 DOI: 10.58190/icat.2023.37

PROCEEDINGS OF

INTERNATIONAL CONFERENCE ON ADVANCED TECHNOLOGIES

https://proceedings.icatsconf.org/

11th International Conference on Advanced Technologies (ICAT'23), Istanbul-Turkiye, August 17-19, 2023.

Classification of Unmanned Aerial Vehicle and Bird Images Using Deep Transfer Learning Methods

Ahmet ÖZDEMİR[1], İlker Ali ÖZKAN[2]

[1]*Department of Defense Technologies , Selcuk University, Konya, Turkey*
228268001001@lisansustu.selcuk.edu.tr, ORCID: 0000-0002-9905-8418

[2] *Department of Computer Engineering, Selcuk University, Konya, Turkey*
ilkerozkan@selcuk.edu.tr, , ORCID: 0000-0002-5715-1040

***Abstract*— The increasing accessibility and affordability of unmanned aerial vehicles (UAVs), commonly known as drones, have led to the emergence of malicious users. In precaution to this perceived threat, various anti-UAV systems are being developed, including electro-optical systems utilizing cameras. It is possible to detect UAVs from images using various machine learning methods. However, the similarity between UAVs and birds poses a challenge, as birds can be mistakenly identified as UAVs, leading to false alarms in a security system. In order to avoid this problem, this study provided the classification of birds and unmanned aerial vehicles over images using deep learning methods. In this study, a data set consisting of 400 birds and 428 UAV images was used. The data were divided into 70% for training, 30% for testing and validation purposes. Three different deep learning models, based on DenseNet, VGG16, and VGG19 architectures, were trained using transfer learning techniques, and their performances were compared. Experimental results on the test data showed an accuracy of 94.64% with the DenseNet model, 89.67% with the VGG16 model, and 90.67% with the VGG19 model.**

***Keywords*— Unmanned Aerial Vehicles, DenseNet, VGG16, VGG19, Deep Transfer Learning**

I. INTRODUCTION

Due to technological advancements, the usage of unmanned aerial vehicles (UAVs), also known as drones, has rapidly increased. While initially predominantly used in military applications, their cost reduction has led to their proliferation in civilian sectors as well [1, 2]. UAVs are now employed in various fields such as agriculture, mining, construction, natural disaster monitoring, meteorology, archaeology, law enforcement, logistics, hobbies and sports, communications, forensic applications, and military operations [3]. They offer advantages such as lightweight design, portability, low cost, high maneuverability, and low energy consumption. However, these advantages also paved the way for unmanned aerial vehicles to be used in malicious activities and brought some security problems with it. Presently, different terrorist organizations utilize UAVs to gather intelligence through aerial surveillance or transform them into weapons by attaching various explosive materials. This poses a significant threat to both states and civilian populations [4]. Anti-UAV systems have been developed to prevent and mitigate such harmful activities. These systems can generally be categorized into two types: prevention, neutralization systems, and detection, identification, tracking systems. Various technologies are employed for detection, identification, and tracking systems, including radar, radio frequency, electro-optical, infrared, acoustic, and multi-sensor systems. Radar systems are primarily designed to detect large aircraft flying at high altitudes. While radio frequency systems can yield successful results, they are ineffective against autonomous UAVs and can be susceptible to electromagnetic interference. Acoustic systems rely on detecting UAVs based on the sound generated by their propellers and are typically long-range. However, differentiating propeller noise in noisy environments can be challenging. Electro-optical systems aim to detect UAVs through image analysis. However, their effectiveness is limited by the restricted field of view at long ranges. However, if the area to be scanned is limited and in close proximity, electro-optical systems offer a cost-effective solution to achieve effective results [5-7]. In the literature, various studies have been proposed to detect and classify UAVs using image analysis in order to mitigate potential threats.

Muhammad Saqib et al. conducted a study based on convolutional neural networks to compare the success of different deep learning methods on drone detection in 2017. Due to the small dataset they used, they resorted to the transfer learning method over ImageNet. They used ZF, VGG16 and VGG_M_1024 architectures. They used the models they trained together with Faster R-CNN. They observed the best result in the VGG16 architecture with a mAP score of 0.66 [8-12].

In 2020, Angelo Coluccia et al. conducted a study as part of the "Drone vs. Bird Detection Challenge," aiming to evaluate different deep learning-based approaches for the detection and discrimination of unmanned aerial vehicles (UAVs) from flying birds. According to their findings, the most significant challenge and error in the proposed solutions occurred when distinguishing birds from UAVs at long distances [13].

Subroto Singha and Burchan Aydin conducted a study aiming to automatically detect drones to mitigate the potential dangers associated with malicious drone usage. They employed the YOLOv4 model [14] for drone detection. The study achieved a mean average precision (mAP) of 74.36%, precision of 0.95, recall of 0.68, and an F1-score of 0.79 [15].

Subroto Singha and Burchan Aydin conducted a study aiming to automatically detect drones to mitigate the potential dangers associated with malicious drone usage. They employed the YOLOv4 model [14] for drone detection. The study achieved a mean average precision (mAP) of 74.36%, precision of 0.95, recall of 0.68, and an F1-score of 0.79 [15].

In 2019, Hyun Min Oh et al. conducted a study aiming to classify drones and birds using machine learning methods for anti-drone systems. They evaluated the results using various convolutional neural network (CNN) architectures, including AlexNet [16], GoogleNet [17], Inception-v3 [18], VGG16 [10], ResNet-18, ResNet-50 [19], and SqueezeNet [20]. They achieved accuracy scores of 98.53% with AlexNet, 97.84% with VGG16 [10], 97.36% with ResNet18, 96.8% with ResNet50, 95.85% with SqueezeNet, 91.57% with GoogleNet, and 91.01% with Inception-v3 [21].

Eren Unlu et al. conducted a study aiming to autonomously detect drones to counter their potential malicious uses. For this purpose, they utilized a fixed wide-angle camera and a rotating tower with a narrow-angle camera. They performed drone detection using an approach based on the YOLOv3 [22] architecture, utilizing the images captured by the cameras [23].

In 2020, Fatemeh Mahdavi et al. collected a dataset consisting of a total of 712 images of birds and drones with the aim of drone detection. Using this dataset, they trained three different machine learning models: CNN (Convolutional Neural Network), SVM (Support Vector Machines), and KNN (k-Nearest Neighbors). They compared the results obtained from these models. According to their findings, the CNN model achieved an accuracy score of 93%, SVM scored 88%, and KNN scored 80% [24].

Hamid R. Alsanad et al. conducted a study to detect drones in order to prevent malicious drone usage. They proposed a new algorithm as a solution to the unreliable methods typically employed for drone detection, given the small size of drones. They improved upon YOLOv3 by creating a CNN model that reduced the number of parameters and decreased computational complexity by using Darknet53 as the starting point. The algorithm they developed achieved an accuracy score of 95.6% [25].

In 2020, Dinesh Kumar Behera and Arockia Bazil Raj conducted a study aiming to detect and classify drones using deep learning methods. They trained a deep learning model based on the YOLOv3 architecture using a dataset consisting of over 10,000 images of drones from various types and species. After training, they achieved an accuracy score of over 90% [26].

In 2022, S. Sethu Selvi et al. conducted a study to achieve real-time detection of drones. They used a dataset consisting of 664 drone images and 236 bird images to train models using YOLOv4 and YOLOv5. They obtained an f1 score of 98% with YOLOv4 and 94% with YOLOv5. The YOLOv4 model achieved a detection speed of 54 fps on GPU and 12 fps on CPU, while the YOLOv5 model achieved a detection speed of 77 fps on GPU and 27 fps on CPU [27].

In 2022, Bhagyashri B. Bhagat and colleagues conducted a study to achieve drone detection using both moving and fixed cameras. They used Focus Measure Operators for feature extraction from the images and performed feature ranking. They trained a random forest classifier using a dataset consisting of bird and drone videos captured by both moving and stationary cameras. In the classification tasks performed with different class numbers based on whether the images contained drones, birds, or both, they achieved accuracy scores above 92% and sensitivity scores above 95% [28].

In 2022, Bhagyashri B. Bhagat and colleagues conducted a study to achieve drone detection using both moving and fixed cameras. They used Focus Measure Operators for feature extraction from the images and performed feature ranking. They trained a random forest classifier using a dataset consisting of bird and drone videos captured by both moving and stationary cameras. In the classification tasks performed with different class numbers based on whether the images contained drones, birds, or both, they achieved accuracy scores above 92% and sensitivity scores above 95% [28].

The problem of detecting and classifying UAVs through camera images in the literature remains up-to-date and studies in this area continue. Details about the existing studies that have been done before are given in Table 1. The table consists of the method used in the relevant studies, the score obtained, the data set used and the reference of the study.

The studies conducted demonstrate that unmanned aerial vehicles (UAVs) can be detected through image analysis, leading to successful outcomes in preventing their malicious use. Additionally, one common challenge encountered in these studies is the similarity between small UAVs and birds, which can result in misclassifying birds as UAVs. It is seen that different deep learning architectures are used to overcome this issue by classifying bird and unmanned aerial vehicles. Some commonly used architectures for this purpose include GoogleNet, AlexNet, Inception, VGG, ResNet, SqueezeNet, and their different versions.

In this study, unlike the literature studies, a DenseNet (Densely Connected Convolutional Networks) architecture is used to train a new model through transfer learning. To compare the model performance with other studies in the literature, two additional models are trained using VGG16 and VGG19 architectures. The trained models are used to classify bird and UAV images, and the results obtained from each model are evaluated.

In Section 2, the collected dataset for the study and the trained deep learning models are described in detail. Section 3 presents the experimental results, and Section 4 evaluates the obtained results.

TABLE 1. SUMMARY OF UAV DETECTION & CLASSIFICATION

No	Target	Method	Score	Metric	Dataset	Reference
1	Drone detection	ZF VGG16 VGG_M_1024	%61 %66 %60	mAP	Bird-Vs-Drone dataset (2727 frames)	[11]
2	Drone detection	YOLOv4	%74,36	mAP	Custom (479 bird and 1916 drone images)	[15]
3	Drone - Bird Classification	AlexNet GoogleNet Inception-v3 VGG16 ResNet-18 ResNet-50 squeezenet	%98,53 %91,57 %91,01 %97,84 %97,36 %96,8 %95.85	ACC	Custom (3000 birds, 1500 drone and 2500 background images)	[21]
4	Drone - Bird Classification	CNN SVM KNN	%93 %88 %80	ACC	Custom (total of 712 images consisting of drone and bird images.)	[24]
5	Drone detection	YOLOv3 (restructured)	%95,6	ACC	Custom (5000 drone images)	[25]
6	Drone detection	YOLOv3	%90 +	precision	Custom (10000 drone images)	[26]
7	Drone detection and Drone - Bird Classification	YOLOv4 YOLOv5	%98 %94	F1 score	Custom (664 drone and 236 bird images)	[27]

II. MATERIALS AND METHODS

A. *Dataset*

The study utilized the "Birds vs Drone Dataset" available on the Kaggle platform for deep learning-based image classification [29]. The dataset consists of a total of 828 images, with 400 bird images and 428 drone images. To ensure balanced training data, 28 drone images were removed during the training process. For each class, 280 out of 400 images were used for training, while the remaining 120 were reserved for testing and validation. Some examples of bird images are shown in Figure 1.A, and examples of drone images are shown in Figure 1.B in the paper.

Fig.1.a Sample bird images used in the dataset

Fig.1.b Sample UAV images used in the Data Set

B. *Deep Learning Architectures*

1. *DenseNet architecture:* DenseNet Architecture: The DenseNet architecture was introduced by Gao Huang et al. and is based on convolutional neural networks [30]. This architecture consists of densely connected blocks, where the feature maps from preceding blocks are used as input to generate new features. This allows for more efficient feature extraction. Additionally, the features from previous blocks are concatenated and passed on to subsequent blocks, making the network structure generally deeper compared to other architectures, while using fewer parameters and enabling higher scores. One of the significant advantages of this architecture is its ability to alleviate the vanishing gradient problem [31]. Due to these advantages, in this study, a deep learning model was trained using the DenseNet architecture through transfer learning. Subsequently, model training was also conducted using the VGG16 and VGG19 architectures, and the results were compared. The structure of the DenseNet architecture and the deep learning model trained using DenseNet with transfer learning are shown in Figure2.

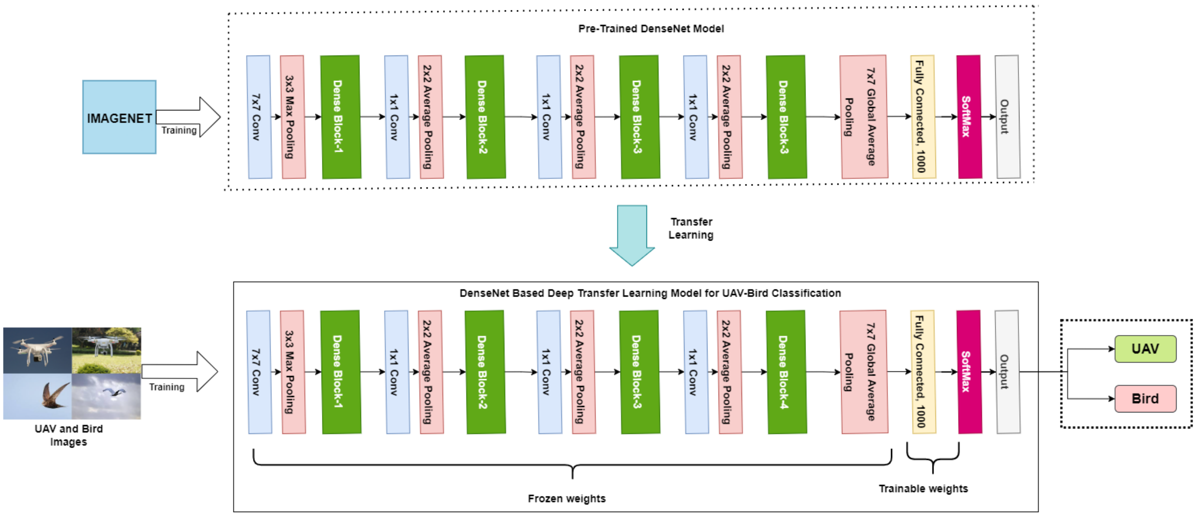

Fig. 2 DenseNet architecture and trained DenseNet-based deep learning model

2. *VGG16 architecture:* The VGG16 architecture was proposed by Karen Simonyan and Andrew Zisserman in 2014. It consists of 13 convolutional layers and 3 fully connected layers. Maximum pooling is applied after consecutive convolution layers. The ReLu activation function replaces negative input values with 0, reducing the complexity of the operation. Thus, the training time is shortened. The main difference of VGG16 compared to previous architectures is the higher number of convolutional layers and smaller kernel size [10]. The structure of the VGG16 architecture and the deep learning model based on VGG16 trained using transfer learning in this study are shown in Figure 3.

3. *VGG19 architecture:* The VGG19 architecture has a similar structure to VGG16. The main difference is that VGG19 has three additional convolutional layers. It consists of 16 convolutional layers and 3 fully connected layers. Therefore, it has a deeper structure than VGG16. This situation causes the training period to be longer than VGG16 [10]. The basic structure of the architecture and the VGG19-based deep learning model trained in this study are as shown in Figure 4.

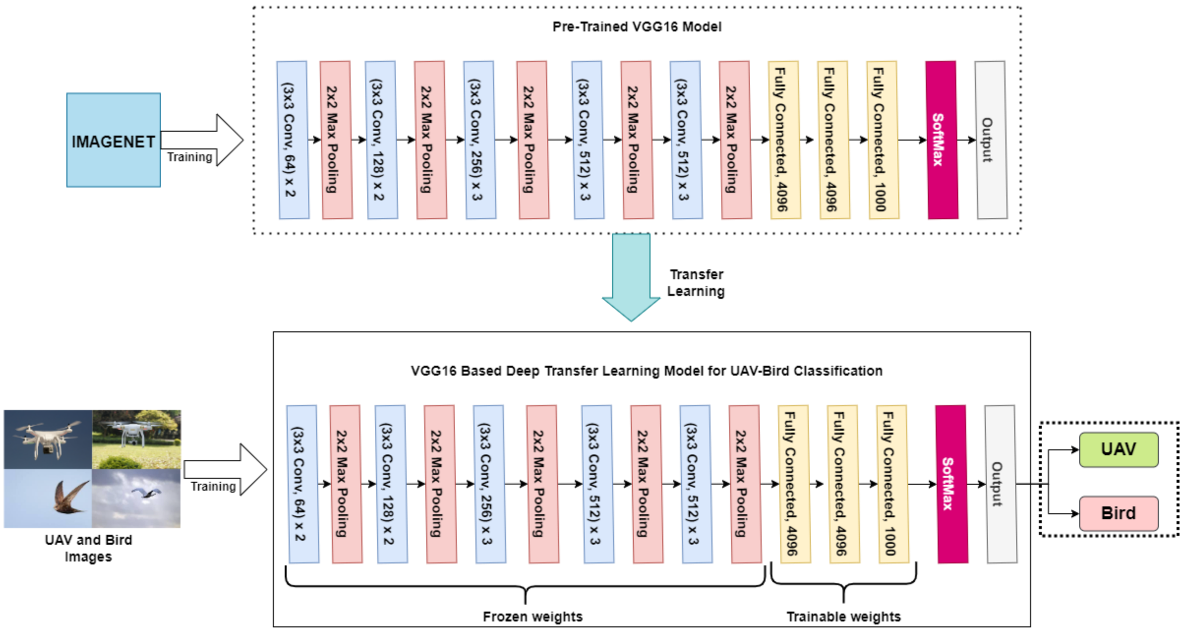

Fig. 3 VGG16 architecture and trained VGG16-based deep learning model

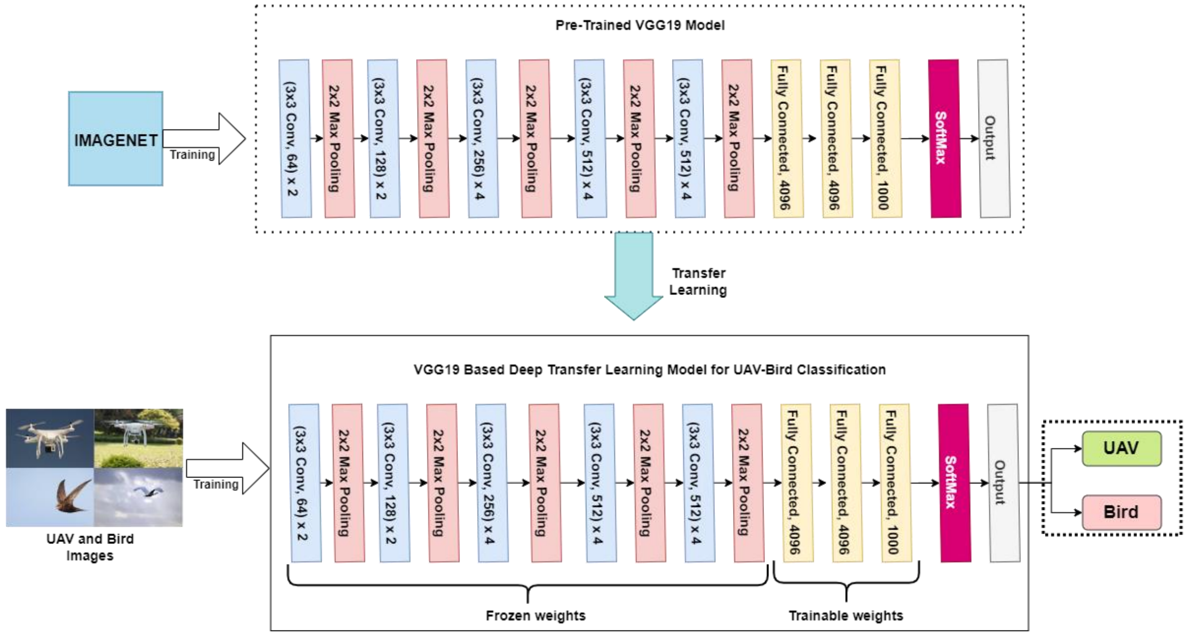

Fig. 4 VGG19 architecture and trained VGG19-based deep learning model

C. *Training Models with Transfer Learning Method*

In this study, three different models were trained using the dataset mentioned in Section II-A and pre-trained DenseNet, VGG16, and VGG19 networks through the transfer learning method. The model training was performed on a computer with a GEFORCE MX150 graphics card and 8GB RAM, using the GPU, for 50 epochs. Each epoch was completed in 20 steps. Python programming language, TensorFlow and Keras libraries were utilized for all processes. The images in the dataset were provided as input to the DenseNet, VGG16, and VGG19 models with a size of 224 x 224 pixels. The learning rate parameter was set to 10^{-5}. Data augmentation was achieved during training by applying operations such as rotation, translation, zooming, and cropping to the data. The parameters used for this operation are shown in Table 2.

Table 2. Data Augmentation Parameters

Parameter	**Value**
rotation_range	40
width_shift_range	0.2
height_shift_range	0.2
shear_range	0.2
zoom_range	0.2
horizontal_flip	True

III. EXPERIMENTAL RESULTS

60 randomly selected images for each class from the images in the collected data set were reserved for testing. Trained DenseNet, VGG16 and VGG19 based deep learning models were tested with these images and the results were evaluated. Accuracy and loss metrics were used for evaluation. Model training was repeated 30 times to ensure the consistency of the results obtained. Box plots of the scores obtained in repetitions are shown in figure 5, figure 6 and figure 7.

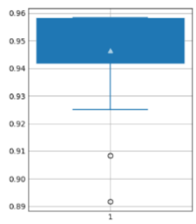

Fig. 5.a DenseNet training repetitions accuracy statistics

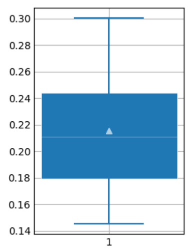

Fig. 5.b DenseNet training repetitions loss function statistics

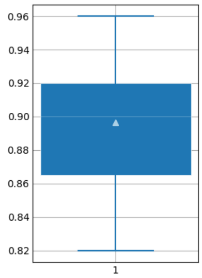

Fig. 6.a VGG16 training repetitions accuracy statistics

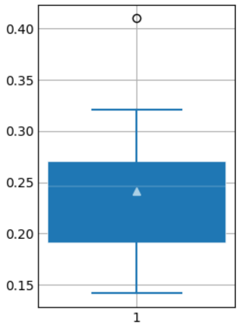

Fig. 6.b VGG16 training repetitions loss function statistics

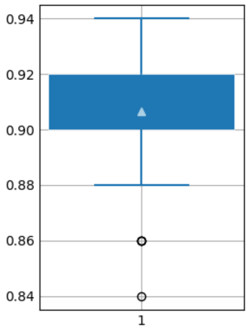

Fig. 7.a VGG19 training repetitions accuracy statistics

The average scores obtained from the training repetitions for each model are presented in Table 3. According to the accuracy scores, the DenseNet model achieved the highest accuracy, while the VGG16 model had the lowest accuracy. When it comes to the loss metric, the DenseNet model again demonstrated the best performance.

Table 3. Test Metrics

No	Method	ACC	Loss
1	DenseNet	%94,64	%21,54
2	VGG16	%89,67	%24,11
3	VGG19	%90,67	%25,04

The loss and accuracy scores for each epoch during the model training were saved as line graphs using the matplotlib library in Python. The saved graphs are shown as follows: Figure 8 for the DenseNet model, Figure 9 for the VGG16 model, and Figure 10 for the VGG19 model.

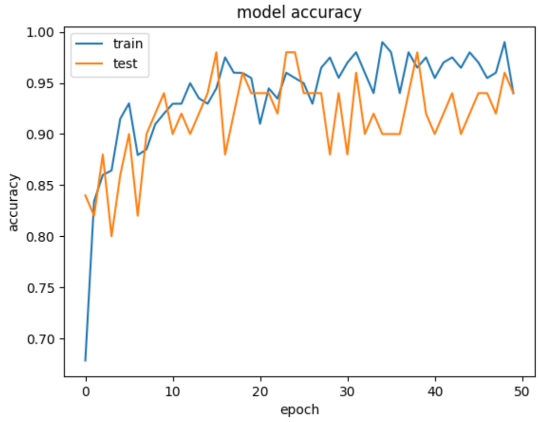

Fig. 8.a DenseNet training accuracy scores

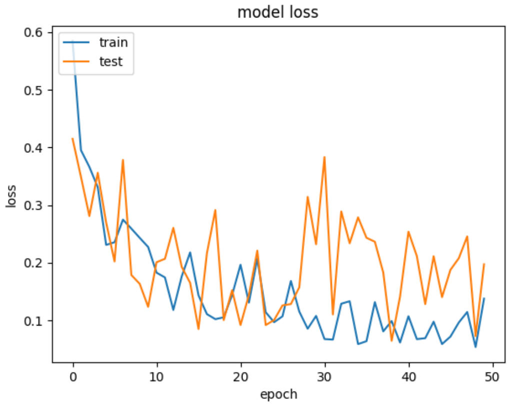

Fig. 8.b DenseNet training loss function

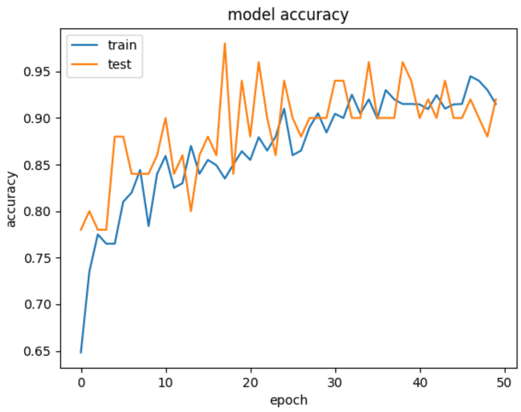

Fig. 9.a VGG16 training accuracy scores

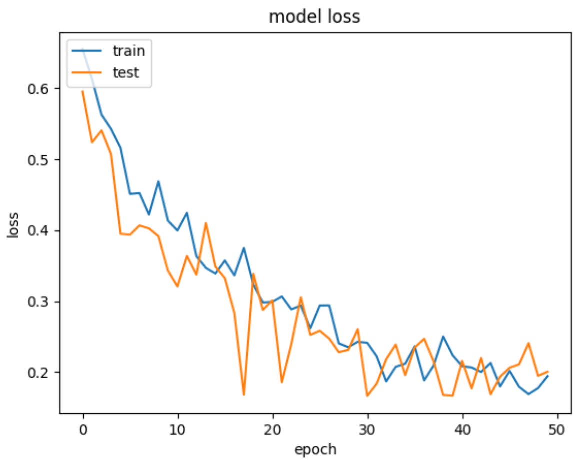

Fig. 9.b VGG16 training loss function

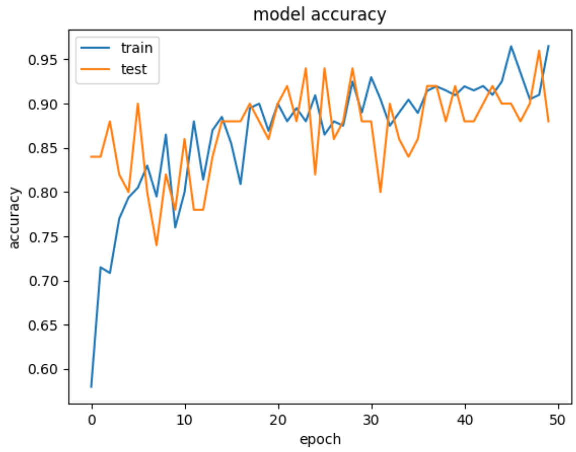

Fig. 10.a VGG19 training accuracy scores

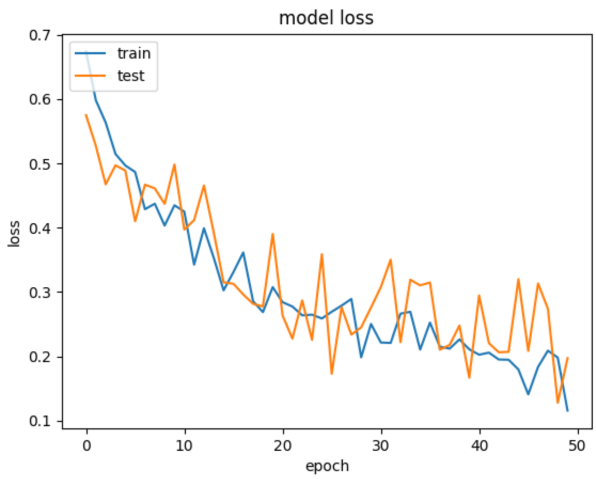

Figure 10.b VGG19 training loss function

IV. CONCLUSION

Unmanned aerial vehicles (UAVs), despite their easy accessibility and portability, also pose risks in terms of security, including malicious uses such as terrorism and espionage. Therefore, it is crucial to take precautions against UAVs in areas prone to security risks. Anti-UAV systems play a significant role in implementing such measures. These systems utilize different techniques based on radio frequency, acoustic, infrared, electro-optical, and other technologies. However, each of these different techniques comes with its unique challenges and disadvantages. In this study, a solution has been sought for the commonly encountered problem of "confusing birds with UAVs" in electro-optical-based Anti-UAV systems. To this end, pre-trained DenseNet, VGG16, and VGG19 networks were used to classify UAVs and birds based on images. To this end, pre-trained DenseNet, VGG16, and VGG19 networks were used to classify UAVs and birds based on images. Three different deep learning models were trained using the "Birds vs Drone Dataset," which consists of bird and UAV images available on the Kaggle platform, through transfer learning. Some of the data are reserved as test data. The results of experiments conducted on the test data using the trained deep learning models were compared. According to the results, the DenseNet model exhibited the highest accuracy, while the VGG16 model showed the lowest accuracy. When compared with CNN, SVM, and KNN models trained with a dataset similar in size to the one used in this study, the DenseNet architecture demonstrated higher accuracy. While outperforming models such as GoogleNet and Inception-V3 trained with a large amount of data, it lagged behind AlexNet, ResNet18, ResNet50, and VGG16 models. The trained DenseNet model in this study performed better than the VGG16 model but fell short compared to the VGG16 model trained with a large dataset. This indicates that the size of the training dataset can affect model performance. This study demonstrated successful results using the DenseNet architecture, which has not been previously attempted in previous studies on classifying birds and UAVs. In future work, the trained classification models can be integrated into a real-time operating system to produce solutions compatible with real-world conditions. Additionally, the network structures can be optimized to be lightweight and perform well in limited-resource environments such as embedded systems. Before conducting model training, identifying regions of interest (ROIs) in the UAV and bird images in the dataset can enhance the scores of classification models.

REFERENCES

[1] E. BUDAK, "Teknolojik Gelişmelerin Habercilik Uygulamaları Üzerine Etkileri: Türkiye'de Drone Haberciliği," *Türkiye İletişim Araştırmaları Dergisi,* no. 33, pp. 119-139, 2019.

[2] H. Özkan, "İNSANSIZ HAVA ARAÇLARININ/DRONE'LARIN TÜRK SİVİL HAVACILIK HUKUKUNA GÖRE STATÜSÜ, UNSURLARI ve CEZA HUKUKU BOYUTUYLA GÜNCEL SORUNLAR," *Türkiye Barolar Birliği Dergisi,* vol. 2016, no. 125, pp. 341-386, 2016.

[3] O. VİLLİ and M. YAKAR, "İnsansız Hava Araçlarının Kullanım Alanları ve Sensör Tipleri," *Türkiye İnsansız Hava Araçları Dergisi,* vol. 4, no. 2, pp. 73-100.

[4] A. Jackman, "Consumer drone evolutions: Trends, spaces, temporalities, threats," *Defense & Security Analysis,* vol. 35, no. 4, pp. 362-383, 2019.

[5] G. Yusuf and E. ERCİYES, "İnsansız Hava Araçları (İHA) Tehditleri ve Güvenlik Yönetimi," *Türkiye İnsansız Hava Araçları Dergisi,* vol. 2, no. 2, pp. 36-42, 2020.

[6] R. J. Kerczewski, J. D. Wilson, and W. D. Bishop, "Frequency spectrum for integration of unmanned aircraft," in *2013 IEEE/AIAA 32nd Digital Avionics Systems Conference (DASC)*, 2013: IEEE, pp. 6D5-1-6D5-9.

[7] M. A. Dini, S. O. Ajakwe, D.-S. Kim, J. M. Lee, and T. Jun, "Droneilliance and Detection Dynamics: A Review of Radar Techniques and Trends."

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems,* vol. 28, 2015.

[9] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*, 2009: Ieee, pp. 248-255.

[10] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556,* 2014.

[11] M. Saqib, S. D. Khan, N. Sharma, and M. Blumenstein, "A study on detecting drones using deep convolutional neural networks," in *2017 14th IEEE international conference on advanced video and signal based surveillance (AVSS)*, 2017: IEEE, pp. 1-5.

[12] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part I 13*, 2014: Springer, pp. 818-833.

[13] A. Coluccia *et al.*, "Drone vs. bird detection: Deep learning algorithms and results from a grand challenge," *Sensors,* vol. 21, no. 8, p. 2824, 2021.

[14] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, "Yolov4: Optimal speed and accuracy of object detection," *arXiv preprint arXiv:2004.10934,* 2020.

[15] S. Singha and B. Aydin, "Automated Drone Detection Using YOLOv4," *Drones,* vol. 5, no. 3, p. 95, 2021.

[16] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM,* vol. 60, no. 6, pp. 84-90, 2017.

[17] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2015, pp. 1-9.

[18] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2818-2826.

[19] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.

[20] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and< 0.5 MB model size," *arXiv preprint arXiv:1602.07360,* 2016.

[21] H. M. Oh, H. Lee, and M. Y. Kim, "Comparing Convolutional Neural Network (CNN) models for machine learning-based drone and bird classification of anti-drone system," in *2019 19th International Conference on Control, Automation and Systems (ICCAS)*, 2019: IEEE, pp. 87-90.

[22] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767,* 2018.

[23] E. Unlu, E. Zenou, N. Riviere, and P.-E. Dupouy, "Deep learning-based strategies for the detection and tracking of drones using several cameras," *IPSJ Transactions on Computer Vision and Applications,* vol. 11, no. 1, pp. 1-13, 2019.

[24] F. Mahdavi and R. Rajabi, "Drone detection using convolutional neural networks," in *2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS)*, 2020: IEEE, pp. 1-5.

[25] H. R. Alsanad, A. Z. Sadik, O. N. Ucan, M. Ilyas, and O. Bayat, "YOLO-V3 based real-time drone detection algorithm," *Multimedia Tools and Applications,* vol. 81, no. 18, pp. 26185-26198, 2022.

[26] D. K. Behera and A. B. Raj, "Drone detection and classification using deep learning," in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2020: IEEE, pp. 1012-1016.

[27] S. S. Selvi, S. Pavithraa, R. Dharini, and E. Chaitra, "A Deep Learning Approach to Classify Drones and Birds," in *2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon)*, 2022: IEEE, pp. 1-5.

[28] B. B. Bhagat, R. R. Sharma, and D. Tilante, "Moving camera-based automated system for drone identification using focus measures," *Signal, Image and Video Processing,* pp. 1-8, 2023.

[29] H. WALIA. "Birds vs Drone Dataset." Kaggle. https://www.kaggle.com/datasets/harshwalia/birds-vs-drone-dataset (accessed.

[30] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE,* vol. 86, no. 11, pp. 2278-2324, 1998.

[31] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700-4708.