SDR-based Drone Detection using Machine Learning Algorithm

WMH Bandaranayake [1#], HHC Gunathilaka[1]

[1]*Centre for Defence Research and Development, Sri Lanka*
[#] <hiroshan.bnayake@gmail.com>

Abstract— *In this new era, misuse of drones and harmful acts that can be done using drones make it hard to detect and classify drones effectively due to the larger bandwidth and real-time processing. The purpose of this research is to find a better machine-learning algorithm to detect and classify the emitting signals from a drone or a remote controller. In the research multiple classification models were built and trained over the dataset obtained using Software Defined Radio (SDR) and drone remote controller. The performances of all these models were compared and their results were in terms of prediction accuracies. Based on the accuracy results, K-Nearest Neighbor classifier has given the highest accuracy among all other models.*

***Keywords* - RF signal classification, detection, software defined radio, machine learning model, neural network, K-nearest neighbor, feasibility**

I. INTRODUCTION

The drones are operated by radiofrequency (RF) and self-contained programmable control devices or operate completely or intermittently and autonomously by an onboard computer that utilizes Global Positioning System (GPS) signals. These days, drones are utilized in civilian applications as well as military context needs. Due to the use of drones, industries have been empowered by more matured technologies, leading to the gradual lowering of the threshold for the use of consumer and industrial-grade drones, making them widely used in civilian applications such as cinematography, agriculture, and mapping. Lack of industrial standards drones has become a high security risk for many public events, civilian infrastructure, and military facilities (Al-Emadi & Al-Senaid, 2020). These drones are used in many unlawful ways for surveillance, suicide attacks, and smuggling drugs. In order to countermeasure the Drones, Centre for Defence Research and Development, Sri Lanka has developed the "Drone Jammer I". But it did not have the capability to detect drones in real-time and the system needed continuous power. Due to this the civilian communication system could be interrupted. The purpose of this paper is to check the feasibility for developing a successful drone detection system that can support a Drone Jammer to detect and identify drones. By doing this, drones in unauthorized areas could be stopped with fewer civilian communication systems interrupted. It will detect the drone first and later identify it as a drone and activate the Drone Jammer which minimizes the effect of the other types of civilian communication disruptions (Huang & Qian, 2022), (Bello, 2019).

Researchers are faced with an intriguing challenge when trying to detect drone RF signals. Among the difficulties are:

1) Since drones can move and appear in all directions, detection and monitoring systems must be able to monitor numerous directions simultaneously.
2) When the drone is far from the detecting module, it is challenging to identify the appearance of the drone from other flying objects such as kites, and birds.
3) Due to their battery life and communication is limited and also drones frequently operate under lower altitudes which makes it harder to detect them.

II. DRONE DETECTION METHODS

Drone detection methods are becoming more significant due to security considerations as Drones are developed for both military and civilian applications. Also, drones can create a significant amount of privacy and ethical issues. Researchers and engineers have created many techniques to detect and identify drones in real time in order to face these problems (Bello, 2019). Following are four techniques which are widely used for drone detection.

A. Visual

Visual drone detection comprises detecting and tracking drones using cameras and computer vision algorithms. This approach detects drones based on visual signatures such as shape, size, and motion patterns or even IR based on heat signatures.

B. Radio Frequency (RF)

Communication signals sent between drones and remote controllers are used in RF-based drone identification systems. These technologies look for certain characteristics generated by drones or remote-control systems in the radio frequency band. By analyzing the frequency, modulation, and power- magnitude of the signal RF detection systems can detect and locate drones within their operational range.

C. Radar

Radar waves are often employed to detect flying objects, and the same methods can be used to detect drones. These devices operate by transmitting electromagnetic signals and then assessing echoes or reflections from flying objects. Radar systems can distinguish drones from other flying objects such as birds or aircraft by analyzing their size, distance, and speed.

D. *Acoustic*

Acoustic drone detection systems use sound waves to identify drones. These devices record the different noise signatures produced by drones using a range of precisely placed microphones. Acoustic detection systems can correctly detect the existence of drones in the vicinity by analyzing noise patterns, namely the frequency, power, and direction of the sound.

Some of these machine learning models might perform well in specific environmental situations, but others must have greater detecting power.

III. HOW DOES DRONE WORK

The drone system is made up of two primary components. There are the remote control and the Drone itself. Both use a radio frequency communication link to communicate with one another. The Figure 1 exhibits architecture and architectural design of the drone.

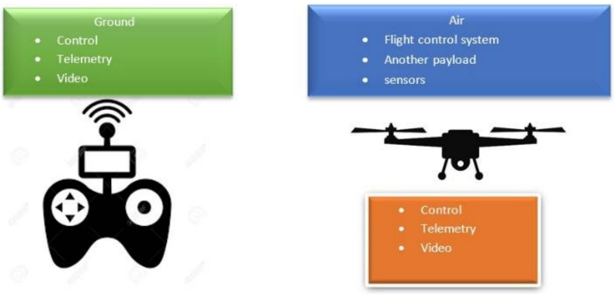

Figure 1: Components of the Drone System

Source: (Bello, 2019)

This paper focuses on detecting the RF classification on communication between the ground (remote controller) and the drone (Air). The remote controller may directly control the drone via control, data, and video transmission to communicate and send information. The majority of controllers operate on the 2.4 GHz frequency band, employing a proprietary frequency hopping spread spectrum (FHSS) modulation (Bello, 2019), (Al-Emadi & Al-Senaid, 2020). Due to that, Drone Jammer typically uses a combination of radio frequency circuits and antenna systems (Bello, 2019).

Table I: Frequency Bandwidth used by Drones and Data Transmission Type

Frequency Bandwidth	**Type of Data Transmission**
2.4 GHz	For communicate with their remote operators
1.2 GHz-1.5 GHz	GPS Navigation
5.8 GHz	video footage Transmission

Source: (Bello, 2019)

RF circuit is a special type of circuit in which Drone Jammers are responsible for generating and transmitting the jamming signal of analog circuits operating at very high frequencies suitable for wireless transmission.

IV. PROPOSED RF-BASED DETECTION METHOD

Drone detection system is based on the Radio Frequency. The system must consist of a device that is capable of capturing specific bandwidth limits from the frequency spectrum. In this experiment, Software Define Radio (SDR) was utilized for capturing and recording drone signals.

A. *BladeRF Micro A9 SDR*

In this experiment, BladeRF Micro A9 given in Figure 2 was utilized for capturing and detecting drone signals. The bladeRF Micro A9 is a software-defined radio (SDR) platform designed by the company Nuand, which specializes in SDR and its application development. The bladeRF Micro A9 consists of Cyclone V A9 FPGA, a 47 MHz to 6GHz radio frequency transceiver, and a high-speed USB 3.0 interface for data transmission to the PC or respective tool (BladeRF 2.0 micro, n.d). The bladeRF is a collection of a variety of open-source software tools and libraries, notably. GNU Radio, a well-known software development toolkit for SDR applications, supports the platform (Bello, 2019). For the BladeRF Micro A9, this enables the user to create software programs and procedures for signal processing.

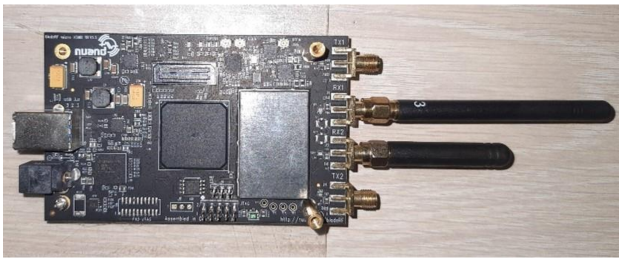

Figure 2: The BladeRF Micro A9 SDR
Source: Generated by the authors

Before analyzing the drone, signal testing was performed on a handheld device and the handheld device signal was captured and observed using BladeRF Micro A9 SDR and SDR console v3 and Nuand bladeRF CLI software.

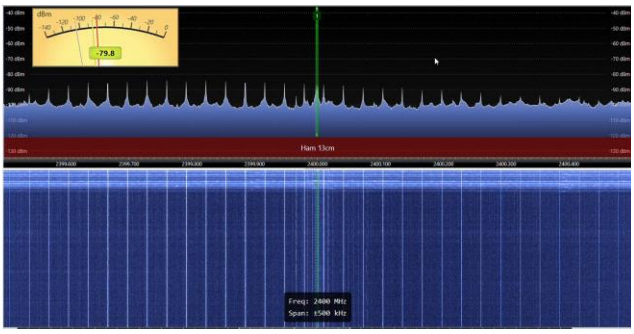

Figure 3: Captured Frequency Waterfall from the SDR consolev3
Source: Generated by the authors

With the help of the SDR Console v3 software program, users of Windows-based computers can manage and operate SDR receivers and transceivers. It offers a simple user interface for demodulating, decoding, and adjusting different radio signal types. The Nuand bladeRF CLI provides a software solution for the bladeRF SDR by the manufacturer. Figure 3 shows the sample of the obtained frequency spectrum using SDR console v3 software for a specific bandwidth.

The SDR console interface consists of a frequency vs. amplitude graph for received signals. The frequency waterfall of the SDR console identifies the intensity of each point on the graph and represents the amplitude of the signal at that particular frequency and time.

B. Captured and recording of the Frequency Bandwidth

Firstly, BladeRF SDR has been connected to the computer with BladeRF CLI software. The BladeRF CLI allows users to interact with the SDR and define and record the Rx channel input for a certain bandwidth in a certain period of time. After obtaining the data from the preferred format, it is utilized to observe and train the Machine learning model later.

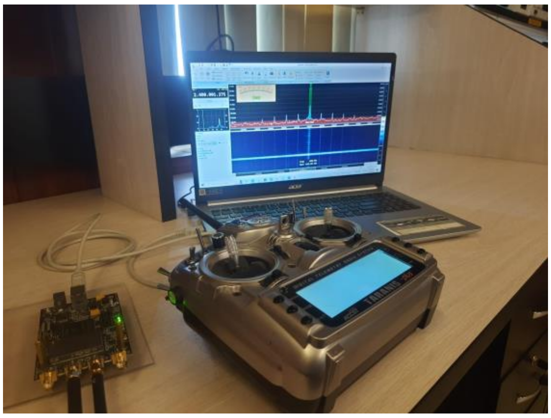

Figure 4: Collecting the data samples using BladeRF and drone remote controller.
Source: Generated by the authors

As shown in Figure 4, the receiving channel (Rx) of the BladeRF SDR can be configured to record specific frequency bandwidth. A way to record the received channel signals was found by using bladeRF CLI. In this approach, RX channel output was obtained for 2.4GHz using bladeRF CLI and bladeRF SDR. For testing, 2.4GHz frequency signals were recorded on the CSV file. The following steps were followed for that. First, connected the bladeRF to the computer and opened the bladeRF CLI program. After the bladeRF CLI program connects to the device, set the Rx channel frequencies to 2.4GHz and the sample rate to 2000 0/1Hz. The bandwidth of the Rx channels was 3 MHz. The automatic Gain Control (AGC) was turned off and set the gain to 60 dB. Afterward, entered the CSV file name and the number of samples wanted to capture. To start the recording rx start and for ending the recording rx wait command can be utilized. To check the parameters before the run print command can be used and it shows all values for parameters (Nuand, n.d.).

These Rx channels are configured to record the received signal in the CSV file. It will be very helpful to find a way to record the Rx channel output for future work in drone detection systems. The data collected in the CSV file can be utilized for training and testing the classification model about to be developed. Both Rx channel outputs have been recorded at the same time (BladeRF Micro A9 has two Rx channels and two Tx channels). The data is represented in two columns the first column represents the Rx1 channel and the second column represents the Rx2 column.

C. Machine Learning classifier

The target of this work is to identify the RF signals communication with drones and remote control. This required the machine learning classification model to separate the normal signal pattern from signal patterns from the drone signals. Using recorded signals obtained from BladeRF CLI, labeled the data prior to the built model. This will enable to building of a supervised learning model using that data (Yoon et al., 2016).

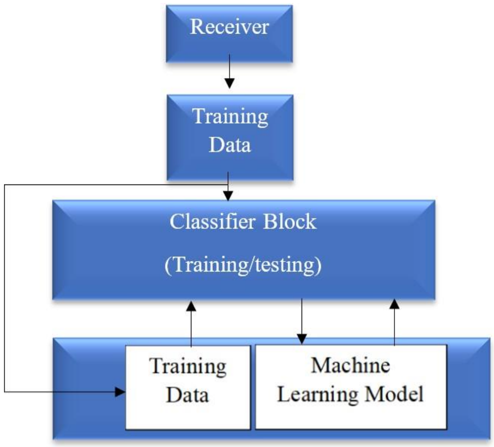

Figure 5: Block diagram for the Machine Learning Model
Source: Generated by the authors

The estimated parameters from the recorded CSV file are first used to generate a training data set before being saved in a data library for subsequent processing. In training mode, the classifier generates models based on the machine learning classification algorithm of choice. The new signal is then tested using the model chosen from the data library.

V. CLASSIFIERS

A. Logistic Regression Model

A logistic regression model is utilized to solve classification problems. It is problematic to identify whether the radio signal pattern contained with the drone is the emitted signal or not. Therefore this is one of the best candidates to solve the binary classification problem. The logistic function to produce a projected binary outcome is the output of the logistic regression which calculates the chance that the outcome variable will fall into a specific category (Yoon et al., 2016).

As shown in Figure 6, if we take an example with age (X) and job (Y), (If it is a job, value is 1 and of that is not a job value is 0), by using the label data and Logistic Regression model the new person can be predicted for whether it is their job or not. The data samples have been labeled for signals that contained the drone controlling signal as "Yes" and signals that did not contain the drone controlling signal as "No". using this label data set Logistic Regression model has been trained and tested (Fatmakursun, 2021).

B. Decision Tree classification

A decision tree classification algorithm is a concept that predicts a model in the shape of a tree. It is a further improvement of the classification and regression model. Decision Tree classification is a popular algorithm for both classification and regression tasks.

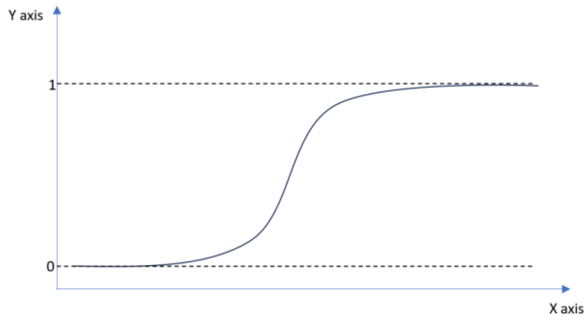

Figure 6: Example graph of a logistic regression curve fitted to data
Source: Generated by the authors

C. Random Forest Classification

Random Forest Classification uses a group of decision trees to create predictions. It is an extension of a decision tree algorithm that is frequently used for classification jobs. The use of randomization techniques and the combination of many decision trees lowers the danger of overfitting in comparison to a single decision tree. It can handle categorical and numerical features and can identify intricate patterns in the data.

D. XGBoosts Classifier

The XGBoost Classifier or Extreme Gradient Boosting Classifier is a machine learning technique from the gradient boosting algorithm category. It is well-known for its strong performance and has seen widespread use in data science com- petitions and real-world applications. The XGBoost Classifier has a number of benefits, such as excellent predicted accuracy, efficient processing, and the ability to handle large-scale data sets.

E.KNN

The k-nearest neighbors (KNN) are a prominent machine-learning technique used for classification and regression applications. It is a non-parametric technique that predicts similarity between instances in a labeled data collection. It is worth noting that the value chosen for 'k' is essential in the KNN algorithm. Over-fitting can occur when the algorithm has a low 'k' value. A big 'k' number can lead to under-fitting. The optimal value for 'k' is determined by experimentation or cross-validation and is based on the data set and situation. The KNN model used for this project has been evaluated for the five nearest neighbors (k=5) (Fatmakursun, 2021).

F. Nave-Bayes classifier

For classification tasks, the Naive Bayes classifier is a simple but powerful probabilistic machine learning algorithm. It is based on Bayes' theorem and makes the "naive" assumption of feature independence. Naive Bayes classifiers are famous for their simplicity, speed, and ability to handle high-dimensional data. They perform particularly well when the notion of feature independence is reasonably well supported.

G. ANN

An Artificial Neural Network (ANN) is modeled by observing the biological neural networks. ANN model processes data and adjusts based on input and output and its own changing the structure as a result. The artificial neural network (ANN) is a powerful tool for analyzing data patterns and modeling complex relationships. The model yields findings similar to logistic regression, a popular classification method.

The ANN model that has been used for this project consists of one input layer, two dropout layers, two hidden layers, and one output layer. The input layer is a fully Connected dense layer with neurons 128. The hidden layer number one is also a fully Connected dense layer with neurons 256. Hidden layer two consists of 128 neurons. Since the model is for binary classification Sigmoid activation has been used for the output layer.

H. CNN

CNN has demonstrated exceptional performance in a variety of computer vision tasks, including matrix or picture classification, object identification, and image segmentation. CNN is well-known for its capacity to learn hierarchical data representations automatically, capturing low-level properties in early layers and higher-level abstract information in deeper layers (Yoon et al., 2016).

The first Conv1D layer with 32 filters doesn't directly translate to the number of neurons as in fully connected layers. Filters are applied to the input data to detect features. In terms of hidden layers, there are 2 convolutional layers, there is one fully connected Layer, and There is one dropout layer. The second Conv1D layer with sixty-four filters also does not directly translate to the number of neurons. The dense layer with one-hundred and twenty-eight neurons. The dropout layer does not contribute to the number of neurons since it is a regularization technique. The neural network goes through the training data set 20 times (20 Epoch).

VI. EVALUATION

Firstly, a training sample set was prepared using signals recorded in the CSV file. The data samples were labeled and registered under drone signals which were not affected by drone transmission and receiving calls. Python-based library was used to split data sets into train sets and test sets. Afterwards, the Machine learning Models were trained. Using testing data, the accuracy of the model was observed. To analyze the model performance, confusion matrix and classification report were obtained for each model.

VII. RESULTS AND DISCUSSION

This section discusses the results of various machine learning classifier modeling. The accuracies obtained are shown in Table 2. The best performance in terms of accuracy was observed for the KNN model that uses five nearest neighbors to predict the music genre with a test accuracy of 60%.

However, the ANN model has 53% accuracy, and the CNN model has 54% accuracy. Other models such as the Logistic Regression Model, Decision Tree Classification, Random Forest Classification, XGBoots Classifier, and Nave- Bayes Classifier have less accuracy than 50%. The reason behind this low test accuracy rate could be the limited data set of 84 recorded signal samples. An increase in the data set might improve the accuracy of these models, especially on the CNN model. When the models are compared, it can be observed that the KNN model and the ANN model give a close test accuracy, even though it is low when we consider the CNN model's accuracy. Testing the systems with a bigger data set that might give better results.

Table 2. Accuracy of the Classifiers Models

Classifiers	**Accuracy**
Logistic Regression	45%
Decision Tree	45%
Random Forest	48%
XGBoosts	48%
KNN	60%
Nave-Bayes classifier	45%
ANN	53%
CNN	54%

Source: Generated by the authors

Compared with other researchers work, this result is adequate when considering the number of data sets that have been utilized. The CNN and ANN models had accuracy over 50% with just eighty-four of the training and data samples and only 33% has been utilized for testing.
Even though image image-based Drone detection system has much more potential to detect drones, the frequency-based Drone detection system has less disturbance from environmental factors, unlike image-based systems. This experiment gave valuable insight into the future development of a frequency-based drone detection system.

VIII. CONCLUSION

In this paper, Drone signal classification is studied using recorded signal samples by the BladeRF Micro A9 SDR and BladeRF CLI software. A simple comparisons were done between multiple other simple and complex, supervised learning models. All the models received an amplitude distribution of the respective bandwidth (2.4 GHz) with time and stored in a CSV for Logistic Regression and ANN model. CNN model gave the best accuracy while the KNN model and the ANN model give a close test accuracy, even though it is comparatively low when we consider the CNN model's accuracy. The KNN model evaluated for five nearest neighbors, and it gave the highest accuracy among all models. However, the CNN model has a promising future since the model was only trained with eighty-four data samples and 33% of samples were the testing set for the model. In future work, a neural network could be utilized to build a successful drone detection system based on radio frequency.

When considering the results shown in other related work (Huang & Qian, 2022), (Taha & Shoufan, 2019) the CNN model shows higher results than our models. The reason for that is the data set used to train our model only consists of amplitude vs. time distribution of the drone controlling signals. Other related work (Huang & Qian, 2022), (Taha & Shoufan, 2019) uses a classifier based on an image (spectrogram) instead of amplitude values of the signal for training and testing. CNN
models that are used in related works show higher accuracy like 99.8% [4]. But accuracy of our neural networks (CNN, ANN) lies in the 50% range. By improving the classifier technique, higher accuracy and precision can be obtained for machine learning models. Also, obtaining the data samples under different environmental factors such as higher, medium, and low RF disturbances will increase the precision of the classifier model. The model can be improved with much more capable feature extraction techniques which will increase the detection capability of the drone detection system.

REFERENCES

Al-Emadi, S. and Al-Senaid, F. (2020) 'Drone detection approach based on radio-frequency using convolutional neural network', 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT). doi:10.1109/iciot48696.2020.9089489.

Bello, A. (2019). Radio Frequency Toolbox for Drone Detection and Classification. doi:https://doi.org/10.25777/9gkm-jd54.

Yoon, J., Lim, H. and Kim, D.-W. (2016). Music Genre Classification Using Feature Subset Search. International Journal of Machine Learning and Computing, 6(2), pp.134–138. doi:https://doi.org/10.18178/ijmlc.2016.6.2.587.

Mo, Y., Huang, J., and Qian, G. (2022). Deep Learning Approach to UAV Detection and Classification by Using Compressively Sensed RF Signal. Sensors, 22(8), p.3072. doi:https://doi.org/10.3390/s22083072.

Taha, B. and Shoufan, A. (2019). Machine Learning- Based Drone Detection and Classification: State-of-the- Art in Research. IEEE Access, 7, pp.138669–138682. doi:https://doi.org/10.1109/access.2019.2942944.

Basak, S., Rajendran, S., Pollin, S. and Scheers, B. (2022). Combined RF-Based Drone Detection and Classification. IEEE Transactions on Cognitive Communications and Networking, 8(1), pp.111–120. doi:https://doi.org/10.1109/tccn.2021.3099114.

BladeRF 2.0 micro (no date) Nuand. Available at: https://www.nuand.com/bladerf-2-0-micro/.

Fatmakursun (2021) Titanic (classification models), Kaggle. Available at: https://www.kaggle.com/code/fatmakursun/titanic-classification-models.

Nuand (no date) Bladerf CLI Tips and tricks, GitHub. Available at: https://github.com/Nuand/bladeRF/wiki/bladeRF-CLI-Tips-and-Tricks.

ACKNOWLEDGMENT

We would like to give our heartfelt thanks to General Brigadier R Elvitigala USP, Director General of Centre for Defence Research and Development for helping us choose the domain of our project and being constant support from Captain (E) DSC Dissanayake psc, Chief Coordinator of Radio & Electronic Wing, and Lieutenant RAMH Ranasinghe SLEME for encouraging us throughout the journey of this project.