e-ISSN : 0976-5166
p-ISSN : 2231-3850

AIRCRAFT RECOGNITION SYSTEM USING DEEP LEARNING BASED EFFICIENTNET

Poojitha K (Student), Dr. Azra Nasreen (Associate Professor)
Department of Computer Science and Engineering, RV College of Engineering,
R V Vidyanikethan Post, Mysuru Road, Bengaluru - 560 059, India

Abstract–- Maintaining airspace safety and minimizing potential hazards have become urgent concerns as a result of the exponential expansion in aviation traffic. Proposed Aircraft Recognition System (ARS) is created to improve aviation safety by properly identifying and classifying aircrafts for the given input. The ARS uses sophisticated computer vision and deep learning techniques to quickly and automatically identify different kinds of military aircrafts. ARS performs recognition task using EfficientNet B3. Comparison with other existing models is done to evaluate the accuracy of proposed system. The accuracy of this model was observed to be 95.76% in comparison with VGG16, InceptionV3 and ResNet50 whose accuracy was observed to be 79.1%, 88.57%, and 56.2% respectively.

Index Terms— Aircraft Recognition system, CNN, EfficientNet B3, VGG16, Inception, Resnet50.

1. INTRODUCTION

The aerospace industry is essential to current military operations, where the quick identification of aircraft is crucial for efficient security and defense. A powerful and trustworthy Aircraft Recognition System (ARS) designed for military applications is becoming more and more necessary as governments deal with an ever-changing panorama of aerial threats and difficulties.

The popularity of unmanned aerial vehicles (UAVs) and the development of advanced stealth technologies have created new difficulties for conventional identification techniques in the field of military aviation. Manual visual recognition by human operators can be prone to mistakes, especially under stressful circumstances, inclement weather, or in crowded airspace. Consequently, the creation of a smart, automated aircraft recognition system has turned into a strategic necessity for contemporary armed forces.

This main goal of the proposed system is to use EfficientNet B3 as the aircraft recognition system and also survey all the available algorithms that is especially suited for military purposes. The system's primary job is to quickly and correctly detect, categories, and track aircrafts for the given input by utilizing cutting-edge artificial intelligence (AI) and computer vision algorithms. The ARS can quickly interpret enormous volumes of visual data by utilizing deep learning models and neural networks. This allows it to distinguish between different aircraft types, including conventional fixed-wing aircraft, helicopters, drones, and stealthy assets.

2. LITERATURE SURVEY

A unique hybrid convolution neural network model for efficient object detection in unmanned aerial vehicle (UAV) images is presented in [1]. The difficulties caused by unequal object distribution, complicated backgrounds, and a range of scales in UAV photos are addressed by the suggested solution. The model uses a feature pyramid network to improve object detection at various levels and a cross-shaped window transformer as a backbone to get multiscale picture features. A hybrid patch embedding module is additionally presented to extract low-level data, such as edges and corners, to enhance feature data. Without changing the original network, a slicing-based inference technique is used to increase small item detection accuracy. The suggested strategy outperforms various established and well-liked object detection techniques, according to experimental results. The paper also notes that several state-of-the-art techniques currently available have not produced perfect outcomes when applied to the training dataset employed in this study, necessitating additional investigation.

The drawbacks of the current anchor-based and anchor-free rotational object recognition techniques for aerial photos is proposed in [2]. Anchor-based approaches produce good results, but they also come with manually set anchors, which adds hyper parameters and computational complexity. However, anchor-free approaches are difficult and have a slow rate of inference. The researchers suggest an anchor-free rotating detector for object detection in aerial photos based on the effective YOLOX approach to get over these problems. A refined rotated module (RRM) for aligning features and getting relevant priors and a novel assigner method called Gaussian distribution sampling optimal transport assignment (GSOTA) method are two significant advancements made by the method. In GSOTA, oriented bounding boxes are modelled using a Gaussian distribution, and a sampling mechanism is added to make label assignment easier. The suggested R2YOLOX models achieve competitive performance in a variety of aerial object identification datasets while maintaining high efficiency, according to experimental findings. Target representation, computational complexity, and training convergence are a few issues with the method that may call for additional optimization. The work concludes by introducing R2YOLOX, a YOLOX-based anchor-free rotating object detection algorithm for aerial photos. The suggested method produces results that are competitive across a range of datasets, providing a decent balance between speed and accuracy. Incorporating new target properties, reducing computational complexity, and improving training effectiveness are just a few of the areas that still require work. To balance detection accuracy and training time, the researchers use a pretrained model on COCO. However, additional work is required to optimize the strategy for even higher performance and practical implementation.

Utilizing deep learning to recognize aircraft using a hybrid attention network model dubbed BA-CNN is proposed in [3]. Due to the wide range of aircraft types, similarities between various models, and texture interference, traditional approaches have difficulty classifying aircrafts. In order to improve the extraction of fine-grained characteristics, BA- CNN uses a two-channel ResNet34 for feature extraction and adds channel attention and spatial attention modules between residual units. The experimental findings on the FGVC-plane BA-CNN model outperforms the majority of currently used commonplace aircraft recognition techniques, with a recognition precision rate of 89.2%. The article does, however, concede that the bilinear characteristic vector outer product combination adds dimensions and processing resources, necessitating further research to reduce dimensionality and enhance the usefulness of the network model.

In order to recognize objects and scenes in machine vision systems, the system [4] offers a novel Scene Semantic Recognition (SSR) framework. The system seeks to address issues such as dynamic backgrounds, occlusion, a lack of labelled data, variations in illumination, direction, and size. De- noising and smoothing of scene data, updated Fuzzy C-Means with super-pixels and Random Forest for object segmentation, and a unique Bag of Features for feature extraction are all included in the proposed SSR framework. For object recognition based on various object patterns, an Artificial Neural Network (ANN) is used, and Maximum Entropy is used to estimate scene labels. The proposed SSR framework surpasses current state-of-the-art systems in terms of computation, segmentation, and accuracy, according to experimental results on a variety of datasets. Autonomous driving, intelligent traffic, remote aerial sensing, augmented reality, military training, engineering design, artificial eyes for the blind, and traffic monitoring are a few possible uses for the system.

The human-computer collaboration framework [5] for identifying aircraft in multi-object and background remote sensing images is introduced. The framework takes advantage of human visual search abilities by using human eye tracking to suggest candidate aircraft. The two-step recognition procedure mimics the methodology used by image analysts. Fully connected features are used in the first step to quickly identify aircraft, and if that fails, encoded convolutional features are used to confirm the types of aircraft. High recall rates, low false alarm rates, and accurate location accuracy show the effectiveness of the suggested framework. Extensive testing demonstrates that the method performs recognition better than cutting-edge techniques.

A novel method for classifying aircraft in synthetic aperture radar (SAR) images called the Scattering Topology Network (ST-Net) [6] is presented. The challenges posed by scattering discreteness and imaging variability at various angles are addressed by this function. The framework makes use of scattering cluster centres to improve discriminative features and update category data. To improve classification accuracy, a scattering topology module (STM) is introduced. It models the spatial relationships and semantic information interactions

of scattering points. To reduce background noise, a context-aware attention excitation (CAE) mechanism collects global and semantic information. The effectiveness and superiority of ST-Net over current CNN-based methods are demonstrated by extensive experiments on the SAR aircraft category dataset (SAR-ACD), which achieves superior classification performance in high-resolution SAR images.

3. CONCEPTS AND THEORIES BEHIND ARS

Computer Vision:

The study of computer vision enables machines to decipher and comprehend visual data from the outside world. In order to preprocess and analyze aircraft images or sensor data, extract pertinent features, and find recognisable patterns that can be used for classification, computer vision techniques are used.

Convolutional Neural Networks:

EfficientNet B3 belongs to the class of deep learning models known as CNNs, which are built to automatically learn hierarchical representations of visual data. Given their ability to accurately detect spatial patterns and features in images, CNNs are well suited for image recognition tasks.

Transfer learning:

It is a technique that uses a pre- trained model, such as EfficientNet B3, as a starting point for a particular task (in this case, aircraft recognition), as opposed to training the model from scratch. It is simpler to modify and fine-tune the pre- trained model for a particular recognition task with a smaller dataset because it has already learned general features from a large dataset.

Feature Extraction:

EfficientNet B3 has been trained to identify intricate patterns and structures in images and extract high-level features from them. The essential qualities necessary for identifying aircraft are captured in these extracted features, which are representations of the input images.

Data Augmentation:

Techniques like rotation, flipping, and random cropping are frequently used on the training dataset to improve the generalization and robustness of the model. As a result, the training data are more diverse, which decreases overfitting and enhances the model's capacity to identify instances of unobserved aircraft.

4. EFFICIENTNET B3 ARCHITECTURE

Convolutional neural network architecture EfficientNet B3 is a member of the EfficientNet family. For image recognition tasks, it is intended to effectively balance model size and performance. The network's depth, width, and resolution are all scaled uniformly by the architecture using scaling coefficients. When compared to smaller variants like B0 and B1, EfficientNet B3 has a larger and more potent model because the scaling coefficients are set to 1.2 for depth and 1.3 for width, respectively.

Multiple blocks with inverted residual connections make up EfficientNet B3, including MBConv blocks with depth wise and pointwise convolutions. These building blocks enable the model to efficiently capture both low-level and high-level features, enabling it to recognize complex patterns in images. A Feature Pyramid Network (FPN) is also incorporated into EfficientNet B3 to enable multi- scale feature extraction, improving the model's ability to recognize objects of various sizes in the input image. EfficientNet B3 is a compelling option for applications in fields like computer vision, autonomous systems, and natural language processing thanks to the combination of these architectural components that enables it to achieve state-of-the-art performance in image recognition tasks, including object recognition.

5. CONCEPTS TO KNOW

The system has two major components. The first is building a model.

Data Preparation: The data directories for training, validation, and testing are taken as input from the user. The dataset is then split into data frames, with file paths and corresponding labels for each image.

Image Data Generators: ImageDataGenerators are set up to pre-process the images, perform data augmentation (e.g., horizontal flip), and convert the images into tensors. Three data generators are created for training, validation, and testing data.

Model Architecture: The EfficientNet B3 model is defined using the pre-trained EfficientNetB3 as the base model. Two dense layers are added, along with batch normalization and dropout, to build the final model for classification.

Custom Callback Function: A custom callback function is created to monitor the training process. This callback adjusts the learning rate based on the model's performance and stops training if there is no improvement.

Training: The model is compiled with the Adamax optimizer and categorical cross-entropy loss. The training is performed using the fit() function, and the custom callback is used during the training process.

Evaluation: After training, the model's performance is evaluated on the training, validation, and testing datasets. The accuracy and loss are displayed.

Prediction: The system allows users to input new images for prediction. First, an aircraft detection algorithm could be used to check if an aircraft is present in the image. If an aircraft is detected, the image is fed to the saved model for classification. Otherwise, a negative output is given.

Visualization: Confusion matrices and classification reports are generated to visualize the model's performance on the testing dataset.

6. IMPLEMENTATION

The system architecture of the Aircraft Recognition System is shown in Fig. 1, which provides a high-level overview of the system. The major components of the system are the pre-processing modules, and EfficientNet model. The image of an aircraft is provided as input, which undergoes pre-processing, which is fed to the trained model to classify.

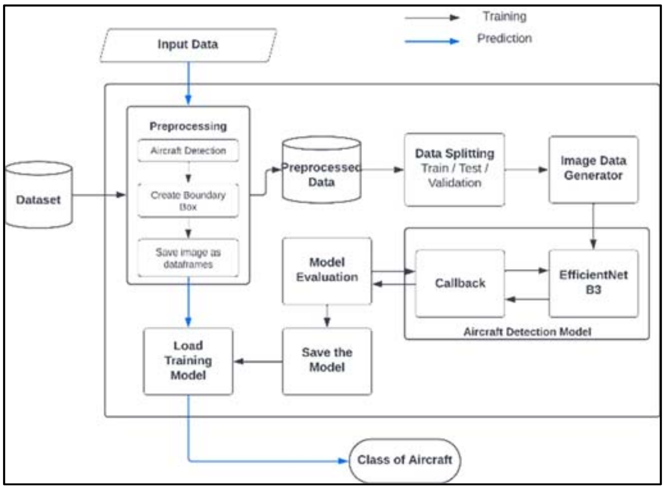

Fig 1: System Architecture

The sign for the classification flow is indicated in blue in Figure 1.

6.1 DATA SET

The dataset used for this project was obtained from a website initially gathered using google images. This data set contains images of 43 aircraft types, which are as follows:

A-10, A-400M, AG-600, AV-8B, B-1, B-2, B-52,Be-200, C-130, C-17, C-2, C-5, E-2, E-7, EF-2000, F-117, F-14, F-15, F-16, F/A-18, F-22, F-35, F-4, J- 20, JAS-39, MQ-9, Mig-31, Mirage2000, P-3(CP- 140), RQ-4, Rafale, SR-71, Su-34, Su-57, Tornado, Tu-160, Tu-95(Tu-142), U-2, US-2(US-1A Kai), V-22, Vulcan, XB-70, YF-23

There are around 11,500 images of these aircrafts. Some images have multiple aircraft with both scenarios of similar and different. The algorithms used for detection is EfficientNet B3. The split up of dataset was 60-20-20 for training, testing and validation of this models. First few images of the dataset along with the aircraft name is shown in fig 2.

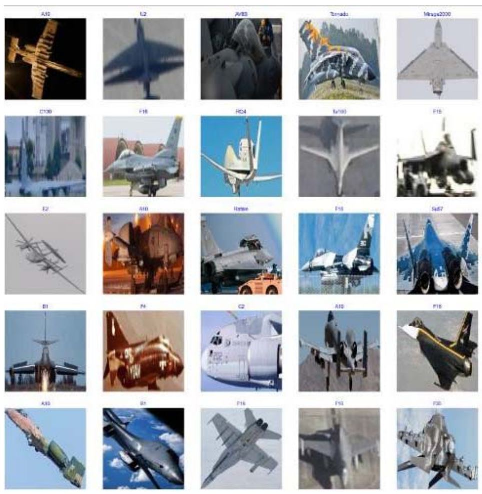

Fig 2: Dataset of aircrafts

Data pre-processing is an essential step that must be completed before the implementation process can begin. The first step in pre-processing is to normalize the photos to a specific dimension. Then, we use normalization and augmentation techniques, which include rotation, height/width changes, shear transformation, and zoom modifications. This cleaned-up data is first fed into an EfficientNet B3 model that has already been trained and operates using predetermined training parameters.

6.2 Training the Model

Split the pre-processed data into training / testing / validation sets based on Table 1 values. The model was trained as shown in fig 3 on the training set of images. Based on the Table 1 values which describes the hyper-parameter data, the model is fine-tuned.

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 efficientnetb3 (Functional)  (None, 1536)             10783535

 batch_normalization (BatchN  (None, 1536)             6144
 ormalization)

 dense (Dense)               (None, 256)               393472

 dropout (Dropout)           (None, 256)               0

 dense_1 (Dense)             (None, 43)                11051

=================================================================
Total params: 11,194,202
Trainable params: 11,103,827
Non-trainable params: 90,375
_________________________________________________________________
```

Fig 3: Model Summary

The evaluation on model is performed using the validation set to determine the accuracy, precision, recall, and F1-score. The hyper-parameters were adjusted to by the model to optimise its performance.

Table 1: Hyper-parameters

Hyper-Parameter	**Value**
Training Data	80% (9200)
Testing Data	20% (2300)
Image size	224 x 224
Channel	3
Base Model - Kernel Regularizer	L2 regularization with a value of 0.016
Base Model – Activity Regularizer	L1 regularization with a value of 0.006
Base Model - Bias Regularizer	L1 regularization with a value of 0.006
Base Model – Rate	Dropout rate of 0.45
Base Model – Optimizer	Adamax optimizer with a learning rate of 0.001
Base Model – loss	Categorical Cross-entropy
Batch Size	40
Epochs	40
Patience	1
Stop Patience	3
Threshold	0.9
Factor	0.5

7. RESULTS AND ANALYSIS

Evaluation matrices used here are Accuracy, Precision, Recall, and F1 score. Validation image data is to validate the model, batch by batch, after it has been sufficiently trained. The accuracy was observed to be 95% with only 5 epochs as shown in fig 4, and the confusion matrix is shown in fig 5

```
history = model.fit(x= train_gen, epochs= epochs, verbose= 0, callbacks= callbacks,
                    validation_data= valid_gen, validation_steps= None, shuffle= False)

Do you want model asks you to halt the training [y/n] ?
y
 Epoch     Loss    Accuracy  V_loss    V_acc     LR      Next LR  Monitor  % Improv  Duration
 1 /40     8.244    32.379   5.26239   66.234   0.00100  0.00100  accuracy    0.00   12142.92
 2 /40     3.989    72.214   2.77536   80.983   0.00100  0.00100  accuracy  123.03   11842.47
 3 /40     2.104    86.654   1.66444   86.278   0.00100  0.00100  accuracy   20.00   11794.63
 4 /40     1.255    92.969   1.20596   86.818   0.00100  0.00100  val_loss   27.55   11974.47
 5 /40     0.875    95.765   0.98181   89.195   0.00100  0.00100  val_loss   18.59   11712.89
enter H to halt training or an integer for number of epochs to run then ask again
H
training has been halted at epoch 5 due to user input
training elapsed time was 16.0 hours, 31.0 minutes, 48.85 seconds)
```

Fig 4: Model Evaluation.

The confusion matrix shown is fig 5 is for the EfficientNet B3 architecture. There are very few false positive and false negative values.

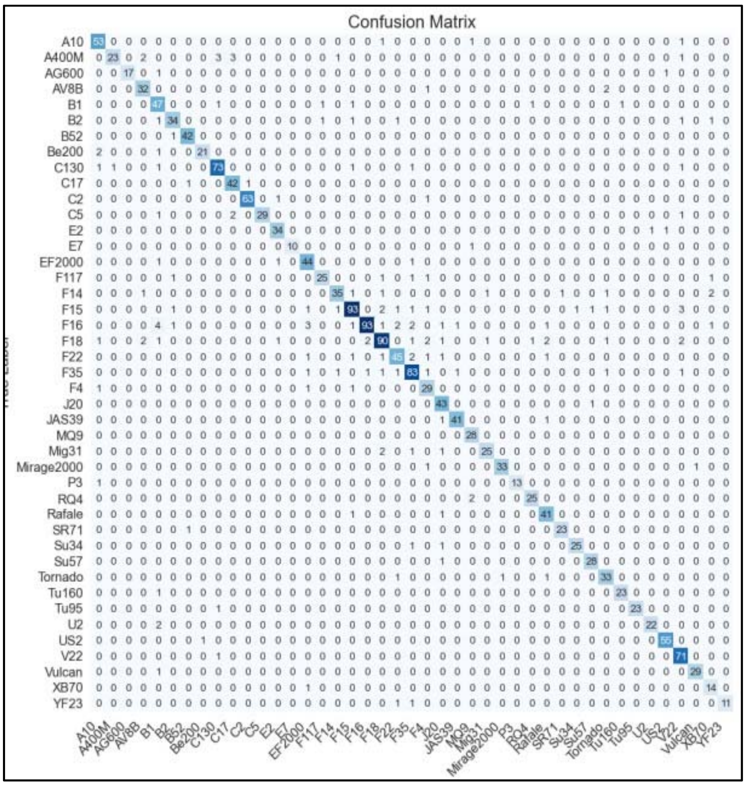

Fig 5: Confusion matrix.

The classification report is shown in figure 6. It shows the precision, recall, and F1-score for each class of aircraft along with the number of images considered in each category. Also, the whole system report with same parameters is displayed after the report for each class is shown.

Classification Report:

	precision	recall	f1-score	support
A10	0.90	0.95	0.92	56
A400M	0.96	0.70	0.81	33
AG600	1.00	0.89	0.94	19
AV8B	0.86	0.91	0.89	35
B1	0.76	0.90	0.82	52
B2	0.89	0.85	0.87	40
B52	0.95	0.98	0.97	43
Be200	0.95	0.88	0.91	24
C130	0.92	0.92	0.92	79
C17	0.89	0.95	0.92	44
C2	0.98	0.97	0.98	65
C5	1.00	0.88	0.94	33
E2	0.92	0.94	0.93	36
E7	1.00	0.91	0.95	11
EF2000	0.85	0.94	0.89	47
F117	0.93	0.83	0.88	30
F14	0.92	0.81	0.86	43
F15	0.92	0.87	0.89	107
F16	0.97	0.84	0.90	111
F18	0.90	0.83	0.87	108
F22	0.87	0.85	0.86	53
F35	0.87	0.90	0.89	92
F4	0.76	0.91	0.83	32
J20	0.84	0.98	0.91	44
JAS39	0.95	0.95	0.95	43
MQ9	0.88	1.00	0.93	28
Mig31	0.93	0.86	0.89	29
Mirage2000	0.97	0.94	0.96	35
P3	1.00	0.93	0.96	14
RQ4	0.93	0.93	0.93	27
Rafale	0.87	0.95	0.91	43
SR71	0.96	0.96	0.96	24
Su34	0.96	0.93	0.94	27
Su57	0.93	0.97	0.95	29
Tornado	0.85	0.92	0.88	36
Tu160	0.96	0.96	0.96	24
Tu95	1.00	0.96	0.98	24
U2	0.96	0.92	0.94	24
US2	0.96	0.98	0.97	56
V22	0.87	0.99	0.92	72
Vulcan	0.97	0.97	0.97	30
XB70	0.74	0.93	0.82	15
YF23	1.00	0.85	0.92	13
accuracy			0.91	1830
macro avg	0.92	0.91	0.91	1830
weighted avg	0.91	0.91	0.91	1830

Fig 6: Classification Report for Validation Dataset

Training accuracy is a metric that indicates the percentage of correctly classified samples from the training dataset during the model training process. Validation accuracy, on the other hand, is a metric used to evaluate the model's performance on a separate validation dataset that it has not seen during training. Fig 7 shows that the trained model exhibits robust performance and quick learning. The fact that the testing accuracy approaches 90% after only 5 epochs shows that the model is effective at generalizing to new, untested samples and quickly grasping the underlying patterns in the data. Additionally, the fact that the validation accuracy regularly stays around 85% and that the model does well on both the training and validation datasets shows that it is not overfitting to the training data.

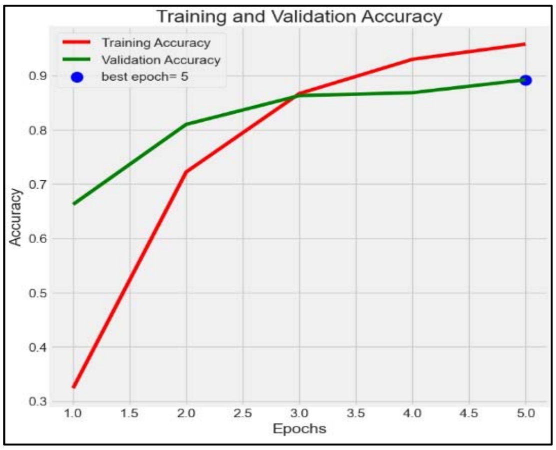

Fig 7: Testing and Validation Accuracy

Training loss is a metric that represents the error or the difference between the predicted output of the model and the actual target values for the samples in the training dataset. Validation loss, on the other hand, is a metric used to evaluate the model's performance on a separate validation dataset that it has not seen during training.

According to Fig. 8, within 5 epochs, the validation loss is kept between 1 and 2% and the training loss is decreased to less than 1%. This demonstrates how effective and efficient the model's training process is. The reduced training loss shows that the model learns to fit the training data with amazing accuracy very soon. The validation loss is also consistently low, ranging from 1 to 2%, which indicates that the model is generalizing effectively to new data.

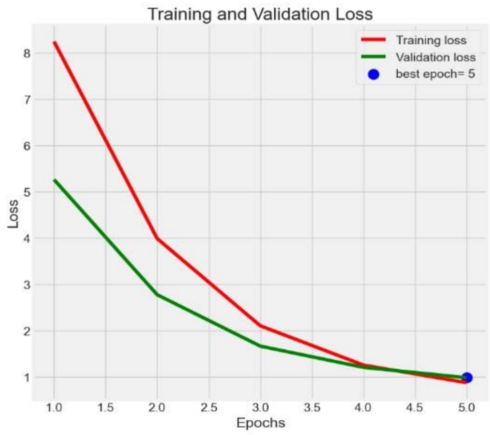

Fig 8: Testing and Validation Loss

To compare the results of this model, three other models were implemented, namely VGG16, Inception, and ResNet50 and the accuracy, loss, Validation accuracy and loss of these models are shown in table 2.

Table 2: Classification report for All the Models

Model Name	Accuracy	Loss	Validation Loss	Validation Accuracy
EfficientNet B3	**95.76**	**0.875**	**0.981**	**89.195**
VGG16	79.10	0.698	0.6208	80.83
InceptionV3	88.57	1.3257	1.1198	90.88
ResNet50	10.21	3.3869	3.6964	4.69

As observed, the Efficient Net B3 algorithm produced better results than the other models. Fig 9 shows a testing of an aircraft 'Rafale' using EfficientNet model.

Fig 9: Testing using EfficientNet Model

8. CONCLUSION

This paper proposes an aircraft recognition system which is capable of recognizing the type of aircraft present in the given input. The model used is EfficientNet B3 architecture. The accuracy of this model was observed to be 95.76%. Other models such as VGG16, InceptionV3 and ResNet50 were implemented with the same data and the accuracy of these models was observed to be 79.1%, 88.57%, and 56.2% respectively.

Since the model can regulate different parameters (such as learning index), it was able to perform better. An improvement to this would be using dataset that contain sensor or radar images of these aircrafts. These data can be included to make the system more versatile.

8.1 Limitations

The project's reliance on Google Images as its main data source is one of its limitations. There may be variations in image resolution, lighting, and item postures since Google Images cannot guarantee consistent data quality. Additionally, the search criteria used could introduce biases or neglect specific aircraft types or modifications, which would affect the dataset's variety and representativeness.

The potential for overfitting is another drawback, particularly when training set accuracy levels are high. It's possible for the model to memorize particular photos rather than learning resilient features, which could limit its capacity to generalize to new data. A larger, more varied dataset and regularization approaches may be able to aid with this problem.

8.2 Future Enhancement

Several crucial steps will be taken in the future to improve the capabilities and applicability of the airplane recognition project using EfficientNet B3. First, the model resilience will be increased by enlarging the dataset to include a greater variety of aircraft types and climatic variables. Broader deployment will be possible once the model has been adjusted to fit real-time processing and edge devices. Insights into model decisions will be gained by addressing interpretability issues using strategies like attention mechanisms.

The model will continue to be relevant through routine updates that include new aircraft models and variants. Finally, working with organizations and aviation professionals can result in insightful improvements and improvements specific to the aviation industry.

Acknowledgments
No funding injected towards the consolidation of this manuscript.

Conflict of interest

The author have no conflicts of interest to declare.

REFERENCES

[1] W. Lu et al., "A CNN-Transformer Hybrid Model Based on CSWin Transformer for UAV Image Object Detection," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 1211-1231, 2023, doi: 10.1109/JSTARS.2023.3234161.
[2] Y. Kang et al., "ST-Net: Scattering Topology Network for Aircraft Classification in High-Resolution SAR Images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1-17, 2023, Art no. 5202117, doi: 10.1109/TGRS.2023.3236987.
[3] Y. Nie, C. Bian and L. Li, "Adap-EMD: Adaptive EMD for Aircraft Fine-Grained Classification in Remote Sensing," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 6509005, doi:10.1109/LGRS.2022.3168581.
[4] K. Gao, H. He, D. Lu, L. Xu, L. Ma and J. Li, "Optimizing and Evaluating Swin Transformer for Aircraft Classification: Analysis and Generalizability of the MTARSI Dataset," in IEEE Access, vol. 10, pp. 134427-134439, 2022, doi: 10.1109/ACCESS.2022.3231327.
[5] X. Sun, Y. Lv, Z. Wang and K. Fu, "SCAN: Scattering Characteristics Analysis Network for Few-Shot Aircraft Classification in High-Resolution SAR Images," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-17, 2022, Art no. 5226517, doi: 10.1109/TGRS.2022.3166174.
[6] H. Meng, Y. Peng, W. Wang, P. Cheng, Y. Li and W. Xiang, "Spatio-Temporal- Frequency Graph Attention Convolutional Network for Aircraft Recognition Based on Heterogeneous Radar Network," in IEEE Transactions on Aerospace and Electronic Systems, vol. 58, no. 6, pp. 5548-5559, Dec.2022, doi: 10.1109/TAES.2022.3175797.
[7] L. -Q. Zhou, P. Sun, D. Li and J. -C. Piao, "A Novel Object Detection Method in City Aerial Image Based on Deformable Convolutional Networks," in IEEE Access, vol. 10, pp. 31455-31465, 2022, doi: 10.1109/ACCESS.2022.3156953.
[8] Y. Wu, Y. Sui and G. Wang, "Vision-Based Real-Time Aerial Object Localization and Tracking for UAV Sensing System," in IEEE Access, vol. 5, pp. 23969-23978, 2017,doi:10.1109/ACCESS.2017.2764419.
[9] A. Jalal, A. Ahmed, A. A. Rafique and K. Kim, "Scene Semantic Recognition Based on Modified Fuzzy C-Mean and Maximum Entropy Using Object-to-Object Relations," in IEEE Access, vol. 9, pp. 27758-27772, 2021,doi:10.1109/ACCESS.2021.3058986.
[10] Z. Ozkan, E. Bayhan, M. Namdar and A. Basgumus, "Object Detection and Recognition of Unmanned Aerial Vehicles Using Raspberry Pi Platform," 2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), Ankara, Turkey, 2021, pp. 467-472, doi: 10.1109/ISMSIT52890.2021.9604698.
[11] [X. Xie and G. Lu, "A Research of object detection on UAVs aerial images," 2021 2nd International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), Zhuhai, China, 2021,pp.342-345, doi: 10.1109/ICBASE53849.2021.00070
[12] F. Zhang, X. Wang, S. Zhou and Y. Wang, "DARDet: A Dense Anchor-Free Rotated Object Detector in Aerial Images," in IEEE

Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 8024305, doi:10.1109/LGRS.2021.3122924.
[13] X. Li, B. Jiang, S. Wang, L. Shen and Y. Fu, "A Human–Computer Fusion Framework for Aircraft Recognition in Remote Sensing Images," in IEEE Geoscience and Remote Sensing Letters, vol. 17, no. 2, pp. 297-301, Feb. 2020, doi: 10.1109/LGRS.2019.2918955.
[14] F. Azam, A. Rizvi, W. Z. Khan, M. Y. Aalsalem, H. Yu and Y. B. Zikria, "Aircraft Classification Based on PCA and Feature Fusion Techniques in Convolutional Neural Network," in IEEE Access, vol. 9, pp. 161683-161694, 2021, doi: 10.1109/ACCESS.2021.3132062
[15] Q. Wu, H. Sun, X. Sun, D. Zhang, K. Fu and H. Wang, "Aircraft Recognition in High-Resolution Optical Satellite Remote Sensing Images," in IEEE Geoscience and Remote Sensing Letters, vol. 12, no. 1, pp. 112-116, Jan.2015, doi: 10.1109/LGRS.2014.2328358
[16] R. Girshick, J. Donahue, and T. Darrell, "Rich feature hierarchies for accurate object detection and semantic segmentation," in Proceedings of the IEEE Conference on ComputerVision and Pattern Recognition, pp. 580–587, Columbus, OH, USA, June 2014.
[17] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1904–1916, 2015.
[18] R. Girshick, "Fast R-CNN," in Proceedings of the IEEE International Conference on Computer Vision, pp. 1440–1448, Santiago, Chile, December 2015.
[19] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN.:Towards real-time object detection with region Proposal networks," 2015, http://arxiv.org/abs/1506.01497.
[20] K. He, G. Gkioxari, and P. Doll´ar, "Mask R-CNN," in Proceedings of the IEEE International Conference on Computer Vision, pp. 2961–2969, Venice, Italy, October 2017.
[21] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779– 788, Honolulu, HI, USA, July 2016.
[22] W. Liu, D. Anguelov, D. Erhan et al., "SSD: Single shot multibox detector," in Proceedings of the European Conference on Computer Vision, pp. 21–37, 2016.
[23] Y. C. Fu, W. Liu, and A. Ranga, "DSSD: Deconvolutional single shot detector," 2017,http://arxiv.org/abs/1701.06659.
[24] Z. Li and F. Zhou, "FSSD: Feature fusion single shot multibox detector," 2017, http://arxiv.org/abs/1712.00960.
[25] T. Y. Lin, P. Goyal, R. Girshick, K. He, and P. Doll´ar, "Focal loss for dense object detection," in Proceedings of the IEEE International Conference on Computer Vision, pp. 2980–2988, Venice, Italy, October 2017.
[26] Yanfeng Wang, Tao Wang, Xin Zhou , Weiwei Cai , Runmin Liu ,Meigen Huang, Tian Jing, Mu Lin, Hua He, Weiping Wang, and Yifan Zhu, "TransEffiDet: Aircraft Detection and Classification in Aerial Images Based on EfficientDet and Transformer", Hindawi Computational Intelligence and Neuroscience Volume 2022.
[27] Liming Zhou , Haoxin Yan , Yingzi Shan , Chang Zheng , Yang Liu ,Xianyu Zuo , and BaojunQiao, "Aicraft Detection for Remote Sensing Images Based on Deep Convolutional Neural Networks", HindawiJournal of Electrical and Computer Engineering Volume 2021.
[28] Alexander Toshev, Ameeshmakadia, Kostas D, "Shape based Object recognition using 3D Synthetic Object Models", IEEE Conference on Computer Vision and PatternRecognition(CVPR) , 2009, Miami, June 20-25 , pp 288 – 295.
[29] Jingjing Meng, Junsong yuan, Gang Wang, "object instance search in videos", in 9th IEEE International Conference on information, Communications and Signal Processing (ICICS), Taiwan 2013 Dec 10- 13, pp 1 – 4.
[30] A. Sharma, P. K. Singh and P. Khurana, "Analytical review on object segmentation and recognition," 2016 6th International Conference - Cloud System and Big Data Engineering (Confluence), Noida, 2016, pp. 524-530, doi: 10.1109/CONFLUENCE.2016.7508176.
[31] Ming Liang , Bin Yang, Shenlong Wang, and Raquel Urtasun, "Deep Continuous Fusion for Multi-Sensor 3D Object Detection", Proceedings of the European Conference on Computer Vision (ECCV), 2018, pp. 641-656
[32] Li, B., Zhang, T., Xia, T.: Vehicle detection from 3d lidar using fully convolutional network. In: RSS (2016).
[33] Luo, W., Yang, B., Urtasun, R.: Fast and furious: Real time end-to-end 3d detection, tracking and motion forecasting with a single convolutional net. In: CVPR (2018)
[34] Yang, B., Luo, W., Urtasun, R.: Pixor: Real-time 3d object detection from point clouds. In: CVPR (2018)
[35] Li, B.: 3d fully convolutional network for vehicle detection in point cloud. In: IROS (2017)
[36] Yu, S.L., Westfechtel, T., Hamada, R., Ohno, K., Tadokoro, S.: Vehicle detection and localization on bird's eye view elevation images using convolutional neural network. In: SSRR (2017)
[37] VGG16 – Convolutional Network for Classification and Detection https://neurohive.io/en/popular- networks/vgg16/
[38] K. He, X. Zhang, S. Ren, and J. Sun, ''Spatial pyramid pooling in deep convolutional networks for visual recognition,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 37, no. 9, pp. 1904–1916, Sep. 2015.