



# On the performance evaluation of object classification models in low altitude aerial data

Payal Mittal<sup>1</sup> · Akashdeep Sharma<sup>1</sup> · Raman Singh<sup>2</sup> · Arun Kumar Sangaiah<sup>3,4</sup>

Accepted: 17 March 2022 / Published online: 5 April 2022

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2022

## Abstract

This paper compares the classification performance of machine learning classifiers vs. deep learning-based handcrafted models and various pretrained deep networks. The proposed study performs a comprehensive analysis of object classification techniques implemented on low-altitude UAV datasets using various machine and deep learning models. Multiple UAV object classification is performed through widely deployed machine learning-based classifiers such as K nearest neighbor, decision trees, naïve Bayes, random forest, a deep handcrafted model based on convolutional layers, and pretrained deep models. The best result obtained using random forest classifiers on the UAV dataset is 90%. The handcrafted deep model's accuracy score suggests the efficacy of deep models over machine learning-based classifiers in low-altitude aerial images. This model attains 92.48% accuracy, which is a significant improvement over machine learning-based classifiers. Thereafter, we analyze several pretrained deep learning models, such as VGG-D, InceptionV3, DenseNet, Inception-ResNetV4, and Xception. The experimental assessment demonstrates nearly 100% accuracy values using pretrained VGG16- and VGG19-based deep networks. This paper provides a compilation of machine learning-based classifiers and pretrained deep learning models and a comprehensive classification report for the respective performance measures.

**Keywords** Machine learning · UAV datasets · Convolutional neural networks · Object recognition · Deep learning

---

✉ Arun Kumar Sangaiah  
arunkumarsangaiah@gmail.com

Extended author information available on the last page of the article

## 1 Introduction

The UAV platform needs visual object classification as a core enabling technology for the deployment of diverse applications in the computing paradigm. The low-altitude aerial images are obtained from drones flying within a certain height from the ground. We have considered aerial images that are captured by drones flying approximately 100 m or less above the land. The applications of unmanned aerial vehicles (UAVs) include autonomous driving cars [1], object detection and classification [2], spotting violent crowd behaviors [3], traffic monitoring [4], and aerial terrain analysis [5]. Low-altitude aerial images retrieved from UAVs incorporate public safety in vehicle accidents [6], ship collisions [7], border-power lines [8], crowd surveillance [9], and energy inspection from solar farms [10]. Low-altitude aerial images in urban settings have different features than remote sensing or standard datasets. These present significant challenges for object classification for low-altitude UAV images, such as payload weight constraints and multiple overlapped or scale-oriented images [11]. In this paper, we perform object classification on multiple low-altitude aerial objects.

### 1.1 Motivation

Research on low-altitude datasets is relatively new, and this paper strives to experimentally compare research in low-altitude aerial datasets by evaluating the performances of leading deep learning methods for object classification. The advent of artificial intelligence technologies has led to a boost in drone-based technologies to perform a wide range of applications. In this paper, we compare machine- and deep learning-based approaches for five different classes of low-altitude aerial objects. The inherent characteristics of low-altitude aerial images are different from standard images, so the challenges encountered in this case are more complicated to solve. The classification algorithms show different behavior when applied to low-altitude aerial images. The versatile applications of UAVs including crowd surveillance [9], traffic monitoring [4], and autonomous navigation [1] are more feasible due to recently formed drone policies. It is worth studying multiobject classification models along with diverse applications in the case of low-altitude aerial images. We aim to provide a suitable model to perform classification in this unexplored domain. This study targeted young audiences working in low-altitude UAV images to compare machine and deep network choices for object classification. The recent technological advancement of the machine and deep learning fields employs visual tasks in which human experts are relatively less efficient in evaluating recognition outcomes with correct visualization. This paper is an attempt in this direction, and the significant offerings of this paper include the following:

- Comparison between machine learning-based classifiers and a deep handcrafted CNN for object classification in low-altitude aerial images.

- Comparison between a deep handcrafted CNN and pretrained deep models for object classification in low-altitude aerial images.
- Performance evaluation of machine learning-based classifiers and pretrained deep models for low-altitude UAV object classification.
- Provide a suitable choice from a machine learning-based classifier and pretrained deep model for recognizing objects in low-altitude aerial images.

The organization of this research paper is as follows: Sect. 2 highlights the challenges of low-altitude UAV objects, machine learning studies, and deep learning-based object classification techniques. Section 3 describes an experimental setup in which the methodology of classification algorithms in low-altitude UAV datasets, the training process, the evaluation parameters, and the description of the low-altitude UAV dataset are discussed. Section 4 analyzes the results obtained from machine learning-based classifiers and pretrained deep models with different parameters. The last section concludes the results achieved and predicts the feasible choice of model for multi-object classification in low-altitude UAV datasets. The future scope of the proposed work is discussed in this section.

## 2 Related work

Over the last decade, convolutional neural networks (CNNs) have emerged as an optimal choice for a range of image manipulation tasks such as object detection, recognition [2], semantic segmentation, and pose estimation [12]. The real-time applications deployed in low-altitude UAV datasets do CNNs work in civilian airspace in a robust manner. The development of complex applications in low-altitude aerial images includes crowd surveillance by estimating violent human poses [12], recycling of plastic waste in wilds [13], monitoring power infrastructures [14], identifying mosquito breeding areas [15], and landslide accidents [16]. In this section, we discuss the challenges of low-altitude UAV-based object classification, machine learning-based classifiers, and deep models.

### 2.1 Challenges of UAV based object classification

Multiple object-based classification in low-altitude aerial images is a crucial problem due to overlapping image resolutions, limited contextual information, scale differences in objects, etc. There are significant challenges in low-altitude UAV-based object detection when related to standard images, such as:

1. Immense variations in the scale of aerial objects.
2. Dense distribution of small objects.
3. Arbitrary orientations of objects in low-altitude aerial images.
4. High illumination underexposes the dark regions of high-resolution images.

## 5. Occlusion in the form of proximity with other present objects.

All the above-discussed challenges have led to object detection and recognition techniques in low-altitude aerial images that used deep features for processing. We first describe machine learning-based classifiers, then a handcrafted CNN model, and finally pretrained deep learning-based models. Object classification-based experiments were performed on models on a low-altitude aerial dataset.

### 2.2 Machine learning-based classifiers

The machine classifiers that have been implemented are K nearest neighbor (KNN), decision trees [17], random forests (RF) [18], and naïve Bayes [19]. These classifiers have become high-performance baseline models in object recognition systems in recent times [20]. K nearest neighbor, the classifier, is the oldest nonparametric algorithm with  $k$  neighbors, determined using a cross-validation vector on an input class. The decision tree classifier attempts to divide the features to yield a suitable generalization. Decision trees are widely used models for classification and numerical data, whereas nonlinear parameters do not affect their performance. In this case, the decision tree classifier is imported with a random state = 0 and then fit on training data into the classifier. The design of decision trees includes attribute selection and pruning method choices. Furthermore, the object is classified by considering the voted class from existing predictors [21]. The most frequently used attribute-related measures are the information gain ratio and Gini index. In a provided training set  $T$ , choosing one pixel at random belonging to some class  $C_i$ , the Gini index is depicted in Eq. (1), where  $f(C_i, T)/|T|$  belongs to the probability of the chosen scenario that belongs to class  $C_i$ .

$$\sum_j \sum_{j \neq i} (f(C_i, T)/|T|)(f(C_j, T)/|T|). \quad (1)$$

The machine learning-based RF classifier consists of a random combination of features at every node of a tree. RF is an ensemble of unpruned decision trees that are built on a bootstrap input using a variable subset. We utilize a random forest without hyperparameter tuning and clustering. The naïve Bayes classifier is based upon the maximum a posteriori principle that calculates probability using the Bayes theorem in Eq. (2):

$$P(C = c | x_1, \dots, x_n) = P(C = c)P(x_1, \dots, x_n | C = c). \quad (2)$$

This approach is extendable to multiple classes and assumes conditional independence. Naïve Bayes classifiers assign the most expected class described by its feature vector and learning through feature independence. We compared machine classifiers with a customized approach, i.e., a deep handcrafted learning-based CNN network in our working methodology. This is intended to design an efficient and lightweight network from the beginning rather than adapt an existing system for low-altitude aerial images. The breakthrough of machine learning-based classifiers is observed in image processing in providing optimized object recognition results.

[22] described a hybrid approach of detecting an object from UAV imagery using the Viola-Jones detection method and a histogram of oriented gradients (HOG) [23]-based support vector machine (SVM) classifier [24] used jointly. The proposed scheme adopted an orientation adjustment method that rotated the UAV image to align in the horizontal direction. The strategy further developed an integrated hybrid approach based on their detection speed to improve efficiency. [25] implemented a cascading classifier that concatenated online learning-based classifiers by exploiting multiscale HOG features. The dimensions of input features were drawn out in multi-scale HOG to supply better and richer information for aerial images. Reference [26] made use of the AdaBoost classifier through a sliding window method of region proposals with integrated channel descriptions to detect independent moving features from aerial views. Different segmentation techniques, such as contour extraction and blob extraction, were evaluated to reduce the merging similarity of motion clusters. References [27, 28] made use of scale-invariant feature transform (SIFT) descriptors [29] for keypoint extraction of vehicle objects in UAV imagery. The number of objects was given by the number of final vital points extracted by the SVM classifier for classification and merging processes. Different combinations of SIFT features with color and morphology were used to calculate detection and false alarms.

Inspired by the above works, we found it interesting to compare machine and deep approaches to classify low-altitude aerial images. A comprehensive explanation of CNN-based deep models for multiple aerial object classification is discussed in the next sections.

### 2.3 Deep learning-based classification models

In the recent era, artificial intelligence has proven to be a revolution in machine learning in computer vision [30]. Later, an advancement of deep learning-based models evolved in image processing, which achieved tremendous object recognition results over traditional approaches in an effective manner [31]. CNNs have been the most successful object classification architectures in deep learning and work analogously to the human brain and embrace neurons that respond to the real-time environment [32]. Deep learning-based well-known CNN architectures have been deployed for object classification-based feature extractors for tuning the classifiers. The training is processed in which filters and parameters have random seeds by performing forward propagation. In low-altitude aerial studies, 2D-based CNNs have been commonly used to extract spatial features from the dimensions for object detection, recognition, and semantic segmentation of high-resolution aerial images [33], medical image-based disease diagnosis [34], and COVID-related measures [35]. Reference [34] proposed a VGG-inspired classification network to study the attention mechanisms for Alzheimer's disease. Eighteen-way data augmentation is proposed to avoid overfitting. The precision and accuracy were  $97.87 \pm 1.53$  and  $97.76 \pm 1.13$ , respectively. Reference [35] identified COVID-19 patients through a novel artificial intelligence model on a chest CT dataset. A novel VGG-style base network was proposed as a backbone network, and a convolutional block attention module was introduced as an attention module. Furthermore, an improved multiple-way data augmentation

method was used to resist overfitting. The proposed model achieved a precision per class above 95% and yielded a micro averaged F1 score of 96.87%, which is higher than 11 state-of-the-art approaches. Reference [36] improved building extraction accuracy in multifaceted building areas through a framework that applies deep learning-based semantic segmentation to UAV images with a digital surface model. The combination identified small buildings that were usually not high and covered partly by tree branches. The proposed method is applied to an open standard dataset to evaluate its strengths, and the results indicate an overall 4% accuracy increase from RGB to RGBD. Reference [37] compared the classification results of three deep models, AlexNet, VGG16, and VGG19, for ten classes of UAV landing sites with respect to different performance parameters. The results offered an understanding of typical false objects among classes of landing sites. Reference [38] proposed a dual inspection mechanism that identified missed targets in suspicious areas to assist single-stage detection branches in producing reliable results. The proposed method improved 2.7% mAP on the VisDrone2020 dataset, 1.0% mAP on the UAVDT dataset, and 1.8% mAP on the MS COCO dataset. Reference [39] provided a review on vehicle detection from UAV imagery using deep learning techniques such as convolutional neural networks, recurrent neural networks, autoencoders, generative adversarial networks, and their impact on improving the vehicle detection task. Reference [40] introduced a novel deep learning CNN architecture to identify anthracnose disease in mangos. A real-time dataset captured in farms of Karnataka, Maharashtra, and New Delhi was used for validation. In comparison with other state-of-the-art approaches, the proposed algorithm gives a higher classification accuracy of approximately 96.16%. Reference [41] evaluated the usage of transfer learning and fine-tuning on several CNN architectures, and the highest accuracy score was obtained by fine-tuning the ResNet50 model, which was 88%. The testing results show that transfer learning helps in generalization and demonstrates strong potential for the real-time application of forest fire detection.

CNNs were explicitly designed for object classification tasks, i.e., assigning single- or multiple-class labels to an entire scene. A breakthrough development in object classification was the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012, where multiple CNNs outperformed the state-of-the-art models based on handcrafted appearance descriptors [42]. Unusual extensions, such as trainable layers, increasing the capacity of the models [43], the introduction of dropout [44], batch normalization [45], and other strategies allowing better propagation of gradients, such as rectified linear unit (ReLU)-based nonlinearities [46], allow efficient training of deeper CNNs. The correctly annotated datasets for training and inference with powerful GPUs made CNNs the right standard for solving object classification problems. The classification of low-altitude UAV images having multiple categories of objects is the primary offering of this study. Pretrained deep models such as VGG16 [43], InceptionV3 [47], ResNet50 [48] and DenseNet121 [49] trained on ImageNet have been implemented for diverse object classification. Furthermore, detailed information about the number of parameters, accuracy rates, and required image size for pretrained deep models is presented in Table 1. The VGG-D models consisted of VGG16 and VGG19 with 13 and 16 convolutional layers, respectively. Their training was regularized by several regularization mechanisms,

**Table 1** Parameters of pre-trained deep models

Deep networks	Image size	No. of parameters (M)	Top-1 accuracy (%)	Top-5 accuracy (%)
VGG16	224×224	138	74.4	91.9
VGG19	224×224	144	74.5	92
InceptionV3	299×299	23.8	78.8	94.4
InceptionResnetV2	299×299	55.8	80.1	95.1
Xception	299×299	22.8	79	94.5
DenseNet121	224×224	33	76.39	93.34

especially for fully connected layers. InceptionV3 [47] eliminates several connections between convolutional layers that are unsuccessful and have redundant information due to the correlation between them. Inception-ResNetV2 [48] takes advantage of both Inception and ResNet networks and outperforms leading deep models. The Xception [54] architecture is built on a linear stack of a depth-wise separable convolution layer with linear residual connections. There are two important layers in architecture: a depth-wise convolutional layer in which a spatial convolution is carried out independently in each channel of input data. A pointwise convolutional layer has a  $1 \times 1$  convolutional layer, which maps the output channels to a new channel space using a depth-wise convolution. The DenseNet [49] network was designed to address the vanishing gradient problem arising from the network depth. The problem of training exists with every deep network due to the large flow of information and gradients. These models were initially trained on ImageNet, and then feature extraction was performed on customized low-altitude UAV datasets by transferring weights only to initial layers.

### 3 Experimental setup

We have considered multiple classes of objects in low-altitude aerial images in which object classification-based experiments have been performed. The methodology of applying machine learning-based classifiers on a low-altitude aerial dataset includes importing the necessary Python libraries, loading image files with their classes, scaling and transforming training and test data, instantiating the classification model, fitting the visualizer, and the model, and evaluating the model on the test data. The discussed machine learning-based classifiers and pretrained deep networks are trained on a customized low-altitude UAV dataset for multiple-object classification. The description of the dataset, training strategies, and performance evaluation methods are presented in subsequent sections.

#### 3.1 Deep network-based handcrafted model

An end-to-end deep object classification model has been trained on multiple objects presented in low-altitude aerial datasets known as a deep handcrafted model. The

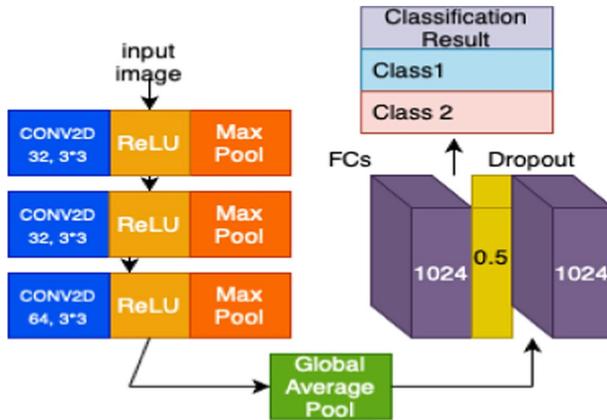


Fig. 1 Used handcrafted deep network

Table 2 Architectural details of handcrafted deep network for object recognition

	Layer type	Activation	Filters	Size	Stride
0	Conv2D	Relu	32	3×3	1
1	Max-Pool				2
2	Conv2D	Relu	32	3×3	1
3	Max-Pool				2
4	Conv2D	Relu	64	3×3	1
5	Max-Pool				2
6	GlobalAveragePooling				
7	Dense	Relu		1024	
8	Dropout			0.5	
9	Dense	Softmax		10	

architectural details of the proposed handcrafted-based model are described in Fig. 1 and Table 2. The network contained six convolutional and pool layers with a size of 150×150 as input images. The low-altitude aerial images with different dimensions were resized before feeding into the proposed algorithm. The filters were used to learn different feature types, and each filter slid over the input images. The layers after a convolution layer in the proposed architecture are global average pooling, dropout, and fully connected layers. The flattened layer converts 3D feature maps to 1D feature vectors. The activation function is ReLU, which accomplishes the threshold operation on the input to purge the effect of dark and noisy regions. Max and GlobalAveragePooling applied a maximum and average operation to each filter by restoring the spatial information of the images. The class values are calculated through a softmax classifier, and activation values correspond to diverse abstraction layers. The top layers of the model consisted of a softmax function class layer

that resulted in function output, and hence, the class layer selects the label with the determined probability.

### 3.2 Training process

The machine learning classifiers are implemented through Python's Scikit-learn library to use the customized low-altitude aerial dataset, which consists of images and corresponding labels. The task is to forecast the low-altitude aerial class to which the related images belong. During the training process, the loading of the dataset takes place, after which the splitting of the dataset into its attributes and labels is performed. The standard scaler function is employed before splitting the data into training and testing as it transforms the data. The final step is to calculate inferences on testing data. The classification report method is utilized to calculate precision, recall, and F-1 score metrics over the employed models. Deep learning-based architectures have been implemented in Keras with a TensorFlow1.10 version backend. We utilized uniform standard data shuffling techniques in all our experiments, including random horizontal, vertical flipping, random scaling, and rotations of the input data images. The input data are shuffled randomly and further split into training and validation (3:1 ratio) for passing into deep learning-based classification models. The same process is repeated multiple times so that a fair evaluation of data can be inferred. Root mean square propagation (RMSProp) was employed to optimize the network loss function, starting with a learning rate of 0.001. The training of each employed network is performed for 1000 epochs. In our case of multiclass classification of low-altitude UAV images, the categorical cross-entropy loss function provides a stable network and significant results. The dropout rate is 0.2 as a regularization technique for deep neural networks, and a batch size of 32 is kept due to the size of the input data. The final trained model was saved to disk for further visualization of the results. Computing on a cluster of 2 NVIDIA Titan XP GPUs was performed for training and validation inputs. Throughout the experiments, platforms of an Ubuntu 16.04 LTS-based Intel Core i7-6850 K CPU @ 3.60 GHz × 12 and 64 GB RAM are used. The main components of the proposed analysis are implemented using the Python language, supported with Sklearn [50], OpenCV libraries [51], Keras [52], and the TensorFlow backend [53]. The deep models utilized the various pretrained CNNs [43, 47, 48], partially fine-tuned with a widely deployed dataset, and implemented with NVIDIA-CUDA toolkits [55] to run on desktop graphical processing units (GPUs).

### 3.3 Evaluation parameters

To evaluate the accuracy of each deep model, popular classification-related evaluation metrics have been employed to visualize the results precisely. The classification report was generated from the predicted data to measure recall, precision, and F-1 score. The metric precision means the fraction of the true positives from the total sum of true positives and false positives. Recall means the fraction of true positives from the total number of true positives and false negatives. The  $F_1$  score describes the harmonic mean of precision and recall.

$$\text{Precision} = \frac{\text{True positives}}{\text{True positives} + \text{False positives}}, \quad (3)$$

$$\text{Recall} = \frac{\text{True positives}}{\text{True positives} + \text{False negatives}}, \quad (4)$$

$$F_1 \text{ measure} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (5)$$

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}}. \quad (6)$$

The accuracy score is the true predictions from the class one having the maximum probability and metrics have been represented in Eqs. 3–6.

### 3.4 Description of low-altitude UAV dataset

We have considered annotated low-altitude UAV datasets such as CARPK [56], Okutama [57], VEDAI [58], and UAVBD [59] and combined them to form five different categories of multiple objects in a single image. A wide variety of low-altitude UAV datasets have been merged to produce multiple classes of objects, such as vehicles, persons, cars, plastic bottles, etc. The description, annotation support, and dataset size-related information are presented in Table 3. The CARPK dataset [56] provides localization and counting of car objects in the parking lot to gather free space information for new entrants. The UAVBD dataset [13] is dedicated to procuring waste plastic bottles from mountains and wild grasses for recycling from a drone's view. The Okutama dataset [57] is specifically dedicated to human action detection between different humans as objects. The Birds dataset [59] captured at a low resolution of 25 pixels from cameras and telephoto lenses detects birds in wind farms for ecological conservation. This combined dataset has five different classes and sizes named birds, cars, persons, bottles, and vehicles, as depicted in Fig. 2. All the above classes make a total of 5000 low-altitude UAV images for implementing machine and deep learning-based classification models. The resizing of the original image was performed according to a pretrained network size, such as  $224 \times 224$  for VGG and  $299 \times 299$  for the Xception model. The low-altitude image data have been shuffled to maximize accuracy, and the performance comparisons of the various machines, as well as deep network-based methods, are made from UAV datasets. The next section describes a training process of multiple object classification in low-altitude UAV datasets.

**Table 3** Low-altitude UAV dataset for object recognition

Dataset	No. of image	Classes	Annotation support	Description
CARPk [56]	1447	Cars	Horizontal BB	Counts and localize target cars in videos
OKUTAMA [57]	77.4 k	Person	Horizontal BB	Concurrent Human action detection
UAV-BD [13]	25,394	Plastic bottle	Oriented BB	Localize waste plastic bottles in the wild spaces
VEDAI [58]	1272	Vehicle	Oriented BB	Small vehicle detection
Birds [59]	4959	Birds	Horizontal BB	Detects birds in wind farms for ecological conservation



**Fig. 2** Multiple classes of low-altitude UAV dataset

## 4 Results and discussion

In this section, a comprehensive quantitative analysis is proposed concerning various machine classifiers and deep learning architectures to predict urban objects in low-altitude aerial images. The experiments suggest that the handcrafted CNN achieved a maximum accuracy score of 92.48 compared with machine classifiers. Out of the KNN, naïve Bayes, decision trees, and random forest classifiers, random forests obtained the highest value of 90% on low-altitude aerial data. Our experimental results helped to conclude that deep networks provide the right choice for achieving significant improvements in low-altitude aerial image-based classification. The overall accuracy score of the handcrafted CNN (92.48), as shown in Table 4, is higher than machine-based classifiers in Table 6. The performance of the handcrafted CNN degraded when compared with pretrained networks. The deep network models are trained for various input sizes of multiple low-altitude aerial datasets. Deep network architectures such as VGG16, VGG19, InceptionV3, Xception, DenseNet121, and InceptionResNetV2 were utilized to perform the experiments. The acquired dataset of low-altitude aerial images was resized to  $224 \times 224$  for the VGG16 & 19 and DenseNet121 networks and to  $299 \times 299$  for the InceptionV3, Xception and InceptionResNetV2 networks.

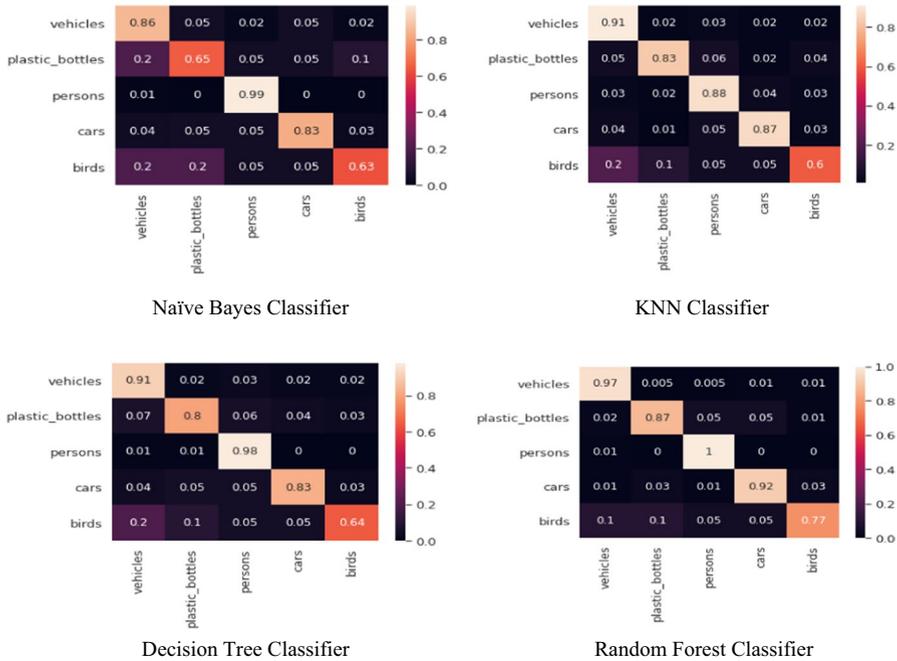
### 4.1 Analysis of performance metrics

In this section, an analysis of performance metrics such as precision, recall, and F-1 score evaluation has been discussed. The confusion matrices for each machine learning-based classifier are utilized to better understand true positives and false positives for multiobject classification in low-altitude aerial images. Table 4 represents the

**Table 4** Performance results for handcrafted CNN

Class	Precision	Recall	F-1 score
Vehicles	0.96	0.85	0.90
Plastic bottles	0.85	0.93	0.89
Persons	0.96	1.00	0.98
Cars	0.99	0.82	0.90
Birds	0.89	1.00	0.94

**Table 5** Confusion matrix of machine learning-based classifiers

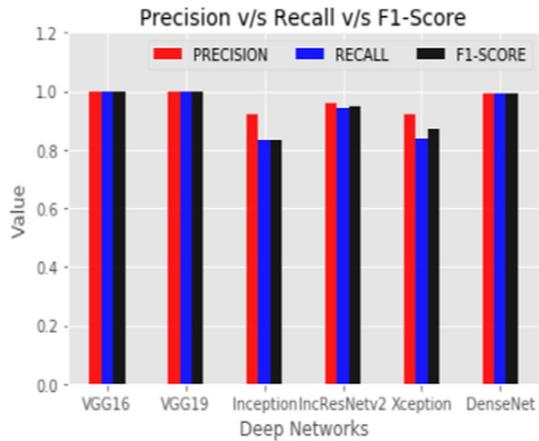


**Table 6** Comparison of classification accuracy in machine learning-based classifiers

Classifier	Individual Class Accuracy Metrics				Classifier	Individual Class Accuracy Metrics			
<b>K Nearest Neighbor Accuracy 82.66</b>	<b>Class</b>	<b>Precision</b>	<b>recall</b>	<b>F1 score</b>	<b>Decision Tree Accuracy 83.26</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>
	Vehicles	0.84	0.99	0.91		Vehicles	0.90	0.92	0.91
	Plastic bottles	1.00	0.70	0.83		Plastic bottles	0.81	0.79	0.80
	Persons	0.78	1.00	0.88		Persons	0.98	0.98	0.98
	Cars	0.85	0.89	0.87		Cars	0.85	0.81	0.83
	Birds	0.68	0.54	0.60		Birds	0.62	0.65	0.64
<b>Naïve Bayes Accuracy 79.00</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>	<b>Random Forest Accuracy 90.00</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>
	Vehicles	0.89	0.82	0.86		Vehicles	0.98	0.96	0.97
	Plastic bottles	0.76	0.57	0.65		Plastic bottles	0.89	0.85	0.87
	Persons	1.00	0.99	0.99		Persons	1.00	1.00	1.00
	Cars	0.85	0.82	0.83		Cars	0.90	0.94	0.92
	Birds	0.54	0.77	0.63		Birds	0.77	0.78	0.77

confusion matrices for the KNN, naïve Bayes, decision tree, and random forest classifiers. The diagonal values in the matrix represent the true predictions out of the total samples. The evaluation of performance metrics in the case of machine classifiers and deep learning-based networks has been done for low-altitude aerial images. Parameters such as precision, recall, F-1 score, and accuracy score were calculated from the classification report (Table 5). Detailed visualization of the classification

**Fig. 3** Performance of various deep learning models

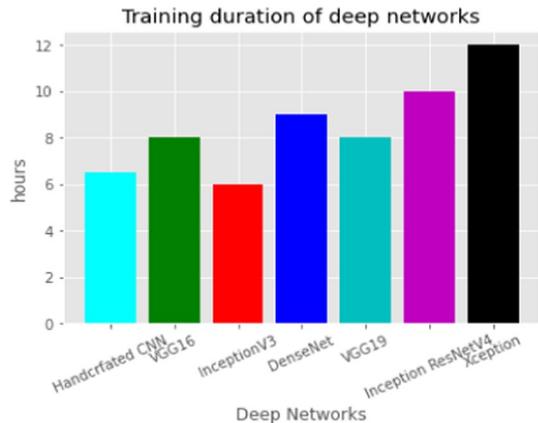


**Table 7** Comparison of classification accuracy between pretrained deep CNNs

Pretrained CNNs	Individual Class Accuracy Metrics				Pretrained CNNs	Individual Class Accuracy Metrics			
<b>Inception ResnetV2</b> Accuracy <b>0.9864</b> Loss <b>0.2041</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>	<b>Inception V3</b> Accuracy <b>0.96</b> Loss <b>0.5740</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>
	Vehicles	1.00	0.93	0.96		Vehicles	1.00	0.80	0.89
	Plastic bottles	0.96	1.00	0.98		Plastic bottles	0.90	1.00	0.94
	Persons	0.99	1.00	1.00		Persons	0.96	1.00	0.98
	Cars	0.99	1.00	0.99		Cars	0.96	1.00	0.98
	Birds	1.00	1.00	1.00		Birds	1.00	1.00	1.00
<b>VGG16</b> Accuracy <b>1.00</b> Loss <b>0.0012</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>	<b>VGG19</b> Accuracy <b>0.9968</b> Loss <b>0.0317</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>
	Vehicles	1.00	1.00	1.00		Vehicles	1.00	0.99	0.99
	Plastic bottles	1.00	1.00	1.00		Plastic bottles	0.99	1.00	1.00
	Persons	1.00	1.00	1.00		Persons	1.00	1.00	1.00
	Cars	1.00	1.00	1.00		Cars	1.00	1.00	1.00
	Birds	1.00	1.00	1.00		Birds	1.00	1.00	1.00
<b>DenseNet121</b> Accuracy <b>0.9968</b> Loss <b>0.0414</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>	<b>Xception</b> Accuracy <b>0.9864</b> Loss <b>0.1680</b>	<b>Class</b>	<b>precision</b>	<b>recall</b>	<b>F1 score</b>
	Vehicles	1.00	0.98	0.99		Vehicles	1.00	0.94	0.97
	Plastic bottles	1.00	1.00	1.00		Plastic bottles	1.00	1.00	0.98
	Persons	1.00	1.00	1.00		Persons	0.98	1.00	0.99
	Cars	0.99	1.00	0.99		Cars	0.96	1.00	1.00
	Birds	1.00	1.00	1.00		Birds	1.00	0.99	0.99

report of machine learning classifiers and handcrafted CNN-based object classification models with individual classes of low-altitude UAV datasets are displayed in Tables 4 and 6, respectively. Furthermore, classification performance metrics with respect to low-altitude UAV objects are presented in these tables. The precision, recall, and F-1 score of each machine classifier and deep learning-based handcrafted

**Fig. 4** Training duration of various deep learning models



CNN model were combined to depict that the deep handcrafted model performed better than machine classifiers. The detailed visualization of the classification report of deep learning-based object classification models with individual classes of low-altitude UAV datasets is displayed in Fig. 3 and Table 7. The experiments suggest that the Xception model needed maximum time when trained for the required number of epochs, as depicted in Fig. 4. The VGG16 and VGG19 models converged quickly; after that, stagnant performance was seen. The analyzed deep networks depicted different behavior when trained on low-altitude aerial datasets compared to standard images. Xception, DenseNet121 and InceptionResNetV2 performed better than InceptionV3 in terms of evaluation parameters. Our experimental results helped gather deep network choices for multiple class-based object classification problems in low-altitude aerial images. The value of the handcrafted-based CNN (92.48) is found to be higher than machine learning-based classifiers such as KNN (82.26), naïve Bayes (83.26), decision trees (79), and random forests (90). Our findings concluded that training a handcrafted deep neural network is feasible compared with machine classifiers, as the accuracy obtained by the CNN (92.48) is higher than each employed machine classifier. High performance has not been achieved, as the discussed machine learning-based classifiers face problems in the case of low-altitude UAV images, such as [26]:

- The features obtained from manual work relying on aerial domain knowledge may not be adequate for object recognition tasks.
- Handcrafted feature engineering is a time-consuming process and quite tedious.
- Machines with related mathematical models and assumptions restrict the flexibility to handle aerial image shapes.

The pretrained networks perform even better than deep handcrafted networks because the handcrafted CNN started with randomly initialized dynamic weights. In contrast, pretrained networks trained on a large ImageNet dataset provide better end-to-end learning. In addition, pretrained deep model even performed better than the

handcrafted CNN and machine models on the same dataset. This is because of the training of the model's weights. The six kinds of pretrained transfer learning-based deep networks show different multiple object recognition results when compared with previous findings. Inception-ResNet-v2 achieved an accuracy of 98.64 and a loss of 0.2041, the same as that of the Xception network. Accurate models such as InceptionV3 obtained 96.00 accuracy and 0.5740 loss, which states that Inception-ResNetV2 has an improved network over InceptionV3 in our settings. Xception also performed better than InceptionV3, with an accuracy score of 98.64%. The recently developed DenseNet121 also showed significant performance due to concatenation of input layers to produce an output layer with an accuracy of 99.68 and loss value of 0.0414.

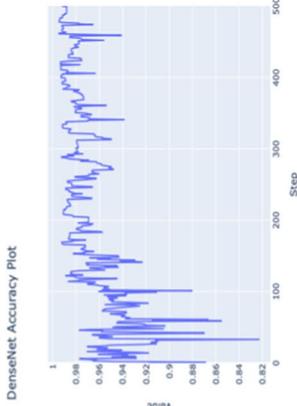
## 4.2 Comparisons of accuracy and loss graphs

The training process of deep networks for multiple object recognition was executed for 500 epochs. For each epoch, a summary of accuracy and loss is generated, and thus graphs obtained from TensorBoard related to deep networks are presented in Figs. 5 and 6. The plots depicted in Fig. 6 show that the validation accuracy models seem to have converged. The line plots for both accuracy and loss show good convergence behavior, although they are somewhat bumpy. All described models are well configured and show no signs of over- or underfitting. The loss and accuracy values depicted almost no convergence after 400 epochs, from which we can assume that the model is trained. DenseNets performed fairly well in multiple object-based UAV datasets and achieved 99.68% accuracy. High convergence can be seen in the accuracy plot of Inception-ResNetV2 due to the learning capacity of the network. InceptionV3 did not perform well in our settings and obtained a loss value of 0.5714, which is higher than other pretrained deep networks. Xception performed better than the InceptionV3 network but relatively poorly when compared with other deep networks trained on low-altitude UAV datasets. The value of loss and accuracy depicted no convergence after 200 epochs, and both VGG16 and VGG19 models performed best on the low-altitude UAV dataset.

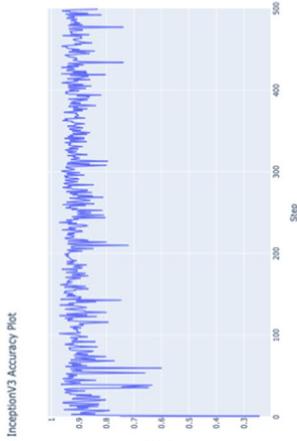
The comparison with the state-of-the-art studies mentioned in Table 8, [60, 63, 64] made use of descriptor-based classification methods. These methods require hand engineering and complex methodology. [66] employed hyperspectral images by developing a hail vegetation index to identify agriculture-based patterns. Our dataset contains multiple size objects, and the impressive results of the VGG networks revealed that the network depth is an important factor in obtaining high classification accuracy. The evaluation presented in Fig. 7 indicates that deep networks trained on standard images have a different scope than those trained on low-altitude aerial views. Due to the inherent characteristics of low-altitude aerial images, such as the small size of objects, captured angle, resolution, orientation, and scale, they differ from natural images.

### Analysis of Accuracy Graphs

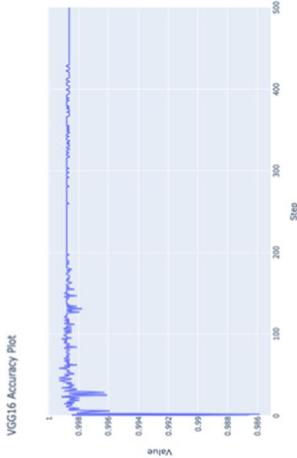
**Densenet121 (a)**



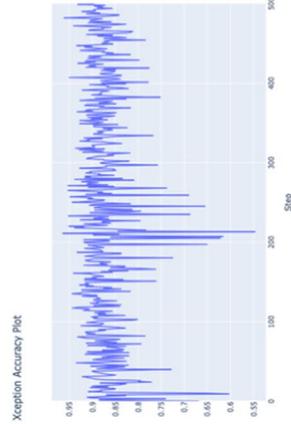
**InceptionV3 (b)**



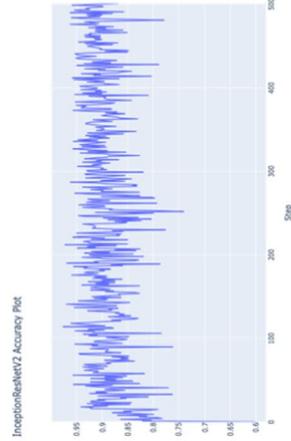
**VGG16 (c)**



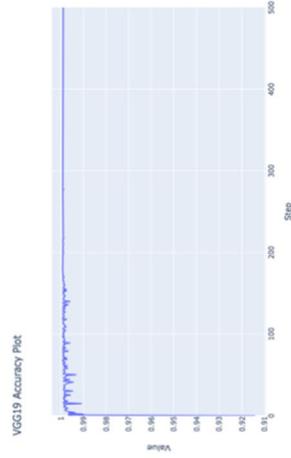
**Xception (d)**



**InceptionResnetV2(e)**



**VGG19 (f)**



**Fig. 5** Validation accuracy graphs of deep learning models

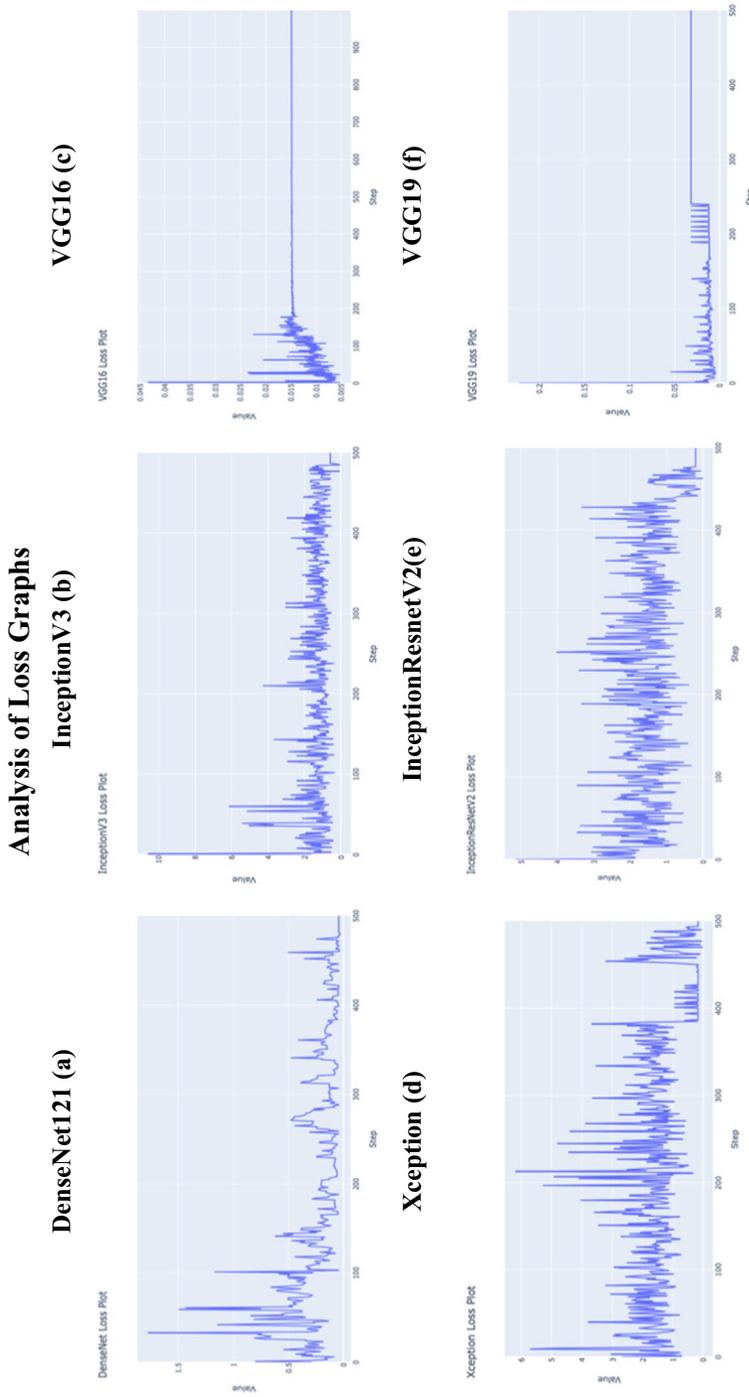
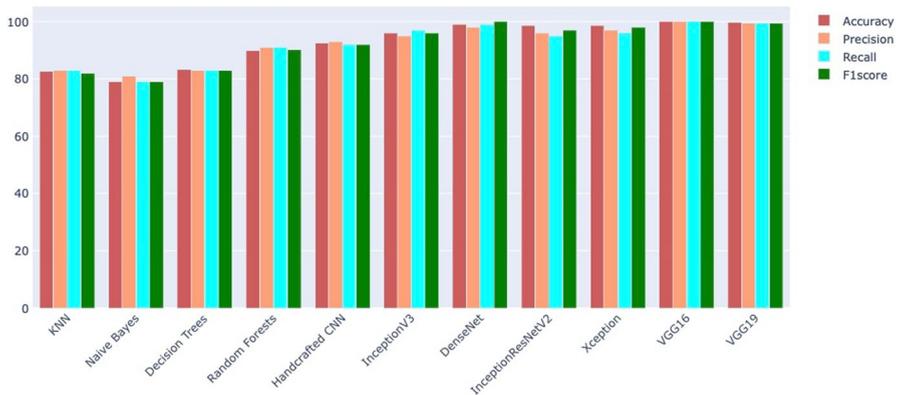


Fig. 6 Validation loss graphs of deep learning models

**Table 8** Comparison with existing classification methods

Studies	Methods	Datasets	Best Results
[60]	Handcrafted shape and color features	Flavia	96.30
[61]	GoogleNet and micro-Doppler signature based classification	Merging MDS and CVD	94.7
[62]	CNN, SVM	TARBIL	97.47
[63]	Fourier and wavelet descriptors, MLP	Flavia	97.50
[64]	Pixel-based image classification method	Fixed wing drone dataset	84.7
[65]	AlexNet/LDA	UCI Leaf	96.20
[66]	Hail Vegetation Index	Hyperspectral Images	> 90
Our work	VGG16/VGG19	Instances of VEDAI, CARPK, OKUTAMA, BIRDS	99.68

**Fig. 7** Performance comparison of all algorithms

## 5 Conclusion

This paper has analyzed various machine learning- and deep learning-based classification networks to recognize multiple objects from low-altitude UAVs. The proposed evaluation compares machine classifiers KNN, naïve Bayes, random forest, decision trees and deep models such as handcrafted-based CNN, VGG16, VGG19, InceptionV3, Xception, DenseNets, etc. Machine- and deep model-based classification was performed to conduct experiments on low-altitude UAV images. Among the employed machine classifiers for classification, random forests achieved better results among KNN, decision trees, and naïve Bayes classifiers. However, when compared with a handcrafted CNN, the performance of leading machine classifier random forests degraded on low-altitude aerial images. In the case of pretrained deep models for object recognition, VGGD, InceptionV3, DenseNet121, Inception-ResNetV2, and Xception depicted different behaviors when trained on low-altitude

aerial datasets. DenseNet121 and Inception-ResNetV2 performed better than InceptionV3 and Xception. However, VGG16 and VGG19 performed better than Xception,

DenseNet121, and Inception-ResNetV2 due to the inherent characteristics of low-altitude data. Our experimental results provide academia and the research community with a medium for dealing with multiple object classification in low-altitude aerial images. The classification reports concerning individual class in terms of precision, recall, and F-1 score are represented to analyze models better.

The progressive approaches of deep learning-based object classification in low-altitude aerial data seem to have a bright future. The vast deployment of applications influenced the aerial imaging market, which is expected to grow at a rate of 14.2% in the coming years. One of the major factors creating advanced prospects in the aerial imaging classification solutions market is the recently published drone policies by the Government of India and the availability of artificial intelligence-based technologies. Furthermore, as a part of our future work, we intend to explore human activity recognition and detect abnormal behaviors in surveillance-based UAV applications.

**Acknowledgements** We acknowledge DIC, Panjab University Chandigarh for funding a workstation that enabled us to perform experiments in the form of NVIDIA TITAN XP GPUs. This research work has been done under UGC NET SRF scholarship, New Delhi, India.

**Funding** The authors did not receive support from any organization for the submitted work.

**Availability of data and material** Applicable on request.

**Code availability** Applicable on request.

## Declarations

**Conflict of interest** The authors have no conflicts of interest to declare that are relevant to the content of this article.

**Ethics approval** This paper does not contain any studies with human or animal subjects and all authors declare that they have no conflict of interest.

**Consent to participate** All authors declare that they have the consent to participate.

**Consent for publication** All authors declare that they have consent for publication.

## References

1. Mohanan MG, Salgoankar A (2018) A survey of robotic motion planning in dynamic environments. *J Robot Auton Syst* 100:171–185
2. Fernandes D, Silva A, Névoa R, Simões C, Gonzalez D, Guevara M, Novais P, Monteiro J, Melo-Pinto P (2021) Point-cloud based 3D object detection and classification methods for self-driving applications: a survey and taxonomy. *Inf Fusion* 68:161–191
3. Tzelepi M, Tefas A (2017) Human crowd detection for drone flight safety using convolutional neural networks. In: 25th European Signal Processing Conference (EUSIPCO), pp 743–747

4. Li L, Du B, Wang Y, Qin L, Tan H (2020) Estimation of missing values in heterogeneous traffic data: Application of multimodal deep learning model. *Knowl Based Syst* 194:105592
5. Du Terrail JO, Jurie F (2017) On the use of deep neural networks for the detection of small vehicles in ortho-images. In: *IEEE International Conference on Image Processing (ICIP)*, pp 4212–4216
6. Huang T, Wang S, Sharma A (2020) Highway crash detection and risk estimation using deep learning. *Accid Anal Prev* 135:105392
7. Ma J, Li W, Jia C, Zhang C, Zhang Y (2020) Risk prediction for ship encounter situation awareness using long short-term memory based deep learning on intership behaviors. *J Adv Trans* 2020:8897700. <https://doi.org/10.1155/2020/8897700>
8. Ren S, Choi TM, Lee KM, Lin L (2020) Intelligent service capacity allocation for cross-border-E-commerce related third-party-forwarding logistics operations: a deep learning approach. *Transport Res Part E Logist Transport Rev* 134:101834
9. Seema S, Goutham S, Vasudev S, Putane RR (2020) Deep learning models for analysis of traffic and crowd management from surveillance videos. In: *Progress in computing, analytics and networking*. Springer, Singapore, pp 83–93
10. Tang W, Yang Q, Xiong K, Yan W (2020) Deep learning based automatic defect identification of photovoltaic module using electroluminescence images. *Sol Energy* 201:453–460
11. Gaszczak A, Breckon TP, Han J (2011) Real-time people and vehicle detection from UAV imagery. In: *Intelligent robots and computer vision XXVIII: algorithms and techniques*, vol 7878, p 78780B
12. Singh A, Patil D, Omkar SN (2018) Eye in the sky: real-time drone surveillance system (DSS) for violent individuals identification using scatternet hybrid deep learning network. <http://arxiv.org/abs/1806.00746>
13. Wang J, Guo W, Pan T, Yu H, Duan L, Yang W (2018) Bottle detection in the wild using low-altitude unmanned aerial vehicles. In: *2018 21st International Conference on Information Fusion (FUSION)*, pp 439–444
14. Varghese A, Gubbi J, Sharma H, Balamuralidhar P (2017) Power infrastructure monitoring and damage detection using drone captured images. In: *2017 International Joint Conference on Neural Networks (IJCNN)*, pp 1681–1687
15. Amarasinghe A, Suduwella C, Elvitigala C, Niroshan L, Amaraweera RJ, Gunawardana K, Kumarasinghe P, De Zoysa K, Keppetiyagama C (2017) A machine learning approach for identifying mosquito breeding sites via drone images. In: *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, p 68
16. Li Z, Shi W, Lu P, Yan L, Wang Q, Miao Z (2016) Landslide mapping from aerial photographs using change detection-based Markov random field. *J Remote Sens Environ* 187:76–90
17. Quinlan JR (1986) Induction of decision trees. *Mach Learn* 1(1):81–106
18. Breiman L (2001) Random forests. *Mach Learn* 45(1):5–32
19. Kohavi R (1996) Scaling up the accuracy of naive-bayes classifiers: a decision-tree hybrid. *Kdd* 96:202–207
20. Li X (2007) Conference 9247: High-Performance Computing in Remote Sensing, *Remote Sens. Secur. Def. Technol.*, p 188
21. Xu Y, Yu G, Wang Y, Wu X, Ma Y (2016) A hybrid vehicle detection method based on viola-jones and HOG+ SVM from UAV images. *Sensors* 16(8):1325
22. Sugimura D, Fujimura T, Hamamoto T (2016) Enhanced cascading classifier using multi-scale HOG for pedestrian detection from aerial images. *Int J Pattern Recognit Artif Intell* 30(03):1655009
23. Mizuno K, Terachi Y, Takagi K, Izumi S, Kawaguchi H, Yoshimoto M (2012) Architectural study of HOG feature extraction processor for real-time object detection. In: *2012 IEEE workshop on signal processing systems*. IEEE, pp 197–202
24. Neumann J, Schnörr C, Steidl G (2005) Combined SVM-based feature selection and classification. *Mach Learn* 61(1–3):129–150
25. Teutsch M, Krüger W, Beyerer J (2014) Evaluation of object segmentation to improve moving vehicle detection in aerial videos. In: *11th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp 265–270
26. Moranduzzo T, Melgani F (2014) Automatic car counting method for unmanned aerial vehicle images. *IEEE Transactions on Geoscience Remote Sensing* 52(3):1635–1647
27. Xu B, Xu X, Own C-M (2017) On the feature detection of nonconforming objects with automated drone surveillance. In: *Proceedings of the 3rd International Conference on Communication and Information Processing*, pp 484–489

28. Fernández-Delgado M, Cernadas E, Barro S, Amorim D (2014) Do we need hundreds of classifiers to solve real world classification problems? *J Mach Learn Res* 15(1):3133–3181
29. Lindeberg T (2012) Scale invariant feature transform, pp 10491
30. Zhao B, Feng J, Wu X, Yan S (2017) A survey on deep learning-based fine-grained object classification and semantic segmentation. *Int J Autom Comput* 14:119–135
31. Baykara HC, Biyik E, Gül G, Onural D, Öztürk AS, Yildiz I (2017) Real-time detection, tracking and classification of multiple moving objects in UAV videos. In *IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI)*, pp 945–950
32. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*, pp 1097–1105
33. Liang Y, Monteiro ST, Saber ES (2016) Transfer learning for high resolution aerial image classification. In: *Applied imagery pattern recognition workshop (AIPR)*. IEEE, pp 1–8
34. Wang SH, Zhou Q, Yang M, Zhang YD (2021) ADVIAN: Alzheimer’s disease VGG-inspired attention network based on convolutional block attention module and multiple way data augmentation. *Front Aging Neurosci* 13:313
35. Wang SH, Fernandes S, Zhu Z, Zhang YD (2021) AVNC: attention-based VGG-style network for COVID-19 diagnosis by CBAM. *IEEE Sensors J*
36. Boonpook W, Tan Y, Xu B (2021) Deep learning-based multi-feature semantic segmentation in building extraction from images of UAV photogrammetry. *Int J Remote Sens* 42(1):1–19
37. Kuchár D, Schreiber P (2021) Comparison of UAV landing site classifications with deep neural networks. In: *Computer Science On-line Conference*. Springer, Cham, pp 55–63
38. Tian G, Liu J, Zhao H, Yang W (2022) Small object detection via dual inspection mechanism for UAV visual images. *Appl Intell* 52(4):4244–4257
39. Bouguettaya A, Zarzour H, Kechida A, Taberkit AM (2021) Vehicle detection from UAV imagery with deep learning: a review. *IEEE Transactions on Neural Networks and Learning Systems*
40. Kumar P, Ashtekar S, Jayakrishna SS, Bharath KP, Vanathi PT, Kumar MR (2021) Classification of mango leaves infected by fungal disease anthracnose using deep learning. In: *5th International conference on computing methodologies and communication (ICCMC)*. IEEE, pp 1723–1729
41. Treneska S, Stojkoska BR (2021) Wildfire detection from UAV collected images using transfer learning
42. Sommer L, Nie K, Schumann A, Schuchert T, Beyerer J (2017) Semantic labeling for improved vehicle detection in aerial imagery. In: *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp 1–6
43. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. <http://arxiv.org/abs/1409.1556>
44. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 1–9
45. Ioffe S, Szegedy C (2015) Batch normalization: accelerating deep network training by reducing internal covariate shift. <http://arxiv.org/abs/1502.03167>
46. Ramachandran P, Zoph B, Le QV (2017) Searching for activation functions. <http://arxiv.org/abs/1710.05941>
47. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA (2017) Inception-v4, inception-resnet and the impact of residual connections on learning. In: *Thirty-First AAAI Conference on Artificial Intelligence*
48. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp 770–778
49. Iandola F, Moskewicz M, Karayev S, Girshick R, Darrell T, Keutzer K (2014) Densenet: implementing efficient convnet descriptor pyramids. <http://arxiv.org/abs/1404.1869>
50. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Prettenhofer P, Weiss R, Dubourg V (2011) Scikit-learn: machine learning in Python. *J Mach Learn Res* 12:2825–2830
51. Bradski G, Kaehler A (2008) *Learning OpenCV: computer vision with the OpenCV library*. O’Reilly Media, Inc.
52. Gulli A, Pal S (2017) *Deep learning with Keras*. Packt Publishing Ltd
53. Abadi M, Barham P, Chen J, Chen Z, Davis A, Dean J, Devin M, Ghemawat S, Irving G, Isard M, Kudlur M (2016) Tensorflow: a system for large-scale machine learning. *OSDI* 16:265–283
54. Chollet F (2017) Xception: deep learning with depthwise separable convolutions. In: *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*

55. Kirk D (2015) NVIDIA CUDA software and GPU parallel computing architecture. In: ISMM, vol 7
56. Hsieh M-R, Lin Y-L, Hsu WH (2017) Drone-based object counting by spatially regularized regional proposal network. In: The IEEE International Conference on Computer Vision (ICCV), vol 1
57. Barekatin M, Martí M, Shih HF, Murray S, Nakayama K, Matsuo Y, Prendinger H (2017) Okutama-action: an aerial view video dataset for concurrent human action detection. In: 1st Joint BM3D-PETS Workshop on Tracking and Surveillance, CVPR, pp 1–8
58. Razakarivony S, Jurie F (2016) Vehicle detection in aerial imagery: a small target detection benchmark. *J Vis Commun Image Represent* 34:187–203
59. Yoshihashi R, Kawakami R, Iida M, Naemura T (2015) Construction of a bird image dataset for ecological investigations. In: 2015 IEEE International Conference on Image Processing (ICIP), pp 4248–4252
60. Caglayan A, Guclu O, Can AB (2013) A plant recognition approach using shape and color features in leaf images. In: International Conference on Image Analysis and Processing. Springer, Berlin, pp 161–170
61. Kim BK, Kang HS, Park SO (2016) Drone classification using convolutional neural networks with merged Doppler images. *IEEE Geosci Remote Sens Lett* 14(1):38–42
62. Yalcin, H. and Razavi, S (2016) Plant classification using convolutional neural networks. In: 2016 Fifth International Conference on Agro-Geoinformatics. IEEE, pp 1–5
63. Yousefi E, Baleghi Y, Sakhaei SM (2017) Rotation invariant wavelet descriptors, a new set of features to enhance plant leaves classification. *Comput Electron Agric* 140:70–76
64. Liu H, Qu F, Liu Y, Zhao W, Chen Y (2018) A drone detection with aircraft classification based on a camera array. *IOP Conf Ser Mater Sci Eng* 322(5):052005
65. Kaya A, Kececi AS, Catal C, Yalic HY, Temucin H, Tekinerdogan B (2019) Analysis of transfer learning for deep neural network based plant classification models. *Comput Electron Agric* 158:20–29
66. Yang W, Xu W, Wu C, Zhu B, Chen P, Zhang L, Lan Y (2021) Cotton hail disaster classification based on drone multispectral images at the flowering and boll stage. *Comput Electron Agric* 180:105866

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## Authors and Affiliations

Payal Mittal<sup>1</sup> · Akashdeep Sharma<sup>1</sup> · Raman Singh<sup>2</sup> · Arun Kumar Sangaiah<sup>3,4</sup>

Payal Mittal  
payalmittal6792@gmail.com

Akashdeep Sharma  
akashdeep@pu.ac.in

Raman Singh  
raman.singh@thapar.edu

<sup>1</sup> UIET Panjab University, Chandigarh, India

<sup>2</sup> Thapar Institute of Engineering and Technology, Patiala, India

<sup>3</sup> School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, India

<sup>4</sup> Department of Industrial Engineering and Management, National Yunlin University of Science and Technology, Douliu, Taiwan