SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V10I11P102

Volume 10 Issue 11, 21-30, November 2023

Original Article

DL Based Multi-Class Drone Classification for Counter Drone Detection Applications

Venkata Subba Rao Pittu[1], Usha Rani Nelakuditi[2], Pavani Bandla[3], Yojitha Thotakura[4]

[1]*Department of Electrical and Electronics Engineering, Vignan's Foundation for Science, Technology and Research, Andhra Pradesh, India.*
[2,3,4]*Department of Electronics and Communication Engineering, Vignan's Foundation for Science, Technology and Research, Andhra Pradesh, India.*

[1]*Corresponding Author : pvsrao.p@gmail.com*

Received: 03 September 2023 Revised: 05 October 2023 Accepted: 04 November 2023 Published: 30 November 2023

Abstract *- Drones are the new disruption technology at present owing to their use in many areas like transport, agriculture, security, surveillance, surveying & mapping, etc., to handle various critical tasks with less complexity and cost-effectively. In this research, drone usage in air surveillance, especially as a counter drone technique, is considered which is a significant threat today at borders. Transfer learning-based multiclass drone detection and classification were implemented using pretrained ResNet-50, VGG-16, Inception and Xception nets. Drone detection and classification performance for drone, bird, helicopter, and aeroplane classes are validated using accuracy, precision, F-score and recall metrics. Xception net is performing well over other nets with an accuracy of 0.98.*

Keywords *- UAV, Transfer learning, VGG-16, ResNet-50, Inception net, Xception net, Drone detection, Drone classification.*

1. Introduction

The drone industry has witnessed significant growth due to wide emerging applications of drones in many areas, such as medicine, agriculture, security, etc., due to the policy taken by the government. Several industries, like construction and the medical domain, are more dependent on drone technology to observe the works and carry drugs and organs, despite being in the infant stage.

As per regulations of DGCA, people can hire a drone and obtain a UID number license. Safety and security concerns are significant when individuals also utilize drones. Few are probing drones for unethical works such as carrying and dropping explosives, drug smuggling, chemical weapons, surveillance into prohibited areas, etc. Detection technologies are now being researched, each with its own set of tradeoffs in complexity, range, and capability [20]. As a result, security teams require a method of detecting drones in the air and being aware of what is flying in their area. Acoustic devices, lasers, infrared sensors, LIDAR, and RADARs are existing technologies for detecting, localizing, and identifying small drones. Targets can be removed once detected by using birds trained to catch drones and jamming target-detected drones, laser guns, water cannons, and laser guns. Drones, on the other hand, can be employed to counter malicious drones. In recent times, academicians and industry working on computer vision-based techniques like object classification and detection methods have been implemented in deep learning-based Convolutional Neural Network (CNN) architectures, which are amicable solutions in surveillance.

In this research, ResNet-50 and VGG-16-based drone detection and classification models were dispensed for the detection of drones and to reduce criminal activities in the geofencing areas. It is more challenging to classify the drone and bird in critical real-time situations like low contrast, less visibility, high range, etc., and even more complicated to use an algorithm to classify drone and bird with maximum accuracy. When flying, the drone should not collide with other birds or drones. Hence, this research concentrates on the detection and classification methods of drones from birdsVGG-16, ResNet-50, Inception V3, and Xception nets, which are employed with greater accuracy and less loss.

2. Related Works

Many researchers are working in this domain, and a few state-of-the-art works are mentioned as follows. Mohammed Javed et al. [12] investigated a ResNet-50 and faster RNN-based real-time drone surveillance system, achieving 79% accuracy. It also uses four distinct object detectors, which detect faces and weapons. Tamer Khattab [13] proposed a model for tracking traffic and various RF signals from drones and classifying them into 2, 4, and 10 classes utilizing DNN techniques. Dong Kyu [11] et al. proposed a model of a drone detection and identification system with an accuracy of 89 percent using AI and OpenCV. This system was equipped with a camera, which infers position via machine classification. In 2016, Dinesh Kumar Behra [4] proposed a deep-learning model for drone identification and classification with an accuracy of 97.4%.

Senthilnath et al. proposed deep learning-based spectral-spatial methods such as splitting and merging in hierarchical categorization for vegetation analysis. R. Girshick proposed RCNN [16] for object detection. It effectively classified the object and was nine times faster in testing. Daniel Tan Wei Xun, Yoke Lin Lim et al. [1] used YOLOv3 and machine learning for drone detection, which had an average accuracy of 88.9% in 2021.

3. Proposed Method

Multiclass drone detection and classification are proposed and implemented using pre-trained deep neural networks such as VGG-16, ResNet-50, InceptionV3, and Xception networks.

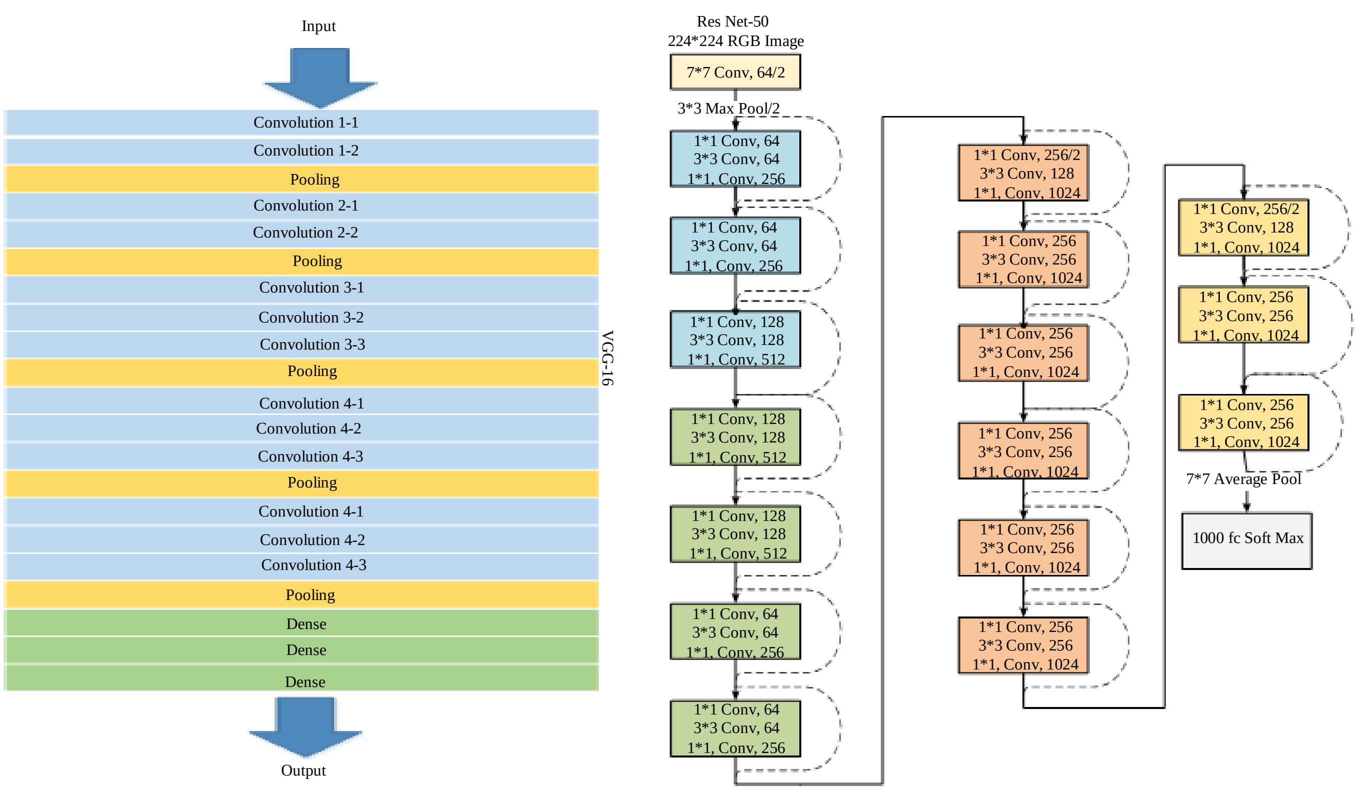
3.1. VGG-16 Architecture

VGG-16 CNN model is a deep, sequential net with 16 layers, as shown in Figure 1. This allows an input image of 224×224×3. It incorporates thirteen convolution layers [14] of fixed filter size (3×3), five pooling layers of 2×2 filter size, and three fully connected layers with about 138 million parameters. But it has the problem of vanishing gradients.

3.2. ResNet Architecture

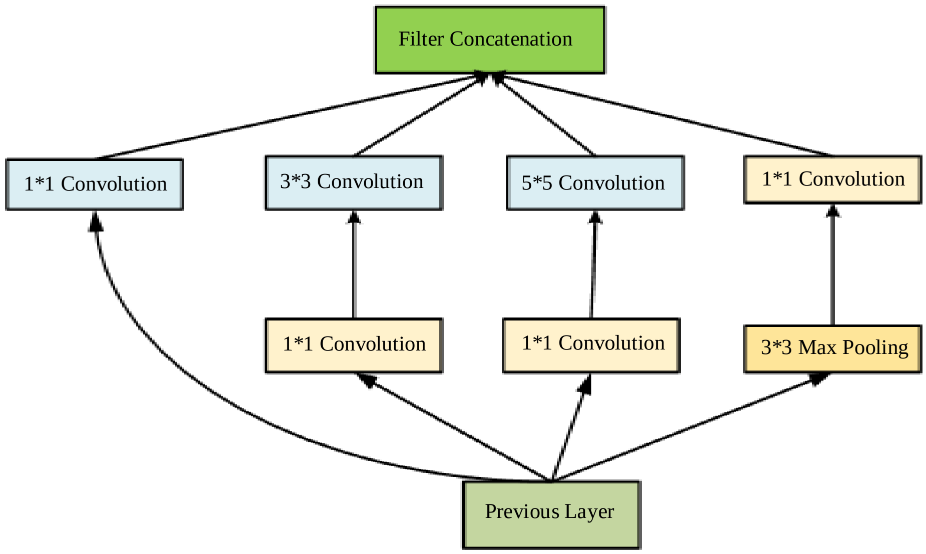
ResNet-50 is a form of exotic architecture, as shown in Figure 1, unlike sequential nets like Lenet and VGG net, where macro architecture is formed by replicating residual modules known as ResNet modules. It is a pre-trained net that can distinguish 1000 classes and consists of 48 convolution layers, a max pool layer, and an average pool layer [10], which takes an input image of size 224×224. It solves the problem of overfitting by using skip connections, and higher-level layers' performance is as good as lower-level layers' due to bypass connections.

The primary inception module is a variant of ResNet, with a multi-level feature extractor, as shown in Figure 1, by computing 1×1, 3×3, and 5×5 convolutions within the same network module and feeding them to the output. It uses the split-transform-merge principle to obtain the feature map by using a point-wise grouped convolutional layer, which divides its input into groups of feature maps and performs standard convolution. Model capacity is determined by using parameter cardinality.

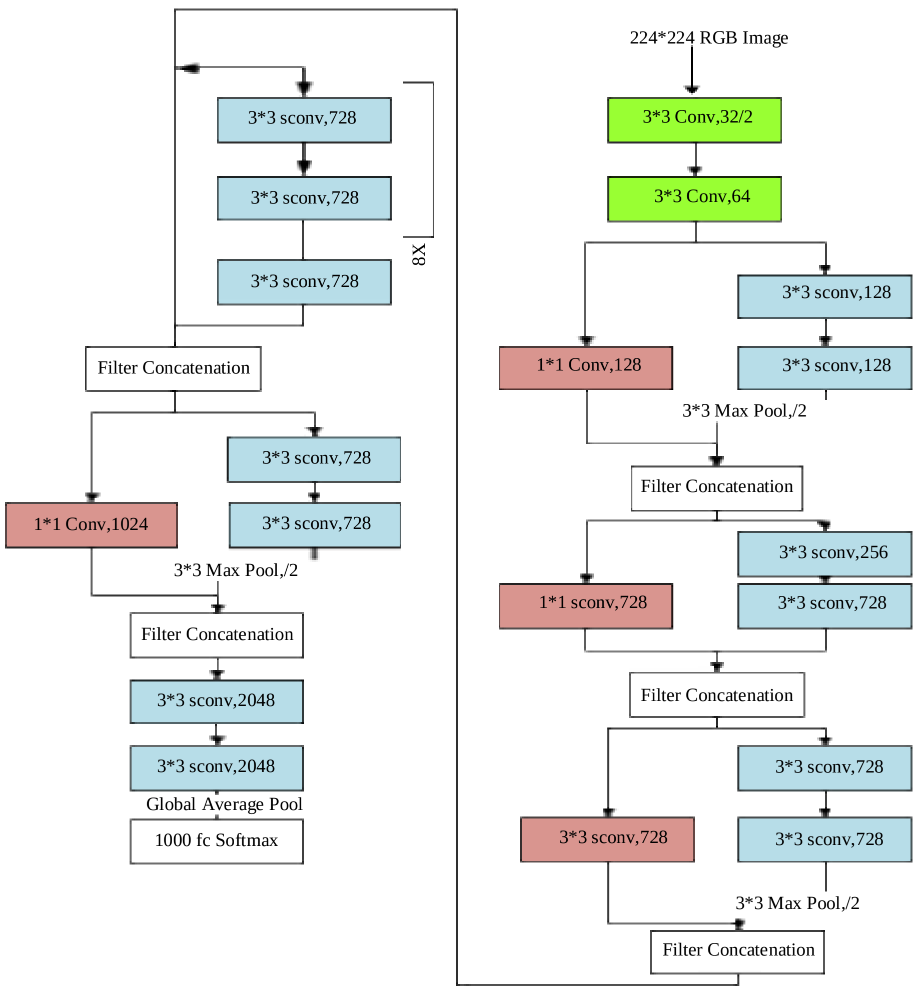


(a) VGG-16

(b) ResNet-50



(c) Inception net



(d) Xception net

Fig. 1 Pre-trained architectures

4. Implementation

The proposed drone identification and classification methodology is displayed in Figure 2, which uses photos from Kaggle. They are pre-processed before being separated into training and testing datasets in an 80:20 ratio, then blended and spewed with pictures and labels.The model is then built by adjusting hyper-parameters, regularisation, and optimization approaches.

The testing dataset is used to assess the model's performance, whereas the training dataset is used to train the model. Accuracy, precision, recall, and the F1-score measure performance. Pre-trained CNN networks such as VGG-16 and ResNet-50, Inception, and Xception are implemented in tensor flow with the Keras API for drone detection as a black box solution for multiclass classification in this paper using four as shown in Table 2 and Table 3. Net is initialized with the ImageNet weights trained with 40 epochs and a batch size 32. Data augmentation is used on models to avoid overfitting due to sample restrictions. Dropout is introduced during training to reduce the problem of overfitting and acceleration of the training phase. Fully Connected (FC) and the softmax layers used will perform classification. This research dataset is prepared using around 2000 images of drones, birds, helicopters, and aeroplanes. Details are represented in Table 1.

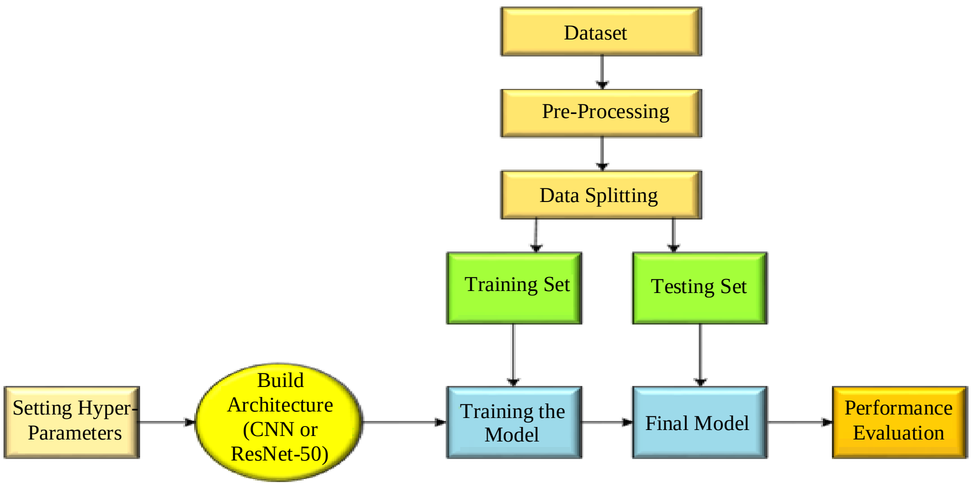


Fig. 2 Methodology for the drone detection

Table 1. Dataset details

Parameter	VGG-16	ResNet-50	Inception V3	Xception
Image Shape	224x224	224x224	224x224	224x224
Dropout Rate	0.5	0.4	0.4	0.4
Classifier	Softmax	Softmax	Softmax	Softmax
Optimizer	RMS Prop	Adam	RMS Prop	Adam
Loss Function	CCE	CCE	CCE	CCE
Regularization	Nil	BN	BN	BN
Batch Size	32	32	32	32
Epochs	40	40	40	40

Table 2. Hyperparameters for classification

Class	No. of Image Used for		
	Training	Testing	Total
Drone	431	143	574
Bird	429	144	573
Aero Plane	105	60	165
Helicopter	152	100	252
Total	1117	447	1564

4.1. Performance Metrics

The following Equations (1-4) provide the accuracy, precision, recall, and F1-score metrics for evaluating the performance of the models generated by VGG-16, ResNet-50, Inception V3, and Xception for the categorization of drones from birds, airplanes, and helicopters. The following Equations (1-4) provide the accuracy, precision, recall, and F1-score metrics for evaluating the performance of the models generated by VGG16, ResNet-50, Inception V3, and Xception for the categorization of drones from birds, airplanes, and helicopters.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (2)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (3)$$

$$\text{F1 Score} = 2*\frac{Recall x Precision}{Recall + Precision} \quad (4)$$

Where,

True Positive (TP) refers to the number of accurately predicted true cases.

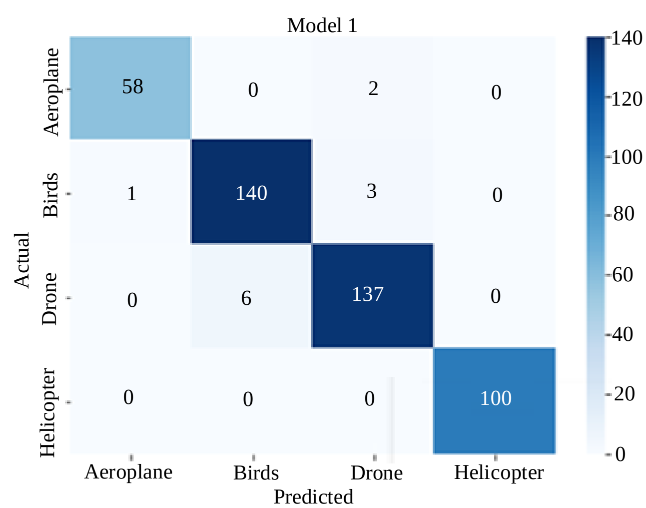
True Negative (TN) is the total number of incorrectly predicted false cases.

The total number of correctly predicted true cases is referred to as the False Negative (FN).

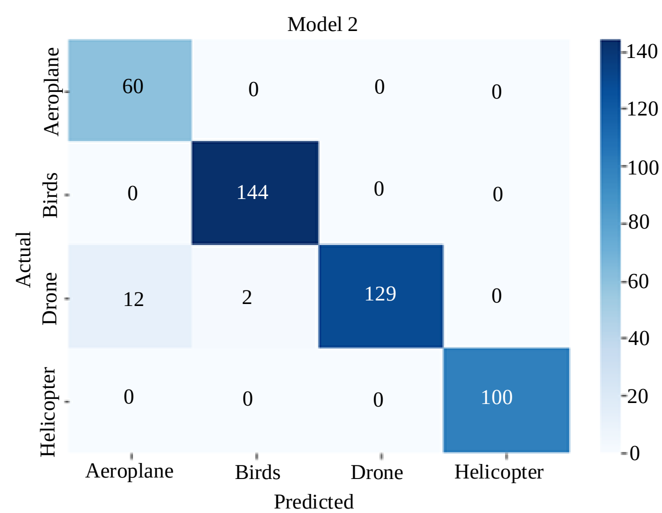
False Positive (FP) refers to the total number of cases accurately predicted but false.

4.2. Confusion Matrix

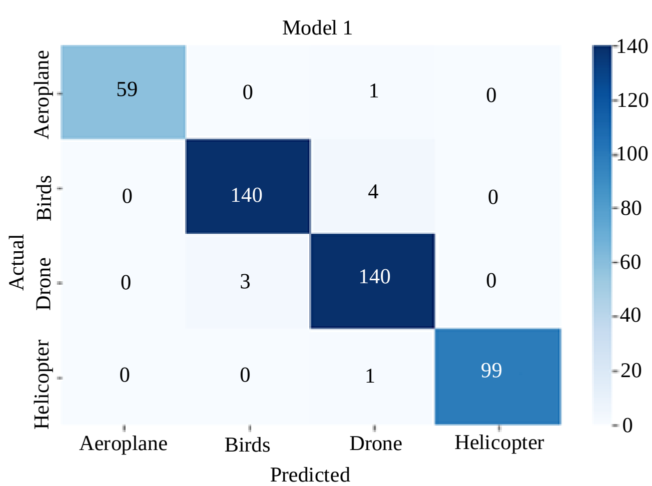
Confusion matrices of VGG-16, ResNet-50, Inception, and Xception nets are shown in Figure 3(a)-(d), respectively. In this case, each network was learned with one thousand four hundred-three samples of the training dataset were validated with a test data set of 447 samples. The aeroplane class is correctly classified without confusion in the exception network. VGG-16 and ResNet-50 performances are the same, and the other two functions are the same.



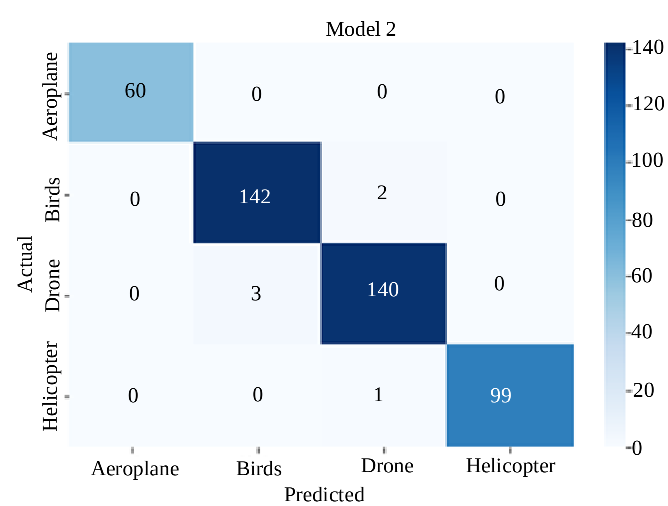
(a) Confusion matrix of VGG-16



(b) Confusion matrix of ResNet-50



(c) Confusion matrix of Inception v3



(d) Confusion matrix of Xception

Fig. 3 Confusion matrix of various network models

In the case of drones, VGG-16 produced more misclassifications due to the gradients vanishing problem. This problem is addressed in ResNet50, as shown in Table 3. Due to extended multi-scale feature characteristic performance further improved in xception architecture, the false rate was also reduced.

Table 3. Performance comparison

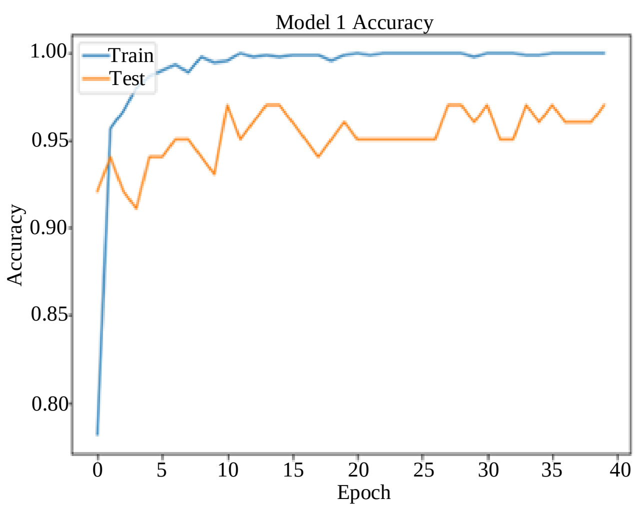
Method	Class	TP	TN	FP	FN	Accuracy	Precision	Recall	F1 Score
VGG-16	Airplane	58	385	2	1	0.99	0.98	0.96	0.97
	Bird	140	296	4	6	0.97	0.95	0.97	0.96
	Drone	137	298	6	5	0.97	0.96	0.95	0.96
	Helicopter	100	346	-	-	1.0	1.0	1.0	1.0
ResNet-50	Airplane	60	374	-	12	0.97	0.83	1.0	0.90
	Bird	144	300	-	2	0.99	0.98	1.0	0.99
	Drone	129	303	14	-	0.96	1.0	0.90	0.94
	Helicopter	100	346	-	-	1.0	1.0	1.0	1.0
Inception	Airplane	59	386	1	0	0.99	1.0	0.98	0.99
	Bird	140	299	4	3	0.98	0.97	0.97	0.97
	Drone	140	297	3	6	0.97	0.95	0.97	0.96
	Helicopter	99	346	1	0	0.99	1.0	0.99	0.99
Xception	Airplane	60	386	0	0	1.0	1.0	1.0	1.0
	Bird	142	299	3	2	0.98	0.97	0.98	0.98
	Drone	140	300	3	3	0.98	0.97	0.97	0.97
	Helicopter	99	346	1	0	0.99	1.0	0.99	0.99

5. Performance Evaluation

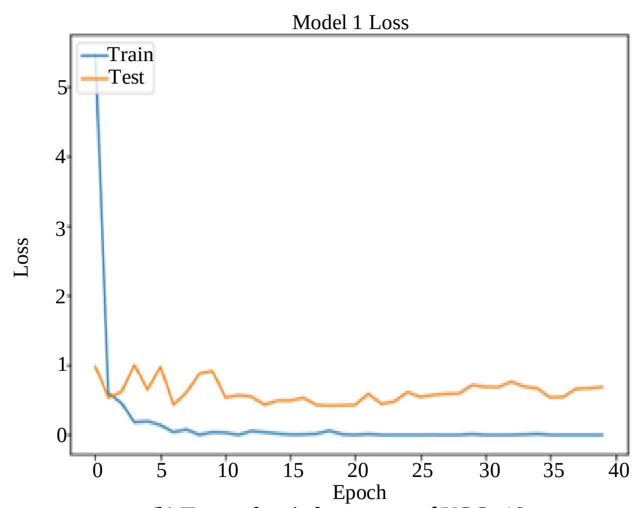
The performance of the proposed work is compared with the help of accuracy and loss functions, as shown in Figure 4 and Table 4, respectively. Xception net with higher accuracy has high entropy loss. Inception net is moderate in terms of accuracy and loss. VGG-16's net loss is shallow. The accuracy and loss functions of four proposed networks for 40 epochs are shown in Figure 4.

Table 4. Accuracy and loss of four models

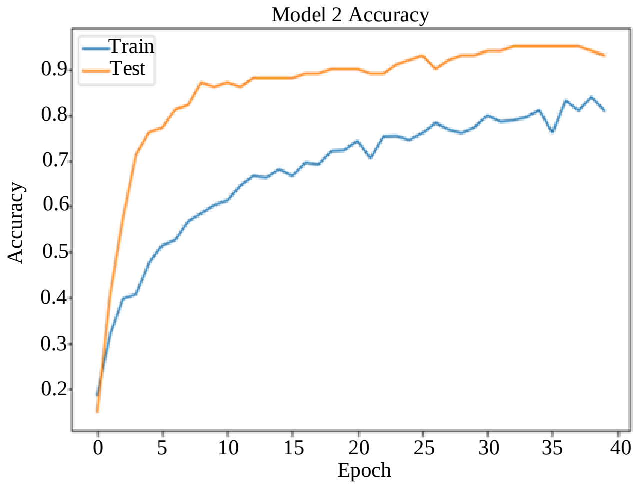
Parameters	VGG-16	ResNet-50	InceptionV3	Xception
Accuracy	0.97	0.96	0.97	0.98
Loss	0.47	0.57	0.48	0.65



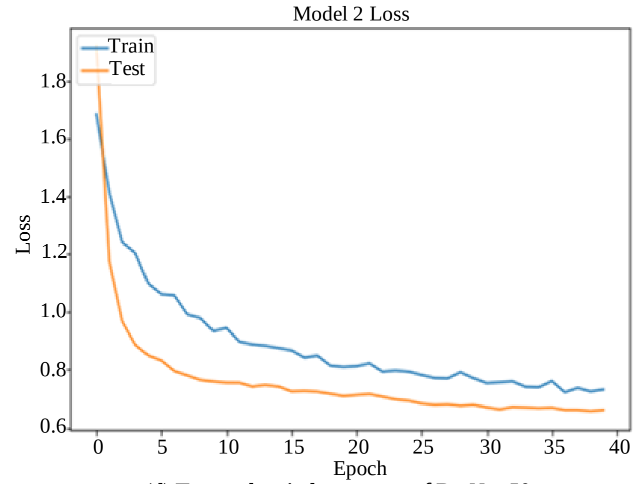
(a) Test and train accuracy curves of VGG- 16



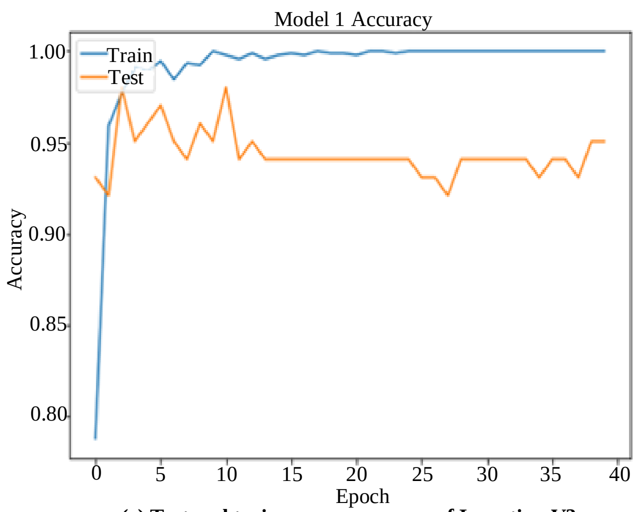
(b) Test and train loss curves of VGG- 16



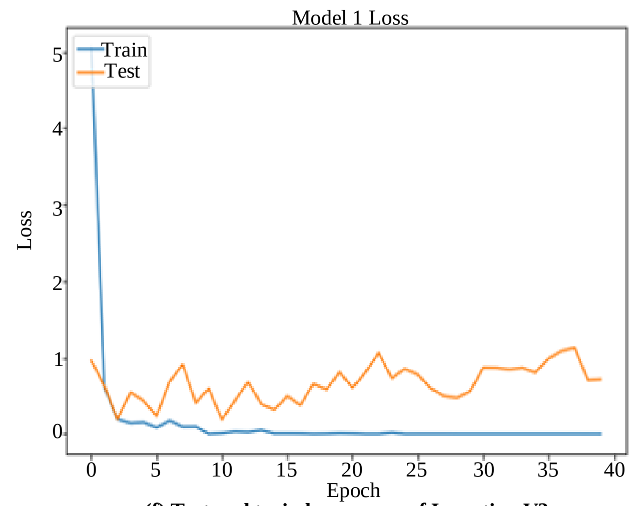
(c) Test and train accuracy curves of ResNet-50



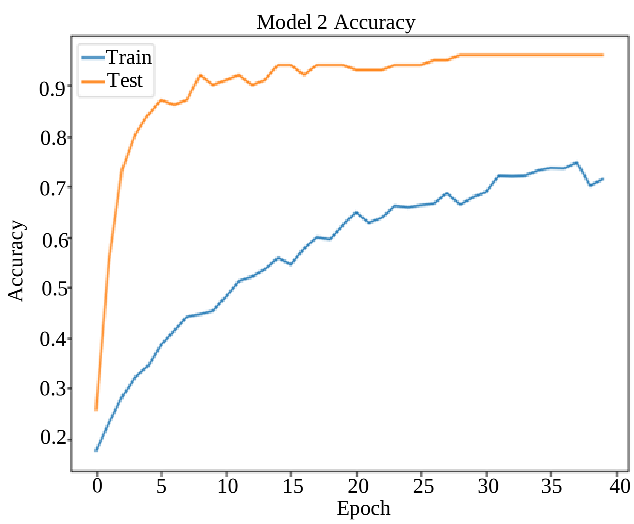
(d) Test and train loss curves of ResNet-50



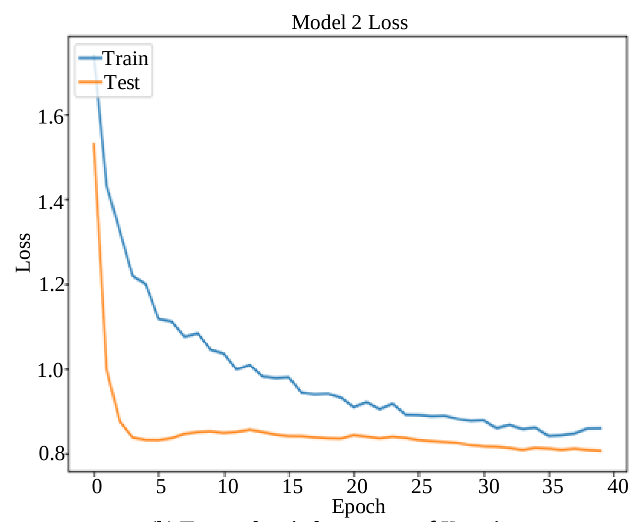
(e) Test and train accuracy curves of Inception V3



(f) Test and train loss curves of Inception V3



(g) Test and train accuracy curves of Xception



(h) Test and train loss curves of Xception

Fig. 4 Training and testing accuracy and loss functions of proposed methods

Table 5. Performance comparison with the state of work

Publication Details	Accuracy	Proposed Classes	Proposed Methods
Muhammad Javed et al. [18]	79%	Drone surveillance	ResNet-50, faster RNN
Tamer Khattabet et al. [13]	97.7%,84.5 %,46.8%	Different RF signals from drones with 2,4,10 classes	DNN
Dong kyu Leeetal [11]	89%	Drone images	AI, OpenCV
Dinesh Kumar Behra et al.	94.74%	Tricopter, Quadcopter, Hexacopter	CNN, YOLOv3
Proposed Models	**0.97**	**Drone, Bird, Aeroplane, Helicopter (Multiclass classification)**	**VGG-16**
	0.96		**ResNet-50**
	0.97		**InceptionV3**
	0.98		**Xception**

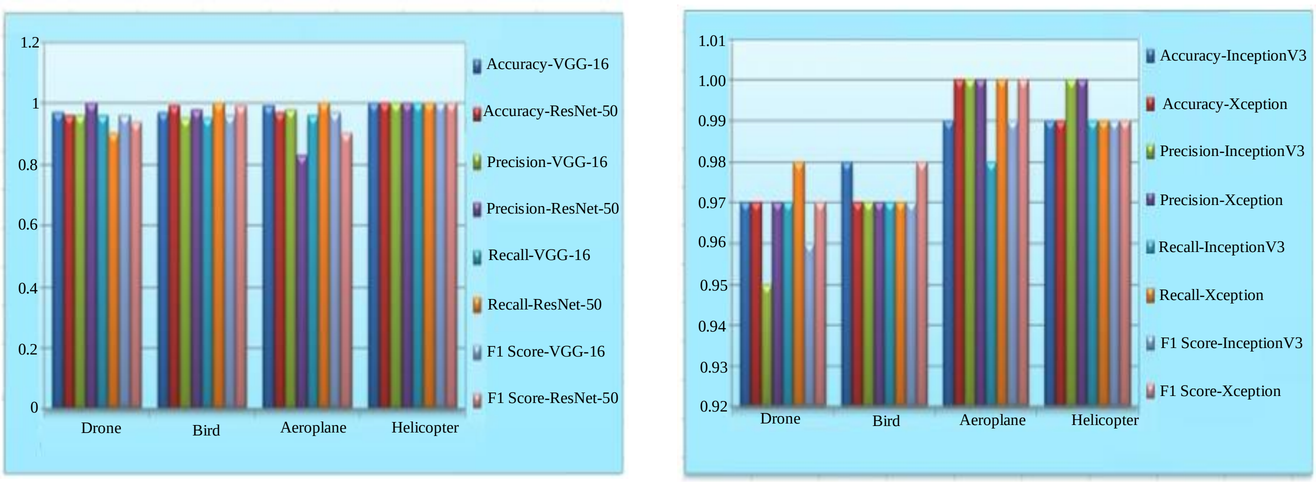


Fig. 5 Performance comparison of four models

6. Performance Comparison

The proposed work is compared with the state-of-the-art work, as shown in Table 5. The proposed method with Xception architecture can classify drones with high detection accuracy.

7. Conclusion

This study suggests two pre-trained models, VGG-16 and ResNet-50, for binary and multiclass classification. These models divide objects into four separate categories: drones and birds for binary classification and aeroplanes and helicopters for Multiclass classification based on their shape.

For multiclass classification, the accuracies for the VGG-16 and ResNet-50 models are 0.97 and 0.96, respectively, while the accuracies for binary classification are 0.9 and 1.0 for each model.

The ResNet-50 model outperformed the VGG-16 model in binary classification due to skip connections in the residual block that decrease overfitting and because lower-dimensional layers operate equally well as higher-dimensional layers. Because fewer kernels exist in the ResNet-50 than in the VGG net, its computational complexity is likewise lower. InceptionV3 and Xception have better accuracies of 0.97 and 0.98.

References

[1] Daniel Tan Wei Xun, Yoke Lin Lim, and Sutthiphong Srigrarom, "Drone Detection Using YOLOv3 with Transfer Learning on NVIDIA Jetson TX2," *2021 Second International Symposium on Instrumentation, Control, Artificial Intelligence, and Robotics (ICA-SYMP)*, Bangkok, Thailand, pp. 1-6, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] Han Sun et al., "TIB-Net: Drone Detection Network with Tiny Iterative Backbone," *IEEE Access*, vol. 8, pp. 130697-130707, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[3] Jing Gong et al., "Vehicle Detection in Thermal Images with an Improved Yolov3-Tiny," *2020 IEEE International Conference on Power, Intelligent Computing and Systems (ICPICS)*, Shenyang, China, pp. 253-256, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[4] Kangquan Ye et al., "Research on Small Target Detection Algorithm Based on Improved Yolov3," *2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE)*, Harbin, China, pp. 1467-1470, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Fuyan Lin, Xin Zheng, and Qiang Wu, "Small Object Detection in Aerial View Based on Improved YoloV3 Neural Network," *2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA)*, Dalian, China, pp. 522-525, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[6] Pompílio Araújo, Jefferson Fontinele, and Luciano Oliveira, "Multi-Perspective Object Detection for Remote Criminal Analysis Using Drones," *IEEE Geoscience and Remote Sensing Letters*, vol. 17, no. 7, pp. 1283-1286, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[7] Jintao Wang, Wen Xiao, and Tianwei Ni, "Efficient Object Detection Method Based on Improved YOLOv3 Network for Remote Sensing Images," *2020 3rd International Conference on Artificial Intelligence and Big Data (ICAIBD)*, Chengdu, China, pp. 242-246, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[8] Zhang Menghan, Li Zitian, and Song Yuncheng, "Optimization and Comparative Analysis of YOLOV3 Target Detection Method Based on Lightweight Network Structure," *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, Dalian, China, pp. 20-24, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[9] Gang Sha, Junsheng Wu, and Bin Yu, "Detection of Spinal Fracture Lesions Based on Improved YOLOV2," *2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA)*, Dalian, China, pp. 235-238, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[10] Syed Ali Hassan, Tariq Rahim, and Soo Young Shin, "Real-Time UAV Detection Based on Deep Learning Network," *2019 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, Korea, pp. 630-632, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[11] Dongkyu Lee, Woong Gyu La, and Hwangnam Kim, "Drone Detection and Identification System Using Artificial Intelligence," *2018 International Conference on Information and Communication Technology Convergence (ICTC)*, Jeju, Korea, pp. 1131-1133, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[12] Muhammad Javed Iqbal et al., "Real-Time Surveillance Using Deep Learning," *Journal of Security and Communication Networks*, vol. 2021, pp. 1-17, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] Mhd Saria Allahham, Tamer Khattab, and Amr Mohamed, "Deep Learning for RF-Based Drone Detection and Identification: A Multi-Channel 1-D Convolutional Neural Networks Approach," *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)*, Doha, Qatar, pp. 112-117, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[14] Dinesh Kumar Behera, and Arockia Bazil Raj, "Drone Detection and Classification Using Deep Learning," *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, India, pp. 1012-1016, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Eren Unlu, Emmanuel Zenou, and Nicolas Rivière, "Using Shape Descriptors for UAV Detection," *Electronic Imaging*, pp. 1-5, 2018. [Google Scholar] [Publisher Link]

[16] Ross Girshick, "Fast R-CNN," *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, pp. 1440-1448, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[17] Juhyun Kim et al., "Real-Time UAV Sound Detection and Analysis System," *2017 IEEE Sensors Applications Symposium (SAS)*, Glassboro, NJ, USA, pp. 1-5, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[18] J. Senthilnath et al., "Application of UAV Imaging Platform for Vegetation Analysis Based on Spectral-Spatial Methods," *Computers and Electronics in Agriculture*, vol. 140, pp. 8-24, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[19] Shaoqing Ren et al., "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Advances in Neural Information Processing Systems*, vol. 28, 2015. [Google Scholar] [Publisher Link]

[20] Brian Hearing, and John Franklin, "Drone Detection and Classification Methods and Apparatus," *U.S. Patent No. US-10032464-B2*, 2016. [Publisher Link]

[21] Muhammad Saqib et al., "A Study on Detecting Drones Using Deep Convolutional Neural Networks," *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, Lecce, Italy, pp. 1-5, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[22] Dar-Shyang Lee, "Effective Gaussian Mixture Learning for Video Background Subtraction," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 5, pp. 827-832, 2005. [CrossRef] [Google Scholar] [Publisher Link]

[23] John Lai, Luis Mejias, and Jason J. Ford, "Airborne Vision-Based Collision-Detection System," *Journal of Field Robotics*, vol. 28, no. 2, pp. 137-157, 2011. [CrossRef] [Google Scholar] [Publisher Link]