure Sciences, University of Basrah, Basrah, Iraq.

اكتشاف الطائرات باستخدام التعلم العميق المعتمد على VGG و SVM

زينب علي خلف[١]، عدي المياحي[٢]، عمار الدير[٢]، فيل بيرش [٣]
[١] كلية التربية للعلوم الصرفة / جامعة البصرة / البصرة – العراق.
[٢] كلية الهندسة /جامعة البصرة / البصرة – العراق.
[٣] كلية الهندسة والمعلوماتية/ جامعة ساسكس / برايتون – المملكة المتحدة.

الخلاصة

اكتشاف الأشياء يستخدم على نطاق واسع في العديد من التطبيقات مثل مراقبة المطارات ،تجنب الاصطدامات المحتملة، ،المساعدة في إدارة المجال الجوي وتعزيز سلامة الطيران بشكل عام. في هذا البحث، تم اقتراح خوارزمية للتعرف على الطائرات بغض النظر عن الفئات المختلفة مثل طراز الطائرات ،حجمها ولونها. يمكن أن تتمثل التحديات الرئيسية في مهام الكشف التلقائي عن الطائرات في الاختلافات في شكل الطائرات بالإضافة إلى وضع اتجاهها ومقدار التشابه مع الكائنات الأخرى. لذلك، يحتاج الى تصميم نظام كشف عن الطائرات بحيث يتم التمييز بشكل فعال دون تأثير مجموعة من الخصائص مثل الدوران، اختلاف الاشكال والنماذج، الدقة، النوع، واللون. النظام المصمم لاكتشاف الطائرات يتكون من ثلاث مراحل رئيسية: مرحلة استخراج الميزات ،اكتشاف الطائرة وتقييم دقة الاكتشاف. لاستخراج الميزات، تم استخدام تقنية التعلم العميق (VGG) لايجاد الخصائص الدقيقة. في حين تم الكشف عن الطائرات باستخدام خوارزمية تعلم الاله (SVM). لغرض تقييم النظام المصمم، تم استخدام مجموعتي بيانات هما Caltech-101 و FGVC-Aircraft حيث بينت النتائج باستخدام درجة (F1) نسبة ٩٩٪ لمجموعة بيانات Caltech-101 و٩٨٪ لمجموعة بيانات FGVC-Aircraft .

الكلمات الدالة: اكتشاف الطائرات، التعلم العميق، استخراج الميزات، VGG، SVM.

1.INTRODUCTION

The hybrid attention network model was introduced to design an aircraft detection algorithm. This approach can address the low precision of aircraft picture detection due to differences in the types of aircraft used, large similarities between various models, and severe texture interference [1]. In the literature, computer vision systems have been applied in various applications, such as robotic systems, security, and surveillance systems. Object detection is one of the major problems that must be well addressed and solved by developing robust algorithms. To solve such a problem, hundreds of captured images must be collected and labeled, which could be used for training purposes. Thus, processing algorithms need to be applied to extract the required information. Detection will occur based on detecting the make, model, type, or even the color of planes [2,3]. State-of-the-art algorithms have been introduced to address similar research problems and related works. Liu et al. [4], a conventional neural network was introduced to solve the problem of airplane detection based on corner clustering. The utilized technique was split into two main steps: region selection and detection. First, candidate boundaries were provided using a mean-shift clustering approach to the corners. Then, the conventional neural network was used for performing feature extraction and region detection of selected candidates to find an aircraft if available in the scene. The images used were from the Remote Sensing Object Detection (RSOD) dataset [5]. The dataset contained “446” aircraft images for “4993” aircraft. After testing, the detection accuracy reached 98.29%. Xu et al. [6], a remote sensing deep learning method was suggested for aircraft detection. They designed the LightNet network unit to extract relevant features, considered an efficient backbone network. Furthermore, the backbone network results were enhanced using spatial pyramid pooling based on residual ideas (to separate more important relevant features). Then, a multiscale prediction fusion network (MFPN) was suggested for fusing features from multiple angles to realize a substantial combination of gradients. The purpose of using a multiscale prediction fusion network was to improve the ability of networks to detect very tiny objects and consequently improve the system’s accuracy. For the experiments in this study, a remote sensing aircraft dataset was used, containing “1916” images of “18,637” aircraft in four public datasets. The F score of this dataset was 99.2%. Another work designed an artificial intelligence technique based on deep neural networks to detect aircraft. This scheme consisted of three parts: hybrid global attribution mapping (HGAM) was used for backbone network selection, and two methods were used for visualization of the detector: class-specific confidence score mapping (CCSM) and path aggregation network (PANet). The main purpose of using HGAMs was to measure the deep neural network feature extraction effectiveness by integrating local and global approaches. For additional feature fusion, PANet was utilized to generate multiscale prediction feature maps. However, the CCSM relies on visualization techniques to test the performance of aircraft detection on SAR images. The metric mAP (mean average precision) was used for evaluation, and the best score obtained was “92.02%” [7]. Furthermore, in Wu et al. [8], a region based on conventional neural networks was proposed to improve the detection of aircraft impacts in high-resolution remote sensing images that include intense targets and complex backgrounds. Object detection and segmentation were used in this study. A modified self-calibrated conventional network based on residual network 101 (ResNet101) was utilized to extract more discriminative features and increase the number of dilated convolutions to a specific size to enhance the influence of segmentation. The dataset called Dataset of Object Detection in Aerial Images (DOTA) was used for evaluation and yielded a "1–2%" improvement in accuracy compared to that of the basic network. Alshaibani et al. [9], an airplane detection system for airports was proposed. In which drone images were used to train and evaluate a mask region convolution neural network (RCNN) model. Microsoft Common Objects in Context (COCO) metrics were used for evaluation, and the highest score was 83% for the test dataset. The airplane objects varied in size and geospatial resolution. Thus, these factors influence the detection performance. Lin and Chen [10] proposed a YouOnlyLookOnce-v3 (YOLOv3)-based detection system for aircraft detection by applying a non-maximum suppression strategy to find reliable and excluded redundant edge boxes revealed in overlapping image blocks. The system evaluated the generalization ability of various training data combinations in several challenging cases. The outcome of this model was a 98% F score. Wang et al. [11], problems of various object sizes and complex aircraft background information were addressed using an approach based on convolutional neural networks. The airport runway areas were first detected to generate a mask and rectangular contour map of the whole airport. Next, a deep neural network named the efficient weighted feature fusion and attention network (EWFAN) was used to recognize aircraft. To enhance feature extraction, the EWFAN integrated the weighted feature fusion module, the spatial attention mechanism, and the CIF loss function. In the final step, the airport runway mask was applied to reduce false alarms and achieve the final aircraft detection results. Large-scale Gaofen-3 SAR images with a '1 m' resolution were used for evaluation in this study. The detection rate and false alarm rate of the EWFAN algorithm were 95.4% and 3.3%, respectively. Zhong et al. [12] suggested combination approaches by combining a local image patch with its corresponding location. In addition, they proposed a feature compression method to reduce the feature vector dimensionality. They also employed principal component analysis (PCA) and singular value decomposition (SVD) separately for the compressed features. In addition, a position estimation scheme was proposed to enhance the accuracy of detection by determining the centroid of the aircraft. To evaluate this system, the Caltech 101 dataset was used. The detection speed in this study was remarkably improved. The compressed feature preserved 96.4% of the PCA information and 95.2% of the SVD information. Moreover, a recent piece of research [13] trained a classical convolution neural network (LeNet-5) using the shuffled frog-leaping algorithm. Four datasets were used for training, and the experimental results were compared with those of three widely used detection datasets. The methods used were the whale optimization algorithm, bacteria swarm foraging optimization, and ant colony optimization. The outcomes of the proposed system in this study demonstrated an improvement in the original LeNet-5 performance. Additionally, deep learning was proposed for aircraft detection [14]. The DenseNet-161 technique was used to classify dense and highly cluttered aircraft images. The model was tested using two datasets: a wild animal camera trap and a handheld knife. The best accuracy achieved was 95.02% on a wild animal camera trap and 95.20% on a handheld knife dataset. Similarly, a deep learning algorithm was applied for aircraft detection [15]. Various signals were gathered and used for communication between airplanes and the related airport tower. One-dimensional signals of the utilized data were converted into two-dimensional feature maps. Afterward, the feature maps were provided by a convolutional neural network for detection. Thus, the obtained detection results were relatively improved, i.e., the accuracy percentage reached approximately 94.1%. Alganci et al. [16] conducted a comparative study of airplane detection via deep learning methods applied to high-resolution satellite images. The deep learning methods used were Faster R-CNN, single shot multibox detector (SSD), and You Look Only Once-v3 (YOLO-v3). These methods automatically detect airplanes with a limited amount of labeled data. Zhang et al. [17] presented a comprehensive survey of classic models for computer vision, feature representation approaches, deep learning methods, and the frequent datasets used. Zhang et al. [18] introduced an arsenic NetPlus neural network for detecting cassava disease using three signals. A depth convolution was applied to remove aliasing signals and down-convert signals. Additionally, an instance batch normalization algorithm was used to maintain the features of the convolutional neural network channels in an appropriate form. Then, the Arsenic Plus block was utilized to generate high-frequency pseudo images in the residual structure. The Cassava and FGVC-Aircraft datasets, i.e., fine-grained visual classification of aircraft, were tested to evaluate the overall performance of the proposed system. The highest accuracy was obtained based on the FGVC-Aircraft dataset, for which the accuracy reached 86.59%. Wang et al. [1] improved the two-channel ResNet34 model based on the 156 characteristic extraction and depth of the network to improve the fine-grained extraction 157 capability without affecting the dimensions of the output characteristics. The FGVC-158 aircraft dataset was used to evaluate the effectiveness and the recognition precision rate 159 (RPR), for which the simulation result reached 89.2%. Based on the conducted literature review, it has been noticed that there is still room to improve the detection accuracy further. Therefore, this paper has utilized a combination of visual graphics group (VGG) and support vector machine (SVM) to improve and enhance object detection of airplanes. The transfer learning VGG was used to extract features from input images, and then SVM was used as a detector to locate an airplane within an image. Two datasets have been used for this purpose, i.e., Caltech-101 and Fine-Grained Visual Classification of Aircraft (FGVC-Aircraft). The contributions of this paper could be understood by, firstly, integrating VGG and SVM to obtain the best detection accuracy compared with the state of the art. Secondly, an accurate boundary box is obtained regardless of the complexity of the input image. Finally, the performance of the proposed system has been effectively enhanced on various volumes of datasets.

2.DEEP LEARNING AND MACHINE LEARNING METHODS

The deep learning technique is based on the VGG methodology, which is required for feature selection. The other machine learning methodology is based on an SVM, which is required for airplane detection. The operation and topology of these two methodologies have been illustrated in the following subsections.

2.1.Visual Graphics Group (VGG)

The VGG network is a convolutional neural network (CNN) architecture designed for image classification. The architecture of VGG mainly consists of 16 layers, i.e., 13 convolutional layers and 3 fully connected layers, as shown in Fig. 1. The VGG is used based on 16 layers to exclude irrelevant features and maintain the required features in the area of interest points [19–21]. The procedure of the VGG can be summarized as follows:

1) The input layer takes an image from three color channels (RGB) to represent the image size (224, 224, and 3).
2) Convolutional Blocks: The convolutional part of the VGG consists of 5 blocks, each followed by a max pooling layer. Each convolutional block contains two or more layers with small 3x3 filters and a rectified linear unit (ReLU) activation function based on the following formula:

$$f(x) = \max(0, x) \qquad (1)$$

3) After the image dimensions have been reduced, the remaining features are fed into one or more fully connected layers. The VGG has three fully connected layers, with the final layer producing the output probabilities for classification.
4) Softmax activation is the last layer used to classify *k* of classes, as in the following formula.

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^{k} e^{z_j}} \qquad (2)$$

Where i=1, ..., k, and z = $(z_1, \ldots, zk_k) \in \mathbb{R}^k$

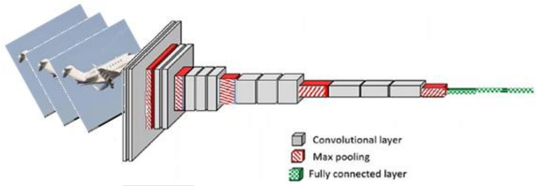

Fig. 1 VGG Architecture.

2.2.Support Vector Machine (SVM)

The SVM classifier is used for detection and is considered an efficient and steady statistical machine-learning algorithm. This algorithm involves separating training data linearly using a hyperplane. The best hyperplane is selected from a set of generated hyperplanes based on the maximum margin, which is the distance between the hyperplane and the nearest points of each class. The points located on the sides that define the margin from each class are termed support vectors, considered the most important vectors regarding classification information, as shown in Fig. 2 [22,23].

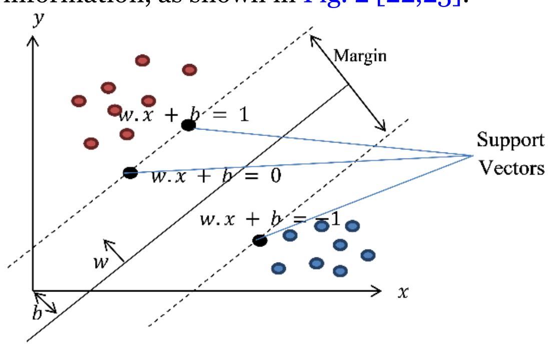

Fig. 2 SVM Hyperplane and Support Vectors.

Suppose the training data are represented by the vectors x_i, where $i = 1,2, \ldots, n$, i.e., and n indicates the dimensional feature space. Each vector belongs to a class label $y_i \epsilon \{-1, +1\}$. The orthogonal vector w and bias b can be determined by finding the hyperplane separator via Eq. (3) [24,25]:

$$w.x + b = 0 \qquad (3)$$

The separator hyperplane is computed by 2/(||w||). Therefore, the following equation is used to measure the maximum margin $min\left(\frac{1}{2}||w||^2\right)$[26]:

$$y_i(w.x_i + b) \geq 1, \ \forall \ i = 1, \ldots, n \qquad (4)$$

Lagrange multipliers $a_i \geq 0$ and $\forall i$ are used to obtain the optimal solution, especially true with nonlinearly separable data. Eq. (5) shows the formula of the Lagrangian, and Eq. (6) shows the Lagrangian with constraints:

$$L(w, b, \alpha) = \frac{1}{2}||w||^2 - \sum_{i=1}^{N} \alpha_i \left(y_i((w.x_i + b) - 1)\right) \qquad (5)$$

$$w = \sum_{i=1}^{N} \alpha_i y_i \qquad (6)$$

The following equation represents the satisfaction optimal solution:

$$\alpha_i[y_i(w.x_i + b) - 1] = 0 \qquad (7)$$

These points are called the support vectors and are used to specify the hyperplane. Therefore, the classifier function can be defined in Eq. (8):

$$f(x) = sgn\{\sum_{i=1}^{n} \alpha_i{}^* y_i(x_i.x + b^*)\} \qquad (8)$$

Where the values of b^* and α^* are obtained from Eqs. (6) and (7), respectively. In large datasets, high dimensionality increases the complexity of detecting an appropriate hyperplane. Thus, the kernel function is used with the SVM algorithm to solve this problem, and as a result, the "curse of dimensionality" is avoided. The kernel function transforms the feature dimension space into a new feature dimension space that can be separated linearly, as demonstrated in Fig. 3 [27,28]. The kernel function is defined in Eq. (9):

$$K(x_i, x_j) = \emptyset(x_i).\emptyset(x_j) \qquad (9)$$

The main advantages of the SVM algorithm include accurate results and the ability to avoid overfitting problems with an appropriate kernel. In addition, it is robust to use with high dimensionality [23,29]. The one-class support vector machine (OCSVM) is a type of support vector machine (SVM) designed for one-class classification that leads to detection. The OCSVM seeks to find an optimal decision boundary that separates a given dataset into one primary class (normal instances) and everything else (outliers).

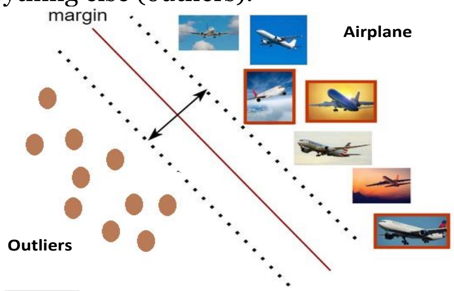

Fig. 3 OCSVM for Airplane Detection.

3.PROPOSED SYSTEM FOR AIRPLANES DETECTION

Object detection is the process of automatically detecting objects of a specific class of visual objects, such as humans, animals, cars, and airplanes. The object detection system splits the given data into training and testing data. During the training phase, the training data are used to obtain the optimal parameters and subsequently utilized to predict the object label to the unseen testing data in the testing phase. In the present paper, airplanes are detected using an automated detection system. This system addresses airplanes in noisy images by identifying airplanes automatically based on predefined image-labeled images. The architecture of the proposed system is based on using a transfer learning approach, in which the VGG extracts features based on '13' convolutional layers only and stops the other remaining three layers, as shown in Fig. 4. In other words, this approach involves removing the fully connected layers. Once the features are extracted using the VGG model, a one-class support vector machine (OCSVM) is used as a classifier for detecting airplane objects. The principle of OCSVM emphasizes a one-class classification trained on a single class of data (airplane instances). Otherwise, any other objects are considered outliers, i.e., not airplanes. The adopted object detection system, shown in Fig. 5, comprises four main phases. These phases include pre-processing, feature extraction, detection, and evaluation phases. Pre-processing is a paramount phase utilized to prepare the object images for the next phase. Therefore, to extract the features and exclude the unimportant features, deep learning filtering is used during the feature extraction phase. After there, these extracted features are passed into the object detection phase to identify the object label by finding the best results. In the final phase, the results are evaluated by computing the system performance. The details of these phases are explained in the following subsections.

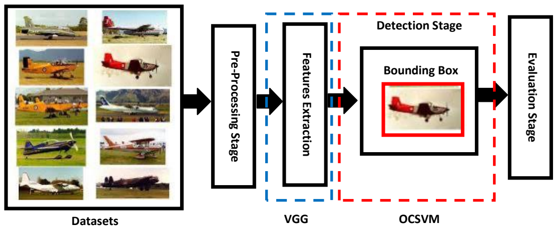

Fig. 5 The Adopted Proposed System Stages of Airplane Detection.

3.1.Datasets

The experiments were conducted using the Caltech-101 dataset and the FGVC-Aircraft dataset. The Caltech-101 dataset comprised 8,677 images of 101 classes, i.e., between 40 and 800 images per class. Although the dataset included 800 images of various airplanes, it contained many images that included various objects, not only airplanes. Additionally, each image had various backgrounds and scenes. Furthermore, the postures of airplanes might differ, and the illumination conditions might vary. The images were also different in terms of color, model, resolution, view angles, and degree of rotation. For training and validation, the airplane class was used. The FGVC-Aircraft dataset consisted of 10,000 images that also share the same properties as those explained for the other dataset. The images for each dataset were divided into 80% of the images for training and 20% for testing. Figure 6 shows an example of a dataset image with multiple objects. In the training phase, the VGG was applied to extract deep features input into the SVM model to construct the object detection system. Then, the test data were automatically used to detect airplanes via the constructed model. The final results demonstrated effective detection with a bounded box.

Fig. 6 Samples of Dataset Images with Multiple Objects.

3.2.Pre-Processing Phase

Pre-processing is an essential phase for differentiating between the features of airplanes and other features in addition to noisy data, requiring the raw data before conducting the feature extraction phase. In the pre-processing phase, the images were resized before being passed as inputs to the next stage. All the images used were warped to a fixed size with minimum shrinkage to avoid accuracy degradation due to deformations. Consequently, the size of the needed memory and computational procedures for image processing would decrease [30,31]. In the present paper, the bilinear interpolation method was used to resize the images. This method is a mathematical technique widely used in image processing to estimate pixel values. The method also derives its name from its interpolation of pixel values by considering the contributions of the four nearest neighboring pixels that exist in an original image [32]. Selecting the four nearest pixels was conducted based on Bilinear interpolation, i.e., integer coordinates (x_1, y_1), (x_2, y_1), (x_1, y_2), and (x_2, y_2) in any original image. The target point (x, y) was surrounded in all corners to form a square based on the aforementioned four pixels. The pixel value at (x, y) coordinates was calculated using the weighted average method for all values of the above four pixels. Assuming the following:

1) $dx=x-x_1$ (fractional part of x)
2) $dy=y-y_1$ (fractional part of y)
3) The points: $(x_1, y_1), (x_2, y_1), (x_1, y_2), and\ (x_2, y_2)$ are the pixel values at the four nearest integer coordinates.

Estimating the pixel value at (x, y) coordinates based on bilinear interpolation equation are found as follows:

$$\textbf{The newly obtained pixel value} = (1-dx).(1-dy).I(x_1,y_1) + dx.(1-dy).I(x_2,y_1) + (1-dx).dy.I(x_1,y_2) + dx.dy.I(x_2,y_2) \quad \textbf{(10)}$$

Where $(1-dx).(1-dy)$ represents the weight for the pixel at (x_1, y_1), $dx.(1-dy)$ represents the weight for the pixel at (x_2, y_1), $(1-dx).dy$ represents the weight for the pixel at (x_1, y_2), and $(dx.dy)$ represents the weight for the pixel at (x_2, y_2). Consequently, the weights were used to calculate the weighted average of the pixel values in which they will obtain the estimated value at the target point (x, y) [33–35].

3.3.Feature Extraction

Feature extraction is essential to properly select and extract the best features from the input images of the utilized dataset. Some features might not be included if irrelevant, so only appropriate features will be retained. In this process, the extracted features were obtained based on the VGG methodology. The features were then mapped into a matrix to be applied and passed to the SVM algorithm to produce the results of classified airplanes [30].

3.4.Detection phase

In this subsection, the VGG model based on transfer learning has been developed to extract deep features effectively. The obtained features were applied as inputs to support vector machine (SVM). The latter has been used to detect airplane objects based on a learning process belonging to the airplane class. This system is subsequently employed to analyze testing images effectively, detecting unseen airplanes and demarcating their positions using bounding boxes. This comprehensive approach harnesses the power of transfer learning to capture intricate feature representations with the VGG model and utilizes SVM for the precise and robust detection of airplanes in a given image.

3.5.Evaluation Phase

Evaluation is necessary for many different purposes and is affected by the number of images used for training and evaluation. The training data were utilized for training the object detection system. Part of these training data was used for error analysis; consequently, the learning algorithm parameters were adjusted according to these errors. The metrics used for evaluation in this study are as follows [36–38]:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad \textbf{(11)}$$

$$F\ Score = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad \textbf{(12)}$$

Where true positive (TP) is the prediction of correct airplane detection. The false positive (FP) denotes the number of detected airplane objects that do not belong to the airplane class. The false negative (FN) is the number of airplane objects not recognized as airplanes, and the true negative (TN) is the number of non-airplane objects correctly classified as negative samples. The precision is the ratio of the number of objects correctly recognized by the proposed system to the total number of objects labeled by the proposed system as positive.

$$Precision = \frac{TP}{TP+FP} \quad \textbf{(13)}$$

The recall (sensitivity) represents the number of correctly detected objects divided by the total number of correct objects in the data.

$$\boldsymbol{Recall} = \frac{TP}{TP+FN} \quad \textbf{(14)}$$

4.RESULTS AND DISCUSSION

This section discusses the results obtained from implementing the proposed system. The experiments were conducted on a PC with a 64-bit Windows 10 operating system, 4 GB of RAM, and an Intel Core i5 1.80 GHz processor. The VGG and SVM detection systems were utilized in this work to obtain the benefits of the CNN and SVM methods. All image features were extracted from the training data to define the class values using VGG. The data images were fed into the CNN to extract the meaningful features necessary for making appropriate decisions. Then, SVM is used for airplane detection. The one-class support vector machine was configured with specific parameters for optimal performance, including a learning rate of 0.01 with a sigmoid kernel and a gamma parameter set to 'scale.' Such parameter choices were strategically selected to balance the trade-off between the fraction of margin errors and the flexibility of the decision boundary provided by the sigmoid kernel. The 'scale' setting for gamma data indicates an automatic adjustment based on the inverse of the dimensionality of the input data, ensuring adaptability to the dataset characteristics. The accuracy, F score, and specifications were used for evaluating the Caltech -101 and FGVC-Aircraft datasets. The total number of images was approximately 800 and 10,000, respectively. These images were separated into two partitions: training and testing. Approximately 600 images for the Caltech -101 dataset and 8,000 images for FGVC-Aircraft datasets were used for learning and validation. However, 200 and 2,000 images were used for the testing model. The experimental results of the proposed system are shown in Fig. 7, illustrating the accuracy and loss. It is observable that the accuracy rates of the training and validation datasets are 100%.

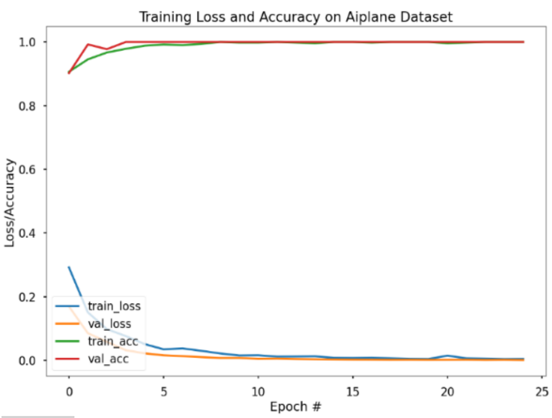

Fig. 7 Loss and Accuracy on Airplane Caltech -101 Dataset.

Figure 8 shows the receiver operating characteristic (ROC) curve, demonstrating the relationship between the true positive and false positive rates.

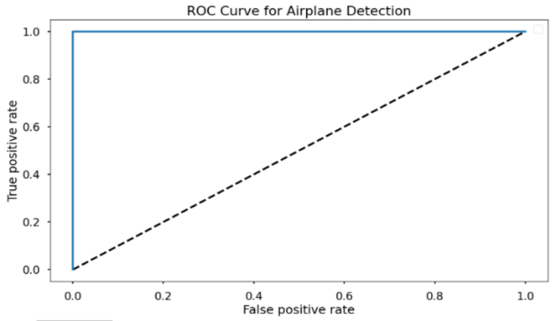

Fig. 8 ROC Curve for Airplane Detection Using Caltech -101 Dataset.

The F scores were 99% and 98% for the testing dataset images from the Caltech-101 and FGVC-Aircraft datasets, respectively. In addition, the experimental results have presented improved detection rates compared to previous relevant works in terms of the type of data used, the number of datasets used, and the detection method used. The data destruction of the Caltech -101 dataset is shown in Fig. 9.

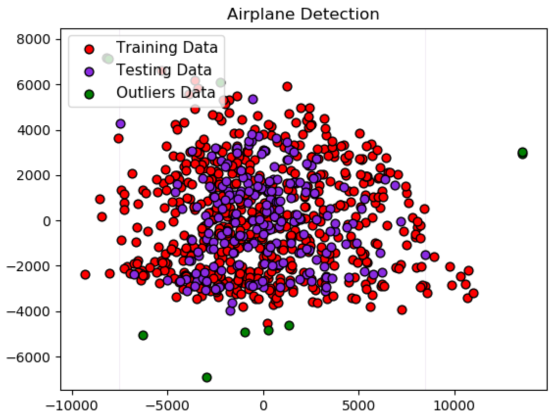

Fig. 9 The Distribution of the Caltech -101 Dataset.

Table 1 reveals comparisons with other relevant works for predicting airplanes. Firstly, in Saadi et al. [13], the obtained accuracy was 97% based on the Caltech-101 dataset using the shuffled Frog-Leaping algorithm. In contrast, the proposed method in this paper achieved a superior F score metric of 99%, demonstrating a significant improvement compared to previous works. Specifically, the F score, balancing precision and recall, underscored the enhanced the proposed method performance. Secondly, for the FGCV dataset, Zhang et al. [18] introduced the Arsenic NetPlus neural network and achieved an accuracy of 86.59%. Wang et al. [1] proposed a hybrid attention network model and attained an accuracy of 89.2%. Notably, the proposed method surpassed these benchmarks, yielding an impressive accuracy metric of 98%. This substantial performance has made observable enhancements compared to existing approaches, showcasing its efficacy in handling the FGCV dataset.

Table 1 Comparisons for the Prediction of Airplanes.

Ref.	Detection method	Dataset	Amount of Data	Detection Performance
[1]	Hybrid attention network model	FGCV dataset	10,000	89.2% Accuracy
[13]	Shuffled Frog-Leaping Algorithm	Caltech 101 airplanes	800	97% Accuracy
[18]	Arsenic NetPlus neural network	FGCV dataset	10,000	86.59% Accuracy
Our proposed system	VGG+SVM	Caltech 101 airplanes	800	99% F score
		FGCV dataset	10,000	98% F score

5.CONCLUSION

This paper presents a deep learning technique based on VGG to perform feature extraction and SVM for airplane detection. To evaluate the proposed methodology's effectiveness and performance, two datasets were utilized, i.e., Caltech-101 and FGVC-Aircraft. After conducting the required experiments, the results of F scores were 99% and 98% for the Caltech-101 and FGVC-Aircraft datasets, respectively. The proposed technique has improved the detection accuracy compared to the-state-of-the-art using the same datasets. In addition, it has been approved that based on the VGGs and SVMs combination effectively contributed to airplane detection in the field of computer vision.

REFERENCES

[1] Wang Y, Chen Y, Liu R. **Aircraft Image Recognition Network Based on Hybrid Attention Mechanism**. *Computational Intelligence and Neuroscience* 2022; **2022**(1): 4189500, (1-9).

[2] Auday Al-Mayyahi, William Wang, Philip Birch AH. **Obstacle Detection System Based on Colour Segmentation Using Monocular Vision for an Unmanned Ground Vehicle**. *International Journal of Intelligent Computing and Cybernetics* 2018; **8**(3): 241–266.

[3] Al-yoonus M, Al-Kazzaz S. **FPGA-SoC Based Object Tracking Algorithms: A Literature Review**. *Al-Rafidain Engineering Journal* 2023; **28**(2): 284–295.

[4] Liu Q, Xiang X, Wang Y, Luo Z, Fang F. **Aircraft Detection in Remote Sensing Image Based on Corner Clustering and Deep Learning**. *Engineering Applications of Artificial Intelligence* 2020; **87**: 103333.

[5] Long Y, Gong Y, Xiao Z, Liu Q. **Accurate Object Localization in Remote Sensing Images Based on Convolutional Neural Networks**. *IEEE Transactions on Geoscience and Remote Sensing* 2017; **55**(5): 2486–2498.

[6] Xu ZF, Jia RS, Yu JT, Yu JZ, Sun HM. **Fast Aircraft Detection Method in Optical Remote Sensing Images Based on Deep Learning**. *Journal of Applied Remote Sensing* 2021; **15**(01):014502-014502.

[7] Luo R, Xing J, Chen L, Pan Z, Cai X, Li Z, Wang J, Ford A. **Glassboxing Deep Learning to Enhance Aircraft Detection from SAR Imagery**. *Remote Sensing* 2021; **13**(18): 3650, (1-19).

[8] Wu Q, Feng D, Cao C, Zeng X, Feng Z, Wu J, Huang Z. **Improved Mask R-CNN for Aircraft Detection in Remote Sensing Images**. *Sensors* 2021; **21**(8): 2618, (1-13).

[9] Alshaibani WT, Helvaci M, Shayea I, Mohamad H. **Airplane Detection Based on Mask Region Convolution Neural Network**. arXiv preprint arXiv:2108.12817 (2021).

[10] Lin YC, Chen WD. **Automatic Aircraft Detection in Very-High-Resolution Satellite Imagery Using a YOLOv3-Based Process**. *Journal of Applied Remote Sensing* 2021; **15**(01): 018502-018502.

[11] Wang J, Xiao H, Chen L, Xing J, Pan Z, Luo R, Cai X. **Integrating Weighted Feature Fusion and the Spatial Attention Module with Convolutional Neural Networks for Automatic Aircraft Detection from SAR Images**. *Remote Sensing* 2021; **13**(5): 910, (1-21).

[12] Zhong J, Lei T, Yao G, Jiang P. **Robust Aircraft Detection with a Simple and Efficient Model**. *Information* 2018; **9**(4): 1–16.

[13] Saadi SB, Sarshar NT, Sadeghi S, Ranjbarzadeh R, Forooshani MK, Bendechache M. **Investigation of Effectiveness of Shuffled Frog-Leaping Optimizer in Training a Convolution Neural Network**. *Journal of Healthcare Engineering* 2022; **2022**: 1–11.

[14] Dhillon A, Verma GK. **A Multiple Object Recognition Approach via DenseNet-161 Model**. In: Sehgal R, Gupta N, Tomar A, Sharma MD, Kumaran V. Smart Electrical and Mechanical Systems: An Application of Artificial Intelligence and Machine Learning. Massachusetts, USA: Academic Press;

2022.

[15] Wang G, Zou C, Zhang C, Pan C, Song J, Yang F. **Aircarft Signal Feature Extraction and Recognition Based on Deep Learning**. *IEEE Transactions on Vehicular Technology* 2022; **71**(9) :9625-9634.

[16] Alganci U, Soydas M, Sertel E. **Comparative Research on Deep Learning Approaches for Airplane Detection from Very High-Resolution Satellite Images**. *Remote Sensing* 2020; **12**(3): 458, (1-28).

[17] Zhang D, Han J, Cheng G, Yang MH. **Weakly Supervised Object Localization and Detection: A Survey**. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2021; **44**(9): 5866-5885.

[18] Zhang J, Qi C, Mecha P, Zuo Y, Ben Z, Liu H, Chen K. **Pseudo High-Frequency Boosts the Generalization of a Convolutional Neural Network for Cassava Disease Detection**. *Plant Methods* 2022; **18**(1): 136, (1-14).

[19] Sulistyowati T, Purwanto P, Alzami F, Pramunendar RA. **VGG16 Deep Learning Architecture Using Imbalance Data Methods for the Detection of Apple Leaf Diseases**. *Moneter: Jurnal Keuangan Dan Perbankan* 2023; **11**(1): 41–53.

[20] Marwaha A, Malik RQ, Beram SM, Rizwan A, Kishore KH, Thakur D, Shabaz M. **Visualisation-Based Binary Classification of Android Malware Using VGG16**. *IET Software* 2023; **17**(4): 717–728.

[21] Khalaf ZA, Hammadi SS, Mousa AK, Ali HM, Alnajar HR, Mohsin RH. **Coronavirus Disease 2019 Detection Using Deep Features Learning**. *International Journal of Electrical and Computer Engineering* 2022; **12**(4): 4364-4372.

[22] Karamizadeh S, Abdullah SM, Halimi M, Shayan J, Rajabi M javad. **Advantage and Drawback of Support Vector Machine Functionality**. *International Conference on Computer, Communications, and Control Technology (I4CT)* 2014; Langkawi, Malaysia. IEEE: p. 63–65.

[23] Akay MF. **Support Vector Machines Combined with Feature Selection for Diabetes Diagnosis**. *Istanbul University - Journal of Electrical and Electronics Engineering* 2017; **17**(2): 3219–3225.

[24] Luo F, Li C, Cao Z. **Affective-Feature-Based Sentiment Analysis Using SVM Classifier**. *2016 IEEE 20th International Conference on Computer Supported Cooperative Work in Design (CSCWD)* 2016; Nanchang, China. IEEE: 276–281.

[25] Han J, Kamber M, Pei J. **Data Mining: Concepts and Techniques**. 4th ed., Massachusetts, USA. Morgan Kaufmann; 2022.

[26] Rustam Z, Nadhifa F, Acar M. **Comparison of {SVM} and {FSVM} for Predicting Bank Failures Using Chi-Square Feature Selection**. *Mathematics, Informatics, Science and Education International Conference (MISEIC)* 2018; Surabaya, Indonesia. IEEE: p. 1-7.

[27] Kesavaraj G, Sukumaran S. **A Study on Classification Techniques in Data Mining**. *Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT)* 2013; Tiruchengode, India. IEEE: p. 1–7.

[28] Byun H, Lee S-W. **Applications of Support Vector Machines for Pattern Recognition: A Survey**. *International Workshop on Support Vector Machines* 2002; Berlin, Heidelberg. Springer: p. 213–236.

[29] Patra A. **A Survey Report on Text Classification with Different Term Weighing Methods and Comparison between Classification Algorithms**. *International Journal of Computer Applications* 2016; **75**(7): 14-18.

[30] Liu YH. **Feature Extraction and Image Recognition with Convolutional Neural Networks**. *First International Conference on Advanced Algorithms and Control Engineering* 2018; National Pingtung University of Science and Technology, Taiwan. IOP Publishing: p. 1-7.

[31] Zhao D, Zhu D, Lu J, Luo Y, Zhang G. **Synthetic Medical Images Using F&BGAN for Improved Lung Nodules Classification by Multi-Scale VGG16**. *Symmetry* 2018; 10(10): 519, (1-16).

[32] Giachetti A, Asuni N. **Fast Artifacts-Free Image Interpolation**. *19th British Machine Vision Conference* 2008; London, UK. BMVC: p. 1–10.

[33] Tam WS, Kok CW, Siu. WC. **Modified Edge-Directed Interpolation for Images**. *Journal of Electronic Imaging* 2010; **19**(1): 013011, (1-20).

[34] Xianming Liu, Debin Zhao, Ruiqin Xiong, Siwei Ma, Wen Gao, Huifang Sun. **Image Interpolation Via Regularized Local Linear Regression**. *IEEE Transactions on Image Processing* 2011; **20**(12): 3455–3469.

[35] Patel V, Mistree K. **A Review on**

Different Image Interpolation Techniques for Image Enhancement. *Emerging Technology and Advanced Engineering* 2013; **3**(12): 129–133.

[36] Xu R, Zeng Q, Zhu L, Chi H, Du X, Guizani M. **Privacy Leakage in Smart Homes and Its Mitigation: IFTTT as a Case Study**. *IEEE Access* 2019; **7**: 63457–63471.

[37] Kodali RK, Gorantla VSK. **RESTful Motion Detection and Notification using IoT**. *International Conference on Computer Communication and Informatics (ICCCI)* 2018; Coimbatore, India. IEEE: p. 1–5.

[38] Khalaf ZA, Shtaet IA. **News Retrieval Based on Short Queries Expansion and Best Matching**. *Journal of Theoretical and Applied Information Technology* 2019; **97**(2): 490-500.