Journal of Technology and System Information Vol: 3, No 1, 2026, Page: 1-17

Classification and Analysis of Real and Fake Aerial Vehicle Images Using Machine Learning

Hasan Aksoy*, Ziya Ozcelik, Yavuz Selim Taspinar

Selcuk University

DOI:
https://doi.org/10.47134/jtsi.v3i1.5345
*Correspondence: Hasan Aksoy
Email: hsnaksoyy@gmail.com

Received: 22-11-2025
Accepted: 22-12-2025
Published: 22-01-2026

Abstract: Aircraft are widely used in both military and civilian fields today. Detecting aircraft in the airspace is of great strategic and societal importance. In recent years, distinguishing images generated by artificial intelligence from real images has become increasingly difficult. This article presents a study on the classification of real aircraft images and AI-generated aircraft images by machine learning algorithms. Six classifications were obtained from 300 images in the dataset. These classifications are: fake commercial aircraft AI, fake military aircraft AI, fake private aircraft AI, real commercial aircraft, real military aircraft, and real private aircraft. These data were classified using common machine learning models such as Artificial Neural Network (ANN), K-Nearest Neighbor (KNN), Support Vector Machine (SVM), and Logistic Regression (LR). Accuracy, Precision, Recall, and F1 Score metrics were used to analyze the classification success of these models. ROC was used for a detailed analysis of the classification success of the models. According to the results obtained, the ANN model achieved a classification success rate of 96.6%, the KNN model 90.4%, the SVM model 96.7%, and the LR model 96.5%. The highest classification success rate was obtained from the SVM model. These results show that all models achieved similar classification success rates, with the KNN model achieving a lower classification success rate than the others. In conclusion, it can be said that all models can be used in the classification of aircraft images.

Keywords: Aerial Vehicle Imagery, Image Classification, Machine Learning, Artificial Neural Networks, False Image Detection

Introduction

The rapid advancements in both manned and unmanned aerial vehicles (UAVs) in today's technology have yielded significant gains in civil and military applications, but also present new challenges in terms of airspace security and surveillance. While aircraft and UAVs are effectively used in many areas such as transportation, reconnaissance, surveillance, and defense, their uncontrolled, unauthorized, or malicious use can lead to serious security problems. Especially in areas with heavy air traffic, the rapid and accurate detection, classification, and tracking of aircraft are of paramount importance. Although traditional aircraft detection systems mostly rely on radar and signal-based solutions, they have limitations in detecting aircraft flying at low altitudes, small in size, or with visually complex backgrounds (Cinar & Taspinar, 2024).

Approaches based on image processing techniques offer a powerful and complementary alternative for the detection and classification of aircraft. Camera-based systems stand out with advantages such as a wide field of view, high resolution, real-time data provision, and low hardware cost. Thanks to developing computer vision methods and machine learning algorithms, aircraft and UAVs of varying sizes, speeds, and structural characteristics can be successfully detected even under varying lighting conditions and complex environmental backgrounds (Cinarer et al, 2025).In particular, Convolutional Neural Networks (CNNs) and deep learning-based models enable the automatic learning of aircraft appearances, movement patterns, and distinctive visual features.

In recent years, with the development of AI-based image generation and manipulation techniques, distinguishing fake aircraft images from real ones has become a significant research topic. The inability to detect fake or altered images of both aircraft and unmanned aerial vehicles can lead to false alarms or security vulnerabilities in both military and civilian systems(Marzukhi et al, 2021). Therefore, projects utilizing image processing and machine learning algorithms are crucial not only for the detection of aircraft but also for the reliable classification of real and fake images.

In a study, Tüzün Tolga Inan used data from 677 aircraft accidents since September 11, 2011, to classify accidents using machine learning algorithms, namely RFE and PCA methods. Using multinomial logistic regression and artificial neural networks, he found that the five most important features for classifying aircraft damage were latitude, wind speed, wind direction, year, and longitude. He presented data to the literature showing that the number of significant accidents has decreased over time as safety measures have increased (İnan, 2023). Wang et al. classified 2558 aircraft image data points into five different types – fighter jets, helicopters, bombers, early warning aircraft, and passenger aircraft – using the EfficientDet method and machine learning methods based on the Transformer module. They achieved a mean average precision (mAP) of 86.6% with the TransEffiDet method. They have provided a publicly available dataset for aircraft detection to the literature (Wang et al, 2022). Taheri Gorji et al. used electroencephalogram (EEG) data from 10 university aviation students in a live flight environment in a single-engine aircraft to classify pilot workload as low, medium, and high using support vector machine, random forrest, and logistic regression methods. The recursive feature elimination (RFE) algorithm achieved an accuracy of 91.67% (±0.11), a sensitivity of 93.89% (±0.09), a recall of 91.67% (±0.11), and an F-score of 91.22% (±0.12), with an average ROC-AUC value of 0.93 (±0.06). Their contributions to the literature have revealed the potential of using power spectral density (PSD) and logarithmic energy entropy, in conjunction with well-designed machine learning algorithms, to enhance aircraft system design, including flight automation features, for distinguishing low, medium, and high workload periods of EEG and improving aviation safety (Taheri Gorji et al, 2023). Sayed et al. used 1200 radar images in the first dataset, 1600 in the second, and 2800 in the third, obtained from various movements of four different drones via a fixed radar, to classify the drones using CNN1, CNN2, CNN3, SVM, KNN, NB, RF, and DT methods. When mechanical control-based machine learning (MCML) was applied, they achieved accuracy rates above 90% in all machine learning methods. They presented to the literature that the accuracy rate of classification can be increased when machine learning is performed using the MCML method (Sayed et al, 2023). AlDosari et al.

performed drone identification using YOLOV3, YOLOV4, YOLOV5, and Detectronv2 deep learning methods with datasets of various drones. In terms of mAP efficiency, they achieved 96.672% accuracy with the Detectronv2 method in the first dataset and 93.57% with the YOLOV4 method in the second dataset. In the third dataset, YOLOV5 achieved the highest accuracy with 99.49% for D3-a and 95.36% for D3-b. Finally, YOLOV3 achieved the highest accuracy with 99.58% for the last dataset. Their work contributes to the literature by comparing deep learning models for drone detection (AlDosari et al, 2024). Ghazlane et al. used selected image data from a dataset containing 20,000 data points to perform identification processes that classify aircraft as friend or foe using eight different convolutional neural network (CNN) methods. They obtained 98.12% accuracy, 98.184% sensitivity, 98.115% F1 score, and 99.85% area under the curve (AUC) results from the EfficientNetB6 method. Their contribution to the literature suggests that a two-class friend-or-foe identification model can achieve higher identification accuracy (Ghazlane et al, 2024). Zohra et al. performed drone and non-drone classification using the VCC16 method with randomly obtained drone image data from the UAV123 dataset on the ImageNet website. They achieved accuracy rates of 97.6% in the first test, 94.205% in the second test, and 94.570% in the third test using the VCC16 method. They have contributed to the literature by developing a drone detection algorithm using CNN (Zohra et al, 2023). Uddin et al. performed drone detection using Independent Component Analysis (ICA) techniques with SVM and kNN methods to separate mixed signals. They concluded that SVM and kNN methods were insufficient for classifying signals due to evolving conditions over time, and to overcome these limitations, they developed the Time-Varying Drone Detection (TVDDT) technique. The TVDDT method offers a more flexible and effective detection process by considering time-dependent changes in interference coefficients. Their simulation studies have shown that the proposed method exhibits superior performance compared to existing approaches (Uddin et al, 2022). Kassab et al. first attempted to measure the classification success of SVM and RF methods using data derived from filtered and unfiltered video frame sequences in both RGB and infrared (IR) formats. They argued that these traditional machine learning algorithms were unsuitable for classification due to background interference and variability. When they applied SVM and RF machine learning algorithms to a new dataset created using the Non-Maximum Suppression (NMS) method, they achieved 25% better results (Kassab et al, 2024). Khurram Shafiq, in order to facilitate drone detection and enhance system security, integrated a sensor fusion approach into his work, implementing the YOLOV3 method for drone detection. He stated that this work demonstrates that the YOLOV3 model exhibits superior performance in terms of speed and durability compared to other models, achieving a high confidence level of over 95% (Shafiq, 2023). Yoshihashi et al. used datasets containing bird and drone images to perform classifications using their proposed recurrent relational network. In their experiments, they suggested that their proposed method performed better than other object classification models (Yoshihashi et al, 2017). Zheng et al. performed air-to-air UAV detection using 13,271 micro UAV images, each with a resolution of 3840x2160 pixels, employing SSD, RetinaNet, YOLOV3, RefineDet, Faster R-CNN, and Grid R-CNN methods. They achieved an 82.4% success rate with the Grid R-CNN method. Their work contributes to the literature by showing how machine learning algorithms used in air-to-air UAV detection perform in different situations (Zheng et al, 2021).

Based on existing studies in the literature, a significant gap has been identified in the classification of fake and real aircraft images. To address this gap, a comprehensive analysis was conducted using machine learning-based classification methods to distinguish between these images. The following steps were performed in the study:

- A structure was created for aircraft images under six different classes: fake commercial aircraft AI, fake military aircraft AI, fake private aircraft AI, real commercial aircraft, real military aircraft, and real private aircraft, and the images were classified according to these classes.
- The aircraft images in the dataset were processed using ANN, KNN, SVM, and LR machine learning models, and the images were classified.

The rest of the study was planned as follows: The second section describes the dataset used, the methods applied, and the performance metrics in detail. The third section explains the experimental results obtained. The fourth section discusses the results of the study, its contributions, limitations, and potential areas of application.

Methodology

Classification Model Dataset (CMD)

The dataset used in this study contains a total of 300 images in 6 classes. Each class contains 50 images. The classes are fake commercial aircraft (AI), fake military aircraft (AI), fake private aircraft (AI), real commercial aircraft, real military aircraft, and real private aircraft. Figure 1 shows sample images in CMD (*Classification Model Dataset*).

Figure 1. Sample images categorized by class in CMD.

Convolutional Neural Network (CNN)

Convolutional neural networks (CNNs) are deep learning models that have achieved significant success, particularly in image recognition and image processing. Unlike traditional artificial neural networks, they have the ability to directly process spatial and local features in images. In CNN architecture, basic feature extraction is performed through convolutional layers. Pooling layers are used to reduce the size of the feature maps obtained in these layers, decrease computation times, and lower computational costs. The most commonly preferred approach among pooling methods is maximum pooling. This method ensures the preservation of the most prominent feature in each region (Taspinar & Cinar, 2024). In this study, the SqueezeNet architecture, one of the pre-trained CNN models, was used to extract features from images.

SqueezeNet Architecture

SqueezeNet is a CNN architecture designed to achieve high classification accuracy using a smaller number of parameters. The architecture is built on firing modules consisting of a compression (squeeze) layer followed by an expansion layer. The compression layer reduces the number of channels in the input feature maps using 1x1 convolution filters, while the expansion layer enables richer feature representations by using 1x1 and 3x3 convolution filters together. This design approach significantly reduces the number of parameters and computational cost of the model, allowing for efficient performance even in environments with limited hardware and computational resources (DERTLI & KOKLU, 2025).

Artificial Neural Network (ANN)

Artificial neural networks (ANNs) are mathematical systems composed of many interconnected processing units (neurons). A processing unit receives signals from other neurons; it combines and transforms them, producing a numerical result. Generally, processing units roughly correspond to real neurons and are interconnected within a network, thus forming artificial neural networks (Ozkan & Koklu, 2017).

K Nearest Neighbor (KNN)

The KNN algorithm is a classification method within the scope of supervised learning. This algorithm aims to determine which class a new instance belongs to by directly using training data. The basic steps of the KNN algorithm are: first, a new instance to be classified is defined. Then, the feature similarity between each instance in the training dataset and this new instance is calculated, and a distance metric is determined accordingly. Based on the calculated distances, the closest 'k' neighboring instances are selected. In the final stage, the new instance is classified using a majority vote method, taking into account the class labels of these neighbors(Kishore et al, 2022).

The main advantages of the KNN algorithm include its simple structure, ease of implementation, and the fact that it does not require complex modeling or parameter learning during the training phase (Cinar & Taspinar, 2024).

Support Vector Mechines (SVM)

SVM (Support Vector Machine) is a supervised learning method used in classification and regression problems. SVM aims to identify a hyperplane that can most effectively separate data points into their respective classes. This hyperplane is defined by choosing a distance (margin) such that the distance between the closest data points belonging to different classes is maximized, and this approach increases the model's generalization ability and accuracy. Following the training process, the model can make predictions with high performance even on previously unseen or unknown data. Thanks to its ability to work effectively with large and complex datasets, SVM is widely used in many fields such as image recognition and text classification (Isik et al, 2024).

Logistic Regression (LR)

LR, an improved version of the linear regression model, is a statistical method widely used, especially in predicting classification results. While the relationship between variables in linear regression is drawn with a straight line, in logistic regression, a curve close to each data point is drawn in datasets with many straight line features. In this way, the model calculates the probability of an observation belonging to a particular class by generating values between 0 and 1 (Yasin et al, 2025). For logistic regression to yield reliable results, the suitability of the data to the model, the relationships between variables, and the sample structure are of great importance.

Confusion Matrix and Performance Matrix

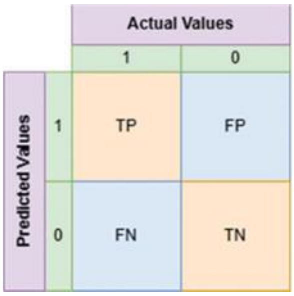

Figure 2. Confusion Martix for binary classes.

ACTUAL \ PREDICT	Fake Commerica Aircraft AI	Fake Military Aircraft AI	Fake Private Aircraft AI	Real Commerica Aircraft	Real Military Aircraft	Real Private Aircraft
Fake Commerical	TP	FN	FN	FN	FN	FN
Fake Military Aircraft AI	FP	TP	FN	FN	FN	FN
Fake Private Aircraft AI	FP	FP	TP	FN	FN	FN
Real Commerical	FP	FP	FP	TP	FN	FN
Real Military Aircraft	FP	FP	FP	FP	TP	FN
Real Private Aircraft	FP	FP	FP	FP	FP	TP

Figure 3. Confusion Martix for six classes.

To obtain performance metrics of classification methods, parameters such as accuracy, recall, sensitivity, and F1 metric were calculated. These parameters were obtained from the confusion matrix. The mathematical expression of each metric is defined below (Aydın & Kızılay, 2022).

Table 1. Mathematical expression of performance metrics (Kilci & Koklu, 2025).

Metrics	Formula		Explanation of Symbols
Avaradged Accuracy	$\frac{\sum_{i=1}^{1}\frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{l}$	(1)	tp_i= True Positives for class i
Avaradged Precision	$\frac{\sum_{i=1}^{1}\frac{tp_i}{tp_i + fp_i}}{l}$	(2)	tn_i= True Negatives for class i
Avaradged Recall	$\frac{\sum_{i=1}^{1}\frac{tp_i}{tp_i + fn_i}}{l}$	(3)	fp_i = False Positives for class i
Avaradged F1-Score	$\frac{2 * \frac{\sum_{i=1}^{1}\frac{tp_i}{tp_i + fp_i}}{l} * \frac{\sum_{i=1}^{1}\frac{tp_i}{tp_i + fn_i}}{l}}{\frac{\sum_{i=1}^{1}\frac{tp_i}{tp_i + fp_i}}{l} + \frac{\sum_{i=1}^{1}\frac{tp_i}{tp_i + fn_i}}{l}}$	(4)	fn_i = False Negatives for class i l = Total number of classes

Accuracy is represented by the terms True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN), Accuracy (A), Recall (R), and Precision (P). Using accuracy as the sole criterion for measuring classification performance is insufficient, especially when using an unbalanced dataset. Therefore, class-based performance measurements are obtained using other criteria (Aydın & Kızılay, 2022).

Cross Validation

Cross-validation is a statistical method used to objectively and reliably evaluate the performance of classification models. In this method, the dataset is divided into equally sized parts based on a parameter called 'k'. In each iteration, a 1/k portion of the dataset is used as the test set, and a k-1 portion is used as the training set. This process is repeated until every subset in the dataset serves as a test set. A total of 'k' iterations are performed. In the final stage, the overall classification performance of the model is calculated by taking the arithmetic mean of the test results obtained in each iteration (Koklu et al, 2021). In this study, the k value was set to 10, and a 10-fold cross-validation method was applied.

Result and Discussion

In this study, 300 aircraft images were analyzed using image processing and machine learning methods. The aim was to identify AI-generated aircraft images and real aircraft images using CMD. Machine learning methods such as ANN, KNN, SVM, and LR were used for classification. A computer with a Core i7™ 12700H 2.3 GHz processor, an NVIDIA GeForce RTX 3060Ti laptop CPU graphics card, and 32 GB of RAM was used. The Python programming language was employed. For the ANN method, the following parameters were used: 100 hidden layers, ReLu activation function, regularization 0.0001, and number of iterations 200. For the KNN method, the following parameters were used: number of neighbors 5, metric euclidean, and weight uniform. For the SVM method, the following parameters were used: cost 1.00, regression loss epsilon 0.1, kernel RBF, and iteration 100. For the LR method, the following parameters were used: regularization type Ridge L2 and Strength C=1. Training and testing procedures were performed using the cross-validation method. In this method, the k value was set to 10. The procedures performed in the study are shown in Figure 4.

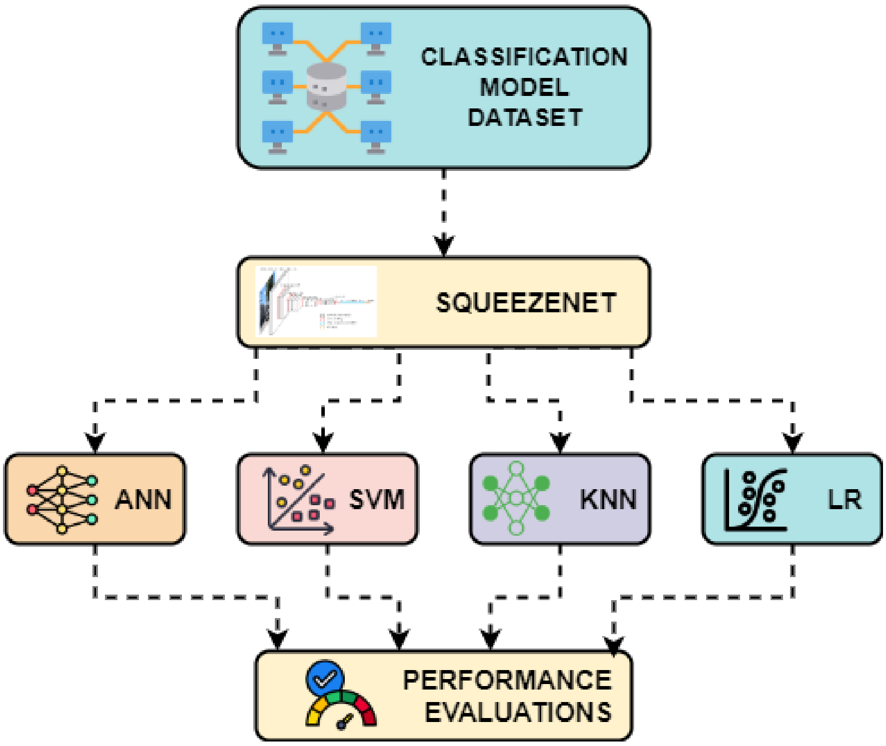

Figure 4. Flowchart of the procedures performed in the study.

Fifty features extracted from each image using the SqueezeNet model were given as an introduction to machine learning methods, and their training was carried out. The complexity matrices obtained from the machine learning methods are given in order. The complexity matrix obtained from the ANN method is shown in Figure 5.

	Fake Commerical Aircraft AI	Fake Military Aircraft AI	Fake Private Aircraft AI	Real Commerical Aircraft	Real Military Aircraft	Real Private Aircraft
Fake Commerical Aircraft AI	40	3	5	1	0	1
Fake Military Aircraft AI	1	43	4	0	2	0
Fake Private Aircraft AI	5	3	42	0	0	0
Real Commerical Aircraft	3	1	1	38	3	4
Real Military Aircraft	0	2	1	3	39	5
Real Private Aircraft	0	1	2	4	5	38

Figure 5. Confusion matrix of ANN model.

According to Figure 5, the ANN method correctly classified 40 data points belonging to the fake commercial aircraft AI class. It classified 3 data points belonging to the fake commercial aircraft AI class as fake military aircraft AI. It classified 5 data points belonging to the fake commercial aircraft AI class as fake private aircraft AI. It classified 1 data point belonging to the fake commercial aircraft AI class as real commercial aircraft. In total, it incorrectly classified 10 data points belonging to the fake commercial aircraft AI class. It correctly classified 43 data points belonging to the fake military aircraft AI class. It classified 1 data point belonging to the fake military aircraft AI class as fake commercial aircraft AI. It classified 4 data points belonging to the fake military aircraft AI class as fake private aircraft AI. It classified 2 data points belonging to the fake military aircraft AI class as real military aircraft. In total, it incorrectly classified 7 data points belonging to the fake military aircraft AI class. The AI correctly classified 42 data points belonging to the "fake private aircraft" category. It classified 5 data points belonging to the "fake private aircraft" category as "fake commercial aircraft" AI. It classified 3 data points belonging to the "fake private aircraft" category as "fake military aircraft" AI. In total, 8 data points belonging to the "fake private aircraft" category were incorrectly classified. The AI correctly classified 38 data points belonging to the "real commercial aircraft" category. It classified 3 data points belonging to the "real commercial aircraft" category as "fake commercial aircraft" AI. It classified 1 data point belonging to the "real commercial aircraft" category as "fake military aircraft" AI. It classified 1 data point belonging to the "real commercial aircraft" category as "fake private aircraft" AI. It classified 3 data points belonging to the "real commercial aircraft" category as "real military aircraft". It classified 4 data points belonging to the "real commercial aircraft" category as "real private aircraft". In total, 12 data points belonging to the "real commercial aircraft" category were incorrectly classified. The system correctly classified 39 data points belonging to the "real military aircraft" category. It classified 2 data points belonging to the "real military aircraft" category as "fake military aircraft AI". It classified 1 data point belonging to the "real military aircraft" category as "fake private aircraft AI". It classified 3 data points belonging to the "real military aircraft" category as "real commercial aircraft". It classified 5 data points belonging to the "real military aircraft" category as "real private aircraft". In total, 11 data points belonging to the "real commercial aircraft" category were incorrectly classified. The system correctly classified 38 data points belonging to the "real private aircraft" category. It classified 1 data point belonging to the "real private aircraft" category as "fake military aircraft AI". It classified 2 data points belonging to the "real private aircraft" category as "fake private aircraft AI". It classified 4 data points belonging to the "real private aircraft" category as "real commercial aircraft". It classified 5 data points belonging to the "real private aircraft" category as "real military aircraft". In total, 12 data points belonging to the "real commercial aircraft" category were incorrectly classified. Figure 6 shows the complexity matrix for the SVM method.

	Fake Commerical Aircraft AI	Fake Military Aircraft AI	Fake Private Aircraft AI	Real Commerical Aircraft	Real Military Aircraft	Real Private Aircraft
Fake Commerical Aircraft AI	36	3	7	3	0	1
Fake Military Aircraft AI	3	43	0	1	3	0
Fake Private Aircraft AI	5	0	45	0	0	0
Real Commerical Aircraft	3	0	1	37	5	4
Real Military Aircraft	0	0	1	2	45	2
Real Private Aircraft	0	0	1	3	11	35

Figure 6. Confusion matrix of SVM model.

According to Figure 6, the SVM method correctly classified 36 data points belonging to the fake commercial aircraft AI class. It classified 3 data points belonging to the fake commercial aircraft AI class as fake military aircraft AI. It classified 7 data points belonging to the fake commercial aircraft AI class as fake private aircraft AI. It classified 3 data points belonging to the fake commercial aircraft AI class as real commercial aircraft. It classified 1 data point belonging to the fake commercial aircraft AI class as real private aircraft. In total, it incorrectly classified 14 data points belonging to the fake commercial aircraft AI class. It correctly classified 43 data points belonging to the fake military aircraft AI class. It classified 3 data points belonging to the fake military aircraft AI class as fake commercial aircraft AI. It classified 1 data point belonging to the fake military aircraft AI class as real commercial aircraft. It classified 3 data points belonging to the fake military aircraft AI class as real military aircraft. In total, 7 data points belonging to the fake military aircraft AI category were misclassified. 45 data points belonging to the fake private aircraft AI category were correctly classified. 5 data points belonging to the fake private aircraft AI category were classified as fake commercial aircraft AI. In total, 5 data points belonging to the fake private aircraft AI category were misclassified. 37 data points belonging to the real commercial aircraft category were correctly classified. 3 data points belonging to the real commercial aircraft category were classified as fake commercial aircraft AI. 1 data point belonging to the real commercial aircraft category was classified as fake private aircraft AI. 5 data points belonging to the real commercial aircraft category were classified as real military aircraft. 4 data points belonging to the real commercial aircraft category were classified as real private aircraft. In total, 13 data points belonging to the real commercial aircraft category were misclassified. 45 data points belonging to the real military aircraft category were correctly classified. 1 data point belonging to the real military aircraft category was classified as fake private aircraft AI. The analysis classified 2 data points belonging to the "real military aircraft" category as "real commercial aircraft". It also classified 2 data points belonging to the "real military aircraft" category as "real private aircraft". In total, 5 data points belonging to the "real commercial aircraft" category were misclassified. It correctly classified 35 data points belonging to the "real private aircraft" category. It classified 1 data point belonging to the "real private aircraft" category as "fake private aircraft AI". It classified 3 data points belonging to the "real private aircraft" category as "real commercial aircraft". It classified 11 data points belonging to the "real private aircraft" category as "real military aircraft". In total, 15 data points belonging to the "real commercial aircraft" category were misclassified. Figure 7 shows the complexity matrix of the KNN method.

	Fake Commerical Aircraft AI	Fake Military Aircraft AI	Fake Private Aircraft AI	Real Commerical Aircraft	Real Military Aircraft	Real Private Aircraft
Fake Commerical Aircraft AI	30	6	14	0	0	0
Fake Military Aircraft AI	2	44	2	1	1	0
Fake Private Aircraft AI	6	1	43	0	0	0
Real Commerical Aircraft	4	1	8	35	0	2
Real Military Aircraft	1	8	5	7	21	8
Real Private Aircraft	1	1	8	11	4	25

Figure 7. Confusion matrix of KNN model.

According to Figure 7, the KNN method correctly classified 30 data points belonging to the fake commercial aircraft AI class. It classified 6 data points belonging to the fake commercial aircraft AI class as fake military aircraft AI. It classified 14 data points belonging to the fake commercial aircraft AI class as fake private aircraft AI. In total, it incorrectly classified 20 data points belonging to the fake commercial aircraft AI class. It correctly classified 44 data points belonging to the fake military aircraft AI class. It classified 2 data points belonging to the fake military aircraft AI class as fake commercial aircraft AI. It classified 2 data points belonging to the fake military aircraft AI class as fake private aircraft AI. It classified 1 data point belonging to the fake military aircraft AI class as real commercial aircraft. It classified 1 data point belonging to the fake military aircraft AI class as real military aircraft. In total, it incorrectly classified 6 data points belonging to the fake military aircraft AI class. The AI correctly classified 43 data points belonging to the "fake private aircraft" category. It classified 6 data points belonging to the "fake private aircraft" category as "fake commercial aircraft" AI. It classified 1 data point belonging to the "fake private aircraft" category as "fake military aircraft" AI. In total, 7 data points belonging to the "fake private aircraft" category were incorrectly classified. The AI correctly classified 35 data points belonging to the "real commercial aircraft" category. It classified 4 data points belonging to the "real commercial aircraft" category as "fake commercial aircraft" AI. It classified 1 data point belonging to the "real commercial aircraft" category as "fake military aircraft" AI. It classified 8 data points belonging to the "real commercial aircraft" category as "fake private aircraft" AI. It classified 2 data points belonging to the "real commercial aircraft" category as "real private aircraft". In total, 15 data points belonging to the "real commercial aircraft" category were incorrectly classified. The AI correctly classified 21 data points belonging to the "real military aircraft" category. The AI misclassified 1 data point belonging to the "real military aircraft" category as "fake commercial aircraft". It misclassified 8 data points belonging to the "real military aircraft" category as "fake military aircraft". It misclassified 5 data points belonging to the "real military aircraft" category as "fake private aircraft". It misclassified 7 data points belonging to the "real military aircraft" category as "real commercial aircraft". It misclassified 8 data points belonging to the "real military aircraft" category as "real private aircraft". In total, 29 data points belonging to the "real commercial aircraft" category were misclassified. It correctly classified 25 data points belonging to the "real private aircraft" category. It misclassified 1 data point belonging to the "real private aircraft" category as "fake commercial aircraft". It misclassified 1 data point belonging to the "real private aircraft" category as "fake military aircraft". It misclassified 8 data points belonging to the "real private aircraft" category as "fake private aircraft". It misclassified 11 data points belonging to the "real private aircraft" category as "real commercial aircraft". It misclassified 4 data points belonging to the "real private aircraft" category as "real military aircraft". In total, 25 data points belonging to the real commercial aircraft class were misclassified. Figure 8 shows the complexity matrix of the LR method.

	Fake Commerical Aircraft AI	Fake Military Aircraft AI	Fake Private Aircraft AI	Real Commerical Aircraft	Real Military Aircraft	Real Private Aircraft
Fake Commerical Aircraft AI	41	2	3	3	0	1
Fake Military Aircraft AI	1	46	2	1	0	0
Fake Private Aircraft AI	3	3	43	0	0	1
Real Commerical Aircraft	4	0	0	42	0	4
Real Military Aircraft	0	0	0	5	39	6
Real Private Aircraft	0	1	1	4	4	40

Figure 8. Confusion matrix of LR model.

According to Figure 8, the LR method correctly classified 41 data points belonging to the fake commercial aircraft AI class. It classified 2 data points belonging to the fake commercial aircraft AI class as fake military aircraft AI. It classified 3 data points belonging to the fake commercial aircraft AI class as fake private aircraft AI. It classified 3 data points belonging to the fake commercial aircraft AI class as real commercial aircraft. It classified 1 data point belonging to the fake commercial aircraft AI class as real private aircraft. In total, it incorrectly classified 9 data points belonging to the fake commercial aircraft AI class. It correctly classified 46 data points belonging to the fake military aircraft AI class. It classified 1 data point belonging to the fake military aircraft AI class as fake commercial aircraft AI. It classified 2 data points belonging to the fake military aircraft AI class as fake private aircraft AI. It classified 1 data point belonging to the fake military aircraft AI class as real commercial aircraft. In total, 4 data points belonging to the fake military aircraft AI category were misclassified. 43 data points belonging to the fake private aircraft AI category were correctly classified. 3 data points belonging to the fake private aircraft AI category were classified as fake commercial aircraft AI. 3 data points belonging to the fake private aircraft AI category were classified as fake military aircraft AI. 1 data point belonging to the fake private aircraft AI category was classified as real private aircraft. In total, 7 data points belonging to the fake private aircraft AI category were misclassified. 42 data points belonging to the real commercial aircraft category were correctly classified. 4 data points belonging to the real commercial aircraft category were classified as fake commercial aircraft AI. 4 data points belonging to the real commercial aircraft category were classified as real private aircraft. In total, 8 data points belonging to the real commercial aircraft category were misclassified. 39 data points belonging to the real military aircraft category were correctly classified. 5 data points belonging to the real military aircraft category were classified as real commercial aircraft. It misclassified 6 data points belonging to the "real military aircraft" category as "real private aircraft." In total, 11 data points belonging to the "real commercial aircraft" category were misclassified. It correctly classified 40 data points belonging to the "real private aircraft" category. It misclassified 1 data point belonging to the "real private aircraft" category as "fake military aircraft AI." It misclassified 1 data point belonging to the "real private aircraft" category as "fake private aircraft AI." It misclassified 4 data points belonging to the "real private aircraft" category as "real commercial aircraft." It misclassified 4 data points belonging to the "real private aircraft" category as "real military aircraft." In total, 10 data points belonging to the "real commercial aircraft" category were misclassified.

Performance metrics for each method were calculated using data from complex matrices obtained from the methods. These metrics are shown in Table 2.

Table 2. Performance metrics of all models

	Accuracy	F1 Score	Precision	Recall
ANN	0,966	0,800	0,801	0,800
KNN	0,904	0,651	0,685	0,660
SVM	0,967	0,803	0,810	0,803
LR	0,965	0,837	0,840	0,837

According to Table 2, the most successful model is SVM with a classification success rate of 0.967. The model with the lowest classification success rate is KNN with 0.904. The classification success rate of the ANN model is 0.966, and the classification success rate of the LR model is 0.965. When other performance metrics are examined, it is seen that the F1 Score, Precision, and Recall metric values also show parallelism with the classification success rates of the models. When the classification success rates and F1 Score, precision, and recall metrics are examined, it can be said that all models show a success rate of 65.1% and above. Figure 9 shows the ROC curves for all models.

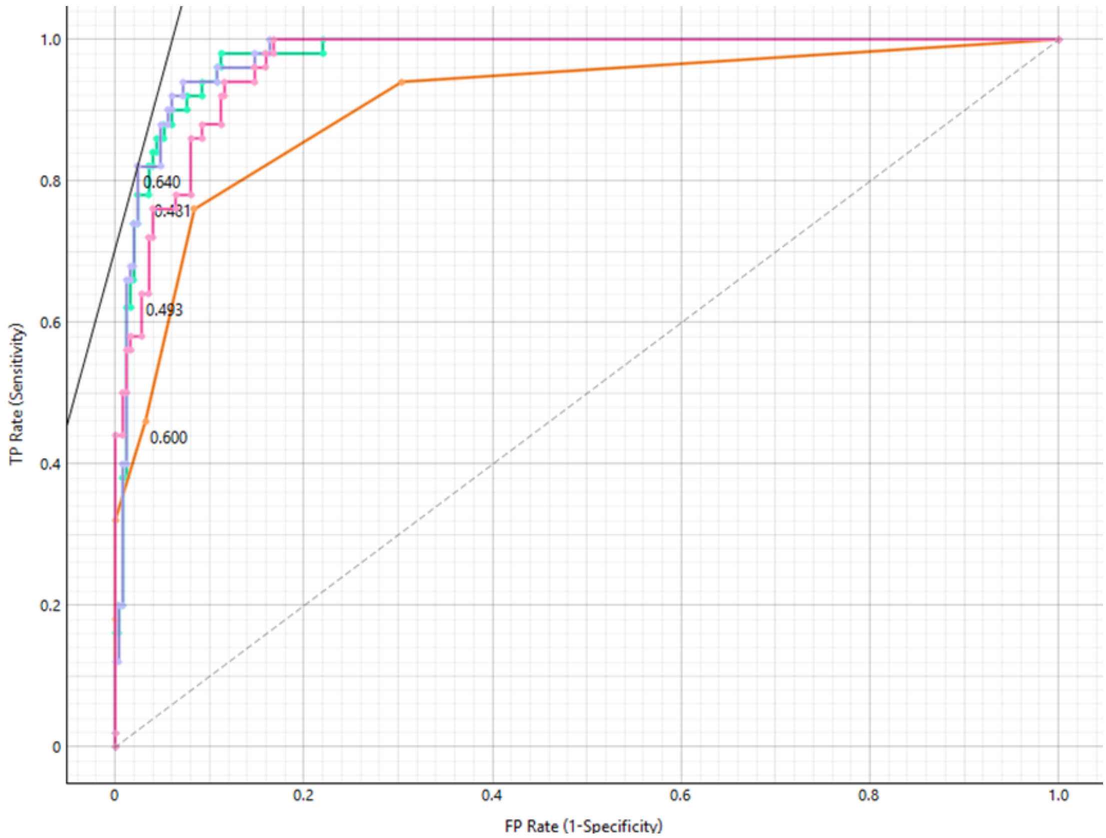

Figure 9. ROC of all models.

Figure 9 shows a correlation between classification success rates and the learning levels of the models. The best-learning model is SVM, while the least-learning model is KNN. In conclusion, it can be said that the classification success of the models is determined by their learning rates.

Conclusion

This study considers the importance of aircraft image classification for the recognition and identification of aircraft. Classification methods for distinguishing between AI-generated images and real images are discussed. The widespread use of AI-generated images today makes it difficult to differentiate between real and fake images. Based on these issues, 300 images in the dataset were classified into 6 different classes using machine learning algorithms. The highest classification success rate was obtained from the SVM model with 96.7%. The lowest classification success rate was obtained from the KNN model with 90.7%.

The results indicate that machine learning algorithms can be used for classifying aircraft images. However, limiting the dataset to 300 images may lead to performance degradation in the models, as different image types may be encountered in real operational conditions. Future studies using larger and more diverse datasets, increasing image data, and evaluating different classification models could lead to higher and more generalizable success rates in aircraft classification.

The classification models proposed in this study are suitable for use in military testing processes, training and education applications where aircraft detection is necessary, and in scenarios where distinguishing between real and fake aircraft is critical. This will enable the rapid and reliable identification of whether an encountered aircraft image is real or generated with the aid of artificial intelligence.

References

AlDosari, K., Osman, A., Elharrouss, O., Al-Maadeed, S., & Chaari, M. Z. (2024). Drone-type-Set: Drone types detection benchmark for drone detection and tracking. 2024 International Conference on Intelligent Systems and Computer Vision (ISCV),

Aydın, İ., & Kızılay, E. (2022). Development of a new light-weight convolutional neural network for acoustic-based amateur drone detection. *Applied Acoustics*, *193*, 108773.

Cinar, I., & Taspinar, Y. S. (2024). Detection of machine failures with machine learning methods. Proceedings of International Conference on Intelligent Systems and New Applications,

Cinarer, G., Taspinar, Y. S., & Zeybek, M. (2025). Detection of Aerial Vehicles Using Satellite Imagery: Comparative Analysis of U-Net Segmentation Model and YOLO Object Detection Model. *International Journal of Aerospace Engineering*, *2025*(1), 5599522.

Classification Model Dataset. https://www.kaggle.com/datasets/samanthaclark55/fake-aircraft-images

DERTLI, B., & KOKLU, M. (2025). Classification Of Hair Types with Deep Learning Methods.

Ghazlane, Y., Gmira, M., & Medromi, H. (2024). Development of a vision-based anti-drone identification friend or foe model to recognize birds and drones using deep learning. *Applied Artificial Intelligence*, *38*(1), 2318672.

İnan, T. T. (2023). Aircraft damage classification by using machine learning methods. *International Journal of Aviation, Aeronautics, and Aerospace*, *10*(2), 4.

Isik, M., Ozulku, B., Kursun, R., Taspinar, Y. S., Cinar, I., Yasin, E. T., & Koklu, M. (2024). Automated classification of hand-woven and machine-woven carpets based on morphological features using machine learning algorithms. *The Journal of The Textile Institute*, *115*(12), 2650-2659.

Kassab, M., Zitar, R. A., Barbaresco, F., & Seghrouchni, A. E. F. (2024). Drone detection with improved precision in traditional machine learning and less complexity in single-shot detectors. *IEEE Transactions on Aerospace and Electronic Systems*, *60*(4), 3847-3859.

Kilci, O., & Koklu, M. (2025). Machine learning-based detection of solar panel surface defects using deep features from InceptionV3. 4th International Conference on Trends in Advanced Research Konya, Türkiye,

Kishore, B., Yasar, A., Taspinar, Y. S., Kursun, R., Cinar, I., Shankar, V. G., Koklu, M., & Ofori, I. (2022). Computer-aided multiclass classification of corn from corn images integrating deep feature extraction. *Computational Intelligence and Neuroscience*, *2022*(1), 2062944.

Koklu, M., Cinar, I., & Taspinar, Y. S. (2021). Classification of rice varieties with deep learning methods. *Computers and electronics in agriculture*, *187*, 106285.

Marzukhi, S., Awang, N., Alsagoff, S. N., & Mohamed, H. (2021). RapidMiner and machine learning techniques for classifying aircraft data. Journal of Physics: Conference Series,

Ozkan, I. A., & Koklu, M. (2017). Skin lesion classification using machine learning algorithms. *International Journal of Intelligent Systems and Applications in Engineering*, *5*(4), 285-289.

Sayed, A. N., Ramahi, O. M., & Shaker, G. (2023). Machine learning for UAV classification employing mechanical control information. *IEEE Transactions on Aerospace and Electronic Systems*, *60*(1), 68-81.

Shafiq, K. (2023). *Drone Detection and Classification using Machine Learning* Université d'Ottawa/University of Ottawa].

Taheri Gorji, H., Wilson, N., VanBree, J., Hoffmann, B., Petros, T., & Tavakolian, K. (2023). Using machine learning methods and EEG to discriminate aircraft pilot cognitive workload during flight. *Scientific Reports*, *13*(1), 2507.

Taspinar, Y. S., & Cinar, I. (2024). Distinguishing between AI images and real images with hybrid image classification methods. 2024 13th Mediterranean Conference on Embedded Computing (MECO),

Uddin, Z., Qamar, A., Alharbi, A. G., Orakzai, F. A., & Ahmad, A. (2022). Detection of multiple drones in a time-varying scenario using acoustic signals. *Sustainability*, *14*(7), 4041.

Wang, Y., Wang, T., Zhou, X., Cai, W., Liu, R., Huang, M., Jing, T., Lin, M., He, H., & Wang, W. (2022). TransEffiDet: aircraft detection and classification in aerial images based on EfficientDet and transformer. *Computational Intelligence and Neuroscience*, *2022*(1), 2262549.

Yasin, E. T., Ropelewska, E., Kursun, R., Cinar, I., Taspinar, Y. S., Yasar, A., Mirjalili, S., & Koklu, M. (2025). Optimized feature selection using gray wolf and particle swarm algorithms for corn seed image classification. *Journal of Food Composition and Analysis*, *145*, 107738.

Yoshihashi, R., Trinh, T. T., Kawakami, R., You, S., Iida, M., & Naemura, T. (2017). Differentiating objects by motion: Joint detection and tracking of small flying objects. *arXiv preprint arXiv:1709.04666*.

Zheng, Y., Chen, Z., Lv, D., Li, Z., Lan, Z., & Zhao, S. (2021). Air-to-air visual detection of micro-uavs: An experimental evaluation of deep learning. *IEEE Robotics and automation letters*, *6*(2), 1020-1027.

Zohra, S. F., Nawal, C., & El Houda, D. N. (2023). Optical Detection of UAVs Using Deep Convolutional Neural Networks. 2023 2nd International Conference on Electronics, Energy and Measurement (IC2EM),