Drone Detection by Neural Network Using GLCM and SURF Features

Tanzia Ahmed
Department of Computer Science and Software Engineering, Concordia University, Montreal, QC, Canada
tanzia.ahmed@mail.concordia.ca
Tanvir Rahman
Department of Computer Science and Engineering, Brac University, Dhaka, Bangladesh
tanvir.rahman@bracu.ac.bd
Bir Ballav Roy
Department of Computer Science and Engineering, Brac University, Dhaka, Bangladesh
roybirballav@gmail.com
Jia Uddin*
Technology Studies Department, Endicott College, Woosong University, Daejeon, South Korea
jia.uddin@wsu.ac.kr

Abstract

This paper presents a vision-based drone detection method. There are a number of researches on object detection which includes different feature extraction methods – all of those are used distinctly for the experiments. But in the proposed model, a hybrid feature extraction method using SURF and GLCM is used to detect object by Neural Network which has never been experimented before. Both are very popular ways of feature extraction. Speeded-up Robust Feature (SURF) is a blob detection algorithm which extracts the points of interest from an integral image, thus converts the image into a 2D vector. The Gray-Level Co-Occurrence Matrix (GLCM) calculates the number of occurrences of consecutive pixels in same spatial relationship and represents it in a new vector- 8×8 matrix of best possible attributes of an image. SURF is a popular method of feature extraction and fast matching of images, whereas, GLCM method extracts the best attributes of the images. In the proposed model, the images were processed first to fit our feature extraction methods, then the SURF method was implemented to extract the features from those images into a 2D vector. Then for our next step GLCM was implemented which extracted the best possible features out of the previous vector, into a 8×8 matrix. Thus, image is processed in to a 2D vector and feature extracted from the combination of both SURF and GLCM methods ensures the quality of the training dataset by not just extracting features faster (with SURF) but also extracting the best of the point of interests (with GLCM). The extracted featured related to the pattern are used in the neural network for training and testing. Pattern recognition algorithm has been used as a machine learning tool for the training and testing of the model. In the experimental evaluation, the performance of proposed model is examined by cross entropy for each instance and percentage error. For the tested drone dataset, experimental results demonstrate improved performance over the state-of-art models by exhibiting less cross entropy and percentage error.

Keywords: Feature Extraction; GLCM Method; Image Processing; Neural Network; SURF Algorithm.

1- Introduction

Drones are, in technical terms, unmanned aircraft. These are also known as unmanned aerial vehicles, or UAVs. Usually they are controlled by remote controlling systems, however, there are some drones that can fly by themselves. While there are numerous merits of drones that be uttered here, drones are also being used in several crimes these days. Crimes are being sophisticated day by day and this is just one angle of it. For instance, it is used for providing unwanted materials to prison [Telegraph, February 16, 2016]. FBI spokesperson says, drug cartel activists are being replaced by drones as there is little fear of getting arrested [1].

Image processing in the field of object detection is getting momentums, and for good reasons. The drones that are being used for crimes need to be detected while they are still in the air and prior to the time crime takes place. Traditional CCTV based security system is in cry of update, hence image processing can come in handy. For that, techniques of image processing have to be chosen carefully for extracting and analyzing features of the objects.

Object detection is one of the major divisions in the field

of computer vision and there are many researches that have been conducted in this area. Object detection has been done using deep learning. The Support Vector Machine (SVM) was used for extracting feature in the field of emotion recognition [2], and also used for dimension reduction in the fields of machine learning [3]. Yan *et. al* detected anomalies using SVM [4]. SVM outperforms then the other state-of-art models in the field of object recognition using entropy theory [2-4]. Blob detection has also been quite popular in the field of feature extraction. Barbosa *et al.* showed a tool to extract meta-data for game sprite using Blob detection aka edge detection. Neural Networks are very sensitive, even to the lowest change of any properties of an object [5]. Often times, it may lead to inefficient generalization of their results. Tay and Lao present how the use of SVM leads to inefficient generalization in the field of financial time series forecasting [6].

Image processing has come a long way from detecting the overall features of a particular image, it is now a field that often works to detect the distinct features. However, this development has its own challenges. Different image processing systems have several challenges as to how they should be analyzed. Kumar and Bhatia showed how to use Fourier analysis to analyze the shapes. Moreover, they also showed the importance of Gray Scale character in both rotation variant and rotation invariant [7]. However, all distinct information may not be relevant. Choras showed how the relevant information can detect and identify the similar images in the field of biometrics. In [8], Content Based Image Retrieval (CBIR) is used for detecting the similarities in the images. Some basic architecture of ZF-NET, and deep normalization and convolutional layers (DNCNN) may use for automatic extracting features. Yin *et al.* suggested a model that shows the previous phenomenon [9]. In [10], authors present an analysis of various ways to recognize the aerial components from images taken by drones on power transmission lines using the neural network and SURF (Speeded-up Robust Features) and BoW (Bag-of-Words) methods as a feature extraction. In [11], Abuzneid *et. al* proposed an enhanced technique of face recognition using traditional methods like back-propagation neural network (BPNN) and feature extraction by correlation between training images of T-Dataset and BPNN.

Binary filtering and Circular Hough Transform (CHT) have been used for circular object detection. Firstly, they filtered the background and after that a gray scale filter is used to prepare the dataset for binary filter and circular Hough transform. It is successful in detecting an object but CHT may not exactly detect the circular object as sometimes it is connected with other object and give an inaccurate result [12]. Color offers potent object recognition data. Swain and Ballard's model is straightforward reconnaissance scheme is the

representation of matching images based on RGB histogram [13]. This color-based recognition method has been extended by Funt and Finlayson and to get extensive flexibility they introduce illumination by indexing on a light-invariant color range [14].

Alex, Ilya and Geoffrey used a very modern technique to train up data's in neural network [15]. Their method of deep convolutional neural network has the capability of doing significant computational power. Their analysis also gives us an insight that regardless of the complication of the dataset it can achieve good result using supervised learning. But subsequent reduce of one layer hinders the performance and accuracy too.

LeCun, Yann, Fu Jie Huang, and Leon Bottou.in their paper points out the lack of flexibility and resource minimization of template-based approaches. Their proposed model is more feature extractable and robust [16].

Chae *et al.*, in their paper "A Wearable sEMG Pattern-Recognition Integrated Interface Embedding Analog Pseudo-Wavelet Preprocessing" have presented a wearable wireless surface electromyogram (sEMG) integrated interface that utilizes a proposed analog pseudo-wavelet preprocessor (APWP) for signal acquisition and pattern recognition [17].

Zupan, in his paper of "Introduction to Artificial Neural Network (ANN) Methods: What they are and how to use them" has explained the selection procedure of training dataset. He has emphasized on this step as to be very important and suggested to divide the dataset into not two but three datasets. According to him, the first dataset should be for training, the second one should be the control set or fine-tuning set and lastly, the third one should be the test dataset. He also suggested that the training dataset should be smaller in size than the test dataset. He also mentioned that the true test set should contain completely 'non-committal' or unbiased set of data [18].

Zupan, when trying to explain artificial neural network simply, he compared it to a black box having multiple input and multiple output which processes large number of parallelly connected simple arithmetic units. ANN methods work best when they are dealing with non-linear dependence between the inputs and outputs.

Youtang and Jianming, in their paper of "Air Target Fuzzy Pattern Recognition Threat-Judgment Model" have tried to establish a threat judgement model that has high reliability in air defense systems in the naval warships. They have used fuzzy pattern recognition model to identify threats from air targets. They have classified the threat degrees considering target distance, type, speed, advent time, cross-point distance and flight altitude. They have theoretically measured the threats using the parameters. For example, distance threat membership function has the feature that the threat degree is inversely proportional to

the distance between targets and warships. Therefore, they described it as a descending ridge distribution [19].

This paper explains the model we came up with after experimenting a number of methods in terms of detecting drones using Neural Network tool. Our goal was to find a better and faster way to detect the object (drones) which leads to better performance and lesser error. Here, we used a hybrid feature extraction method using the SURF and GLCM features which is utilized for detecting a drone by Neural Network. SURF method is popularly used for its faster image matching property and GLCM extracts the best features of a set of images. Thus, combining the two methods we were able to form a dataset that gave us our desired result. The complete process of combining the methods have been explained in proposed model section.

Rest of the paper is organized as follows. Chapter II presents the proposed model, details result analysis is in Chapter III. Finally, conclude the paper in Chapter IV.

2- Proposed Model

The proposed model is a hybrid model where both SURF and GLCM methods have been used to extract features from the input and target datasets. Then the newly created dataset has been fed into SCG function of neural network to obtain the output set.

Four different versions have been developed in order to achieve a better result on the basis of cross entropy and percentage of error. They include applications of various methods for feature extraction such as, MSER, SURF, GLCM and SURF and GLCM combined. Different versions gave different results but the best model that was observed for both SURF and GLCM feature extraction algorithms. At first, images for training were pre-processed, then feature extraction algorithms, SURF and GLCM were applied to extract attributes of drones. This dataset was fed into the neural network for training and testing. From the output, it was possible to analyze the performance and error percentage of the model. The model is described step by step with a flow chart in Fig. 1.
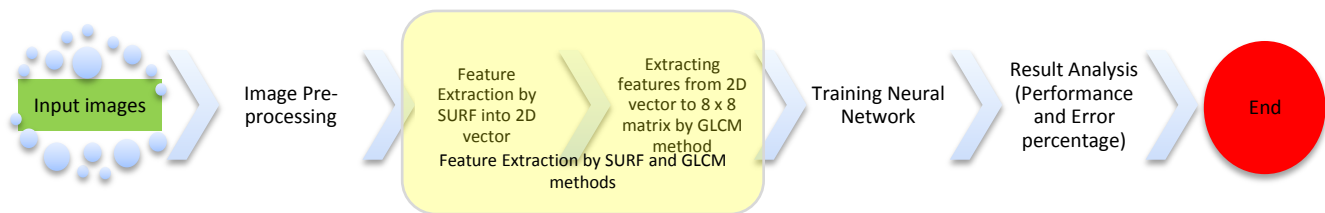

Fig. 1 Flow chart of the hybrid model of drone detection.

2-1- Image Pre-Processing

About 600 images of drones were collected and resized to 227×227 pixels. Among the total images, 75% were used for training and rest of the 25% for testing. All the images were converted into a uniform size. There are barely any scholarly articles to explain the reason for resizing the images to exactly same dimensions. Although, Nikhil *et. al.* mentioned that many Neural Network models expect or assume input images to be square shaped, therefore, images have to be reshaped or cropped [20]. The input images were of true color, having clear sky in the background so that no attributes of other components interrupt the feature extraction of the drones.
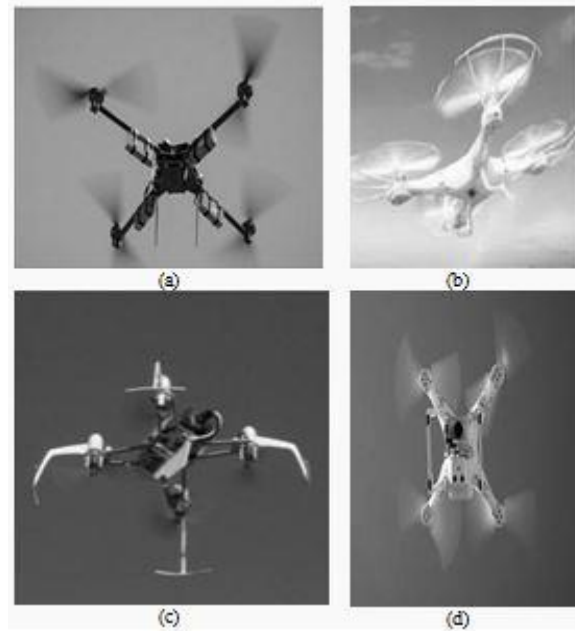

Fig. 2 (a), (b), (c) and (d) are the gray-scale images of drone #79, #98, #100 and #113, respectively.

Here in Fig. 2, the gray-scale or 2D images of few sample drones are shown. The test dataset was also of true color. At first, all the input (training) images have been taken randomly. Then each image was converted from true color or 3D to gray-scale or 2D image. 3D image is of RGB scale and consists of 3D numeric array and similarly, 2D image consists of 2D numeric array. The conversion removes the hue and saturation of the image keeping the luminance intact and returns a 2D array of double values. The method rgb2gray has been used for the conversion which is done by calculating a weighted sum of Red, Green and Blue. This follows an algorithm which is, 0.2989*R + 0.5870*G + 0.1140*B, where the values of R, G and B of a pixel are multiplied with their respective specific co-efficient and then summed together to provide a gray-scale pixel corresponding to that true color pixel [21]. Each value of the 2D array generated from this algorithm is in the range of 0 to 1, it can be positive or negative. The pixels with values greater than 0 are displayed as white and the pixels that are equal to 0 or less than 0 are displayed as black [22]. Similarly, the test image datasets are converted to gray-scale images one by one.

The 2D arrays were rotated to 90° angle for both datasets right after RGB to Gray-scale transformation. The rotation was required because in this model GLCM algorithm was applied as one of the feature extraction methods, where an offset had to be fixed that depends on the angle of rotation. This offset was later delivered to graycomatrix method [23].

Fig. 3 shows some of the images of drones which were used to train our network.

Fig. 3 Sample images of drones for the training dataset.

2-2- Feature Extraction by SURF and GLCM Methods

Since it is a hybrid model, the feature extraction took place in two steps. At first, attributes were extracted from the images using SURF algorithm. Then the 2D array of double values achieved by concatenating the features set of all the images was used to populate into GLCM method to obtain the best features, better performance and minimal error percentage. The steps are described elaborately below:

Speeded-up Robust Features (SURF) is a blob detection algorithm which means it detects the corners of the object and the locations where the reflection of light is higher (light speckles) [24]. This method is popularly used because of faster calculation of interest points due to use of integral images and it can detect locations best where there is illumination [25]. There are three main steps to this algorithm - interest point detection, local neighborhood description and matching. Firstly, the interest points are calculated using Hessian matrix. They can be found at different scales as the algorithm uses comparison images and the corresponding interest points can be found in different levels. To resolve this issue, Gaussian filter is used that smoothed the images repeatedly. Then they are subsampled to get the next level of the hierarchy of the pyramid (scale space) [26-29]. Since images of drones were taken from different scales, filters had to be used and the faster method to do that was provided by SURF algorithm. The levels are calculated by:

$$\sigma = \left(current\ filter\ size\ \times \frac{base\ filter\ scale}{base\ filter\ size} \right) \qquad (1)$$

If p(x,y) is point in an image and σ is the scale where the Hessian matrix is H(p, σ) and $L_{ij}(p,\sigma)$ is the convolution of the second order derivative of Gaussian, then-

$$H(p,\sigma) = \left(L_{xx}(p,\sigma)\ L_{xy}(p,\sigma)\ L_{yx}(p,\sigma)\ L_{yy}(p,\sigma) \right) \qquad (2)$$

Secondly, the local neighborhood descriptor is to be detected. Descriptors provide unique and robust features by describing the intensity or the orientation of the pixels. They are computed from the local neighborhoods of the interest points. To extract descriptors, a circular region around the point of interest of radius 6S is used, where S is the scale of the point. Then a square region is constructed around it aligned to the orientation to obtain scale invariance. Haar wavelet responses in horizontal and vertical directions are calculated within this squared region for each sample point [26], [27-30]. Finally, the descriptors are compared for matching among the images and the common points are taken as the matched attributes of the images [26, 28].
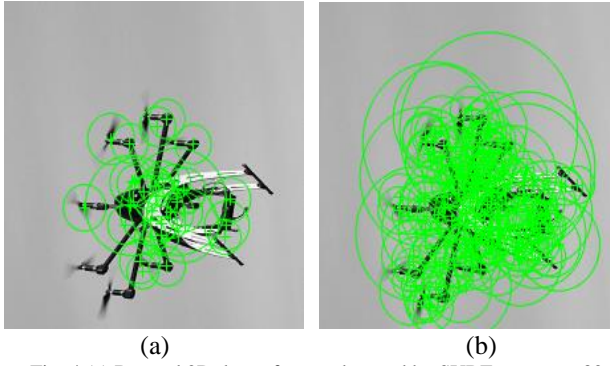
(a) (b)

Fig. 4 (a) Rotated 2D drone feature detected by SURF-strongest 30 points, (b) All possible points

Thus, the original image is fragmented into several squared regions for the use of integral image and then summed up for faster calculations [26-29]. To sum up the integral images, it uses equation (3), where I(i,j) is the interest point of the image I.

$$S(x.y) = \sum_{i=0}^{x} \sum_{j=0}^{y} I(i,j)$$

(3)

Fig. 4 represents the extracted features of a rotated drone by SURF method. In Fig. 4(a) only 30 strongest features are detected and in Fig 4(b) all the features are detected. In this study, all the features have been used. After incorporating the processed image datasets into SURF extraction method, it was possible to obtain the best features out of each image which is represented by a 2D array. The features set of each image have double type values with a size of $n \times m$.

Once the features are collected in a 2D array, it is fed into GLCM algorithm along with proper offset. Gray-Level Co-occurrence Matrix (GLCM) is generated by calculating the number of times a pixel with the gray-level intensity value at i occurs in a specific spatial relationship with the pixel j[m], where, Q(x,y) = i and Q(x+1, y+1) = j when diagonally right pixels are considered, and Q(x+1, y+1) = j when horizontal neighboring pixel is considered. Here, x and y are offsets [23, 31-33]. The equation is determined by using the dimension of the offset matrix. An 'Offset' is the distance between a pixel of interest and its neighbor. It is a p×2 matrix, where p is the number of pairs that pixels of interest make with their neighbors [23]. By default, it is [0 1]. The graycomatrix function takes two parameters where the image points and the offset are used. For image points the feature set obtained from SURF feature extraction method is taken and for the second parameter [2 0] offset is populated and this is the reason why the images are rotated to 90-degree angle in prior [23]. [2 0]-offset means that the sequence of pair of adjacent pixels which is to be considered (as feature), lies in 2 consecutive rows of

the same column. The size of GLCM matrix is determined by number of gray-level intensities which is by default 8. It usually returns an n × n matrix of extracted features. For this study, the function returned an 8 × 8 matrix that means the best 64 features were obtained for each image [31]. The equation (equation 4) for calculating GLCM features is given below [33-34]:

$$C_{\Delta x, \Delta y}(i,j) = \sum_{x=1}^{n} \sum_{y=1}^{m} \{1, if\ i\ and\ j\ true\ 0, if\ i\ and\ j\ false$$

(4)

Fig. 5 represents the extracted features graph of sample images of drones before pre-processing and their corresponding gplots of GLCM feature extraction method.
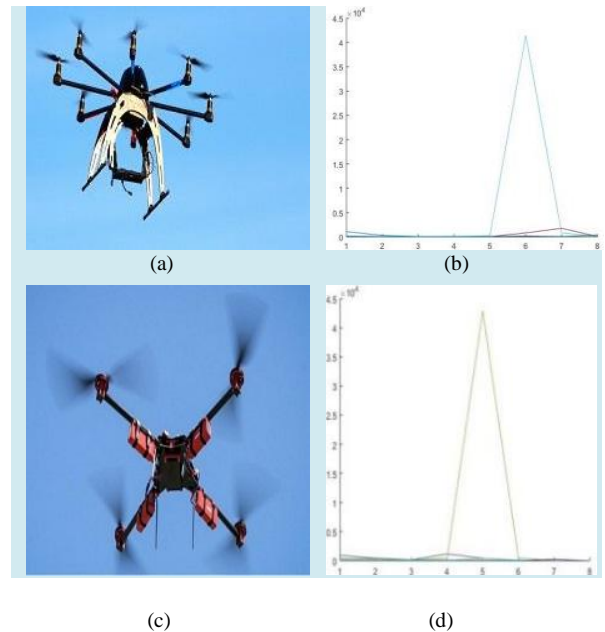

(a) (b)

(c) (d)

Fig. 5 (a) Drone #10 shown in RGB before pre-processing, (b) corresponding plot of features extracted by GLCM after pre-processing of (a), (c) drone #79 shown in RGB before pre-processing, (d) corresponding plot of features extracted by GLCM after pre-processing of (c).

The extracted feature set for each image is converted to 1D array. Then each of these arrays is arranged in another parent array which is the final dataset for training in the Neural Network. This is the newly created hybrid dataset that have to be populated in Scale Conjugate Gradient (SCG) function of the neural network. The separate datasets for training and testing are incorporated in the network, each having $n \times 64$ size 2D array where n is the number of images.

The dataset obtained have the final features of an image. It is arranged in a 2D array. This dataset is used for training the neural network.

2-3- Training Neural Network

A neural network is a collection of connected nodes called the artificial neurons loosely modeled like the neuron connections of the brain [35]. Like the biological neurons, the artificial neurons receive signal (input), combines it with their internal state (activation) and an optional threshold using an activation function and signal other neurons connected to it. The final output finishes the task, such as recognizing an object in the image. The important characteristic of the activation function is that it provides a smooth and differentiable transition as input value changes.

The network consists of connections, each connection provides an output of one neuron as an input to another. Each connection is assigned a weight that represents its relative importance [36].

Artificial neural network was chosen for the proposed model's dataset training. As a machine learning tool, neural network for pattern recognition algorithm has been used for this model.

It is a fully connected neural network which is open to various customization. It uses the basic equation of modelNN = learnNN(X, y) for training, and p = predictNN(X_valid, modelNN) for prediction. There is a chance for an arbitrary number of layers and different activation functions. We used an arbitrary number of layers and the activation function was set to default [42].

Pattern recognition is the algorithm which identifies or classifies object based on their key features [37]. For its fast and optimum classification method, it is not only used for object identification but also used in the fields like speech recognition, text classification, and radar processing. The classification by pattern recognition can be both supervised and unsupervised. Supervised classification is the one where classifiers are created from different object classes. On the other hand, unsupervised classification is the method where hidden structures or patterns are identified within the unlabeled data using segmentation and clustering techniques.

Since the aim of the study is to identify drones from unclassified images, we have used the unsupervised classification method of Pattern Recognition tool. The pattern of the images had to be trained to the system's network, so that on testing it could determine the drones with optimal performance and accuracy. Pattern recognition algorithm matches all the inputs' features with test images' features and try to calculate how much alike they are, considering their statistical variation [38]. And pattern recognition, when implemented with neural network, resolves complex recognition in real time. Real

time response is what we need in case of a drone is identified in the clear sky. Moreover, neural network is well known for its adaptive learning which other tools offer less. No wonder, the leading companies like DeepMind, Google AI, Facebook uses neural network as a machine learning tool. The datasets prepared in earlier step, are passed to the network which have 10 neurons or hidden layers and trained by 'trainscg' function (which uses SCG algorithm) suitable for low memory usage [39]. Scale Conjugate Gradient (SCG) Backpropagation function is an algorithm with superlinear convergence rate. It requires $O(n)$ space complexity, where n is the number of weights in the network, therefore, it is suitable for the system also to get a faster result [29]. SCG is evaluated considering 3 algorithms' performance as standard they are – Backpropagation algorithm (BP), the Conjugate Gradient Propagation (CGP), and the one-step Broyden-Fletcher-Goldfarb-Shanno memoryless quasi-Newton algorithm (BFGS). The speed-up of SCG depends on convergence criterion. If the demand for reduction in error is more, the speed-up will be boosted. SCG is user independent unlike CGP and BFGS, and the weight complexity also favors SCG in terms of showing long ravines in sharp curvature than BP where the ravines are short. Therefore, the overall performance of SCG is better than other training functions considering the low memory space and that is why it has been chosen as the training function of this network [40].

The network took training dataset and trained itself to recognize the pattern of the images of drones. Then by testing with the test dataset, it learned as well as gave output to the number of drones it could detect. The system, however, cannot determine the type of the drone but can identify drones and differentiate between other aerial objects – the output will show greater cross-entropy. The goal is to find the better method to extract features for training the system and it is possible to come up with a better algorithm.

This proposed model is using the best of two already very popular models. It was proven that extracting blob features into a 2D array was necessary. Hence the usage of SURF came into action, however, detecting a drone is a different matter altogether. Hence the GLCM method was thought of.

3- Result Analysis

In the results analysis, we have considered the performance and error percentage of the network. The performance is calculated by cross-entropy per epoch; the minimum is the cross-entropy, the better is the performance [41]. If the system takes all the properties into account then the performance will be 0. If it does not take any properties into account then the performance will be 100. Low performance mean the system works with the high number of properties when it runs the algorithm. Our

focus is to have this performance as low as we can, that means we wanted to take higher number of properties while detecting the drones. The percentage of error is calculated as,

$$percentError = \left(\sum (tind \cong yind) \right) \Big/ numel(tind)$$ (5)

Here tind is a 2D target vector indices and yind is 2D output vector indices. While running the algorithm our model leaves some portion of the dataset that is to say we cannot consider their properties; that portion is our percent error. We have to keep that low as much as we can. Now we give performance preference over percent error; because performance deals with all the properties and percent error deals with portion of the data set. If we do not take all the properties into account of a dataset, it does not matter how big our dataset is.

3-1- Comparative Analysis

Table 1 presents a comparative performance comparison of proposed hybrid models with other state-of-art models for our tested drone dataset.

Table 1 Comparative analysis of state-of-art models

No.	Version No.	Performance	Error percent
1.	MSER [v.3]	29.09	3.5
2.	GLCM Method [v.4]	1.88e-16	89
3.	SURF Feature [v.5]	30.5	2.34
4.	Proposed Model [v.6]	7.34e-17	33

With the trend analysis in Fig. 6, we can assume that SURF feature with GLCM is the better way to detect drones while it is in the air. This way, we can detect the drones with a minimum amount of time and less complexity; that too with accepted error percentage rate. Here, one question may arise how 33% error is better than having 2.34% error. The answer to the question is, it is not better. However, the performance is better when we use the proposed hybrid model. Moreover, 33% of error means the system leaves 33% of the input dataset while matching.

It is acceptable because it still is compared with 67% of the dataset where we know all of the pictures are of drones.
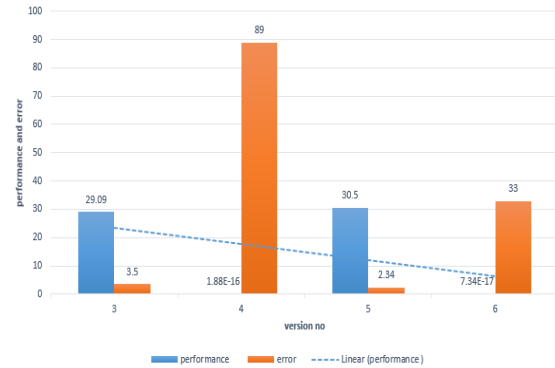

Fig. 6 Comparative analysis with state-of-art models by side with trend analysis.

Besides, we are doing it with all possible extracted features. We are giving priorities to the extracted features rather than how many of the dataset does it go through while comparing them and it does not leave too many of the dataset either in the proposed model. Therefore, the proposed hybrid model is better. Fig. 6 shows the trend analysis of performance and error percentage of all the four state-of-art models.

4- Conclusions

This paper presented a hybrid model to detect drones by extracting the attributes from the image dataset provided by SURF feature and populating the extracted information by GLCM algorithm. This dataset is fed into the network that uses scale conjugate gradient algorithm to recognize the pattern of the drones. The SCG function makes the system faster to detect the desired components. As a result of this model, the system is able to capture any kind of image of drones in the air and it can identify those with as much accuracy as possible and as fast as it can while it remains in the sky. Although, there are various modern neural networks like Alex-Net, ZF-Net, VGG Net, etc., we have chosen to provide a better way for drone detection using the traditional methods and tools for higher performance and lower percentage error. In addition, the proposed model exhibits better results than the state-of-art models.

References

[1] M. Hicks, "Criminal Intent: FBI Details How Drones are used in crime," Techradar-the source for tech buying advice, May 2018. [online]. https://www.techradar.com/news/criminal-intent-fbi-details-how-drones-are-being-used-for-crime.

[2] F. P. George, I. M. Shaikat, P. S. F. Hossain, M. Z. Parvez, and J. Uddin, "Recognition of emotional states using EEG signals based on time-frequency analysis and SVM classifier," International Journal of Electrical and Computer Engineering, 2019, vol. 9, no. 2, pp. 1012-1020

[3] R. Dong, H. Meng, Z. Long and H. Zhao, "Dimensionality reduction by soft-margin support vector machine," IEEE International Conference on Agents (ICA), Beijing, China, 2017, pp. 154-156.

[4] G. Yan, "Network Anomaly Traffic Detection Method Based on Support Vector Machine," 2016 International Conference on Smart City and Systems Engineering (ICSCSE), Zhangjiajie, Hunan, China, 2016, pp. 3-6.

[5] M. d. Barbosa, C. d. Barbosa, and A. F. Barbosa, "MuSSE: A Tool to Extract Meta-Data from Game Sprite Sheets Using Blob Detection Algorithm," 14th Brazilian Symposium on Computer Games and Digital Entertainment (SBGAMES), Piauí, Brazil, 2015, pp. 61-69.

[6] F. E. H Tay and L. Lao, "Application of support vector machines in financial time series forecasting," omega, vol. 29, no. 4, Aug. 2001, pp. 309-317.

[7] G. Kumar, P. K. Bhatia, "A detailed Review of Feature Extraction in Image Processing Systems," Fourth International Conference on Advanced Computing and Communication Technologies, IEEE Computer Society, Washington DC, USA, 2014, pp. 5-12.

[8] R. S. Choras, "Image Feature Extraction Techniques and Their Application for CBIR and Biometric Systems," International Journal of Biology and Biomedical Engineering, 2007, vol. 1, no. 1, pp. 6-16.

[9] Z. Yin *et al.,* "A Deep Normalization and Convolutional Neural Network for Image Smoke Detection," IEEE Access, vol. 6, 2018, pp. 4287-4296.

[10] J. Chen *et al.*, "Analysis of the recognition and localization techniques of power transmission lines components in aerial images acquired by drones," The Institute of Engineering and Technology Journals, IEEE Access, 2017, pp. 29-32.

[11] M. A. Abuzneid and A. Mahmood, "Enhanced Human Face Recognition Using LBPH Descriptor, Multi-KNN, and Back-Propagation Neural Network," IEEE Access, vol. 6, 2018, pp. 20641-2065.

[12] R. Hussin, M. R. Juhari, N. W. Kang, R. C. Ismail, A. Kamarudin, "Digital Image Processing Techniques for Object Detection from Complex Background Image," Procedia Engineering, 2012, pp. 340–344.

[13] M. J. Swain and D. H. Ballard, "Color indexing," International Journal of Computer Vision, vol. 7, no. 11, 1991, pp. 11-32.

[14] B. V. Funt and G. D. Finlayson, "Color constant color indexing," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 17, no. 5, May 1995, pp. 522-529.

[15] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Neural Information Processing Systems, 2012, pp. 1-9.

[16] M. Nakib, R. T. Khan, M. S. Hasan and J. Uddin, "Crime Scene Prediction by Detecting Threatening Objects Using Convolutional Neural Network," International Conference on Computer, Communication, Chemical, Material and Electronic Engineering, Bangladesh, 2018, pp. 1-4.

[17] H. Y. Chae, K. Lee, J. Jang, K. Park, and J. J. Kim, "A Wearable sEMG Pattern-Recognition Integrated Interface Embedding Analog Pseudo-Wavelet Preprocessing," IEEE Access, 2019, vol. 7, pp. 151320-151328.

[18] J. Zupan, "Introduction to Artificial Neural Network (ANN) Methods: What They Are and How to Use Them," Acta Chimica Slovenica, 1994, vol. 41, no. 3, pp. 327-352.

[19] T. Youtang and W. Jianrning, "Air target fuzzy pattern recognition Threat-judgment model," Journal of Systems Engineering and Electronics, 2003, vol. 14, no. 1, pp. 41-46.

[20] B. Nikhil, "Ïmage Data Pre-Processing for Neural Networks," Becoming Human: Artificial Intelligence Magazine, 2017.

[21] rgb2gray-Convert RGB image or colormap to grayscale, MathWorks, v: R2018a, 2018. [online]. https://www.mathworks.com/help/matlab/ref/rgb2gray

[22] mat2gray-Convert matrix to grayscale image, MathWorks, v: R2018a, 2018. [online]. https://www.mathworks.com/help/images/ref/mat2gray.

[23] R.M. Haralick and L.G. Shapiro, "Computer and Robot Vision," vol. 1, Addison-Wesley, 1992, pp. 1-459

[24] A. Xu and G. Namit, "SURF: Speeded-up Robust Features," 2008, Project Report: McGill University.

[25] T. Das, R. Hasan, M. R. Azam and J. Uddin, "A Robust Method for Detecting Copy-Move Image Forgery Using Stationary Wavelet Transform and Scale Invariant Feature Transform," International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2), Bangladesh, 2018, pp. 1-4.

[26] G. S. Rabbani, S. Sultana, M.N. Hasan, S. Q. Fahad, J. Uddin, "Person identification using SURF features of dental radiograph," 3rd International Conference on Cryptography, Security and Privacy, 2019, pp. 303

[27] Detect SURF Features-Detect SURF features and return SURF Points object MathWorks, v: R2018a, 2018. [online]. https://www.mathworks.com/help/vision/ref/detectsurffeatures.

[28] H. Bay, A. Ess, T. Tuytelaars, L. V. Gool, "SURF: Speeded Up Robust Features," Computer Vision and Image Understanding, 2008, vol. 110, no. 3, pp. 346–359.

[29] E. Oyallon and J. Rabin, "An Analysis of the SURF Method," 2015, Image Processing On Line (IPOL), pp.176-218.

[30] B. Fan, Z. Wang, F. Wu, "Local Image Descriptors: Modern Approaches," Springer, 2015, vol. 12, pp. 1-99.

[31] R. M. Haralick, K. Shanmugan, and I. Dinstein, "Textural Features for Image Classification," IEEE Transactions on Systems, Man, and Cybernetics, 1973, vol. SMC-3, pp. 610-621.

[32] A. Uppuluri, "GLCM texture features- Calculates texture features from the input GLCMs," version:1.2.0.0, MathWorks, v: R2018a, 2018. [online].

https://www.mathworks.com/matlabcentral/fileexchange/22 187-glcm-texture-features.

[33] P. Cosman, "Gray-Level Co-occurrence Matrices (GLCMs)," [online]. http://www.code.ucsd.edu/pcosman/glcm.pdf.

[34] J. Uddin, R. Islam, J. M. Kim, "Texture Feature Extraction Techniques for Fault Diagnosis of Induction Motors," Journal of Convergence, 2014, vol. 5, no. 2, pp. 15-20.

[35] Image Recognition- Recognition methods in image processing, MathWorks, v: R2018a, 2018. [online]. https://www.mathworks.com/discovery/pattern-recognition.

[36] "The Machine Learning Dictionary," available at: www.cse.unsw.edu.au. Retrieved at 4 November 2009.

[37] A. Zell, "chapter 5.2," Simulation neuronaler Netze [Simulation of Neural Networks] (in German) (1st ed.), Addison-Wesley, 2003.

[38] C. M. Bishop, "Pattern Recognition and Machine Learning," Springer, 2006, pp. 1-758

[39] Classify Patterns with a Shallow Neural Network, MathWorks, v: R2018a, 2018. [online]. https://www.mathworks.com/help/nnet/gs/classify-patterns-with-a-neural-network.

[40] M. F. Moller, "A Scale Conjugate Gradient Algorithm for Fast Supervised Learning," Neural Networks, 1993, vol. 6, no. 4, pp. 525-533.

[41] Crossentropy- Neural network performance, MathWorks, v: R2018a, 2018. [online]. https://www.mathworks.com/help/nnet/ref/crossentropy.

[42] V. Tshitoyan (2021). Simple Neural Network (https://github.com/vtshitoyan/simpleNN), GitHub. Retrieved January 20, 2021.

Tanzia Ahmed received the Bachelor of Science in Computer Science and Engineering degree from BRAC University, Dhaka, Bangladesh in 2019. She has experience on Software Analysis and Software Development for about two years. Currently, she is a master's candidate of Concordia University, Canada. She is enrolled in the program of Master of Applied Computer Science under Department of Computer Science and Software Engineering. Her research interests include Machine Learning, Image Processing, Computer Interface and Software Design and Analysis.

Tanvir Rahman received the B.Sc. degree in Computer Science and Engineering from BRAC University, Dhaka, Bangladesh in 2018 and M.S. degree in Computer Science and Engineering from BRAC University, Dhaka, Bangladesh in 2021. Currently He is a full-time lecturer in BRAC University, Dhaka, Bangladesh. His research interests include Information retrieval, Forecasting, Algorithms, Machine Learning and Data mining.

Bir Ballav Roy received the B.Sc. degree in Computer Science from BRAC University, Dhaka, Bangladesh in 2019. Currently he is working as a Software Engineer in a data centric software Firm. He is a data enthusiast and likes to get insights from data using machine learning models and Deep Learning and Statistical Models. His research interest includes Data Mining, Computer Vision, NLP, Artificial Intelligence and Machine Learning.

Jia Uddin is an Assistant Professor of Technology Studies Department in Endicott College, Woosong University, Daejeon, South Korea. He is an associate professor and undergraduate coordinator (On leave) in the department of Computer Science and Engineering, BRAC University, Dhaka, Bangladesh. He did Ph.D. in Computer Engineering from University of Ulsan, South Korea and M.Sc. in Electrical Engineering emphasis on Telecommunications from Blekinge Institute of Technology, Sweden. His research interest includes Multimedia Signal Processing, Fault Diagnosis, Bangla Language Processing, IoT based Intelligent System Design.