

Article

Hierarchical Classification Method for Radio Frequency Interference Recognition and Characterization in Satcom

Sahar Ujan *, Neda Navidi  and Rene Jr Landry

LASSENA Laboratory, Ecole de Technologie Superieure (ETS), 1100 Notre-Dame Street West, Montreal, QC H3C 1K3, Canada; neda.nav3@gmail.com (N.N.); renejr.landry@etsmtl.ca (R.J.L.)

* Correspondence: sahar.ujan@lassena.etsmtl.ca

Received: 21 May 2020; Accepted: 25 June 2020; Published: 3 July 2020



Abstract: The Quality of Service (QoS) and security of Satellite Communication (Satcom) are crucial as Satcom plays a significant role in a wide range of applications, such as direct broadcast satellite, earth observation, navigation, and government/military systems. Therefore, it is necessary to ensure that transmissions are incorruptible, particularly in the presence of challenges such as Radio Frequency Interference (RFI), which is of primary concern for the efficiency of communications. The security of a wireless communication system can be improved using a robust RFI detection method, which could, in turn, lead to an effective mitigation process. This paper presents a new method to recognize received signal characteristics using a hierarchical classification in a multi-layer perceptron (MLP) neural network. The considered characteristics are signal modulation and the type of RFI. In the experiments, a real-time video stream transmitted in the direct broadcast satellite is utilized with four modulation types, namely, QPSK, 8APSK, 16APSK, and 32APSK. Moreover, it is assumed that the communication signal can be combined with one of the three significant types of interference, namely, Continuous Wave Interference (CWI), Multiple CWI (MCWI), and Chirp Interference (CI). In addition, two robust feature selection techniques have been developed to select more informative features, which leads to improving the classification precision. Furthermore, the robustness of the trained techniques is assessed to predict unknown signals at different Signal to Noise Ratios (SNRs).

Keywords: supervised learning; time series classification; jamming detection; Automatic Modulation Classification; feature selection; Genetic Algorithm; Principal Component Analysis; QPSK modulation; APSK modulation

1. Introduction

In an optimal radio communication system, the design objective is to allow users to share a medium with minimum or no interference [1]. RFI is one of the most critical issues facing Satcom since it corrupts radio communication networks, disrupting the transmitter channel and Signal of Interest (SoI) reception capacity [2]. Nevertheless, a critical capability is concerning robust RFI recognition and characterization through effective real-time monitoring [3]. Since civilian non-intentional RFI could compromise industry revenues and military interference may cause mission failure and put lives in danger [4]. In this context, it is essential to deploy effective jamming detection methods in both civil and military applications. The jamming detection technologies need to be adapted to the kind of RFI such that their efficiencies are maximized [4–6]. Designing a highly accurate detection technique, which can deal with different jamming types under severe channel distortions, represents an interesting avenue for research.

Several spectrum sensing techniques, such as Cyclo-Stationary Feature Detection (CFD) [7], Energy Detection (ED) [8], and matched filtering-based detection [9], have been proposed for RFI monitoring in Cognitive Radios (CR) [10]. Among these techniques, CFD [7] can recognize the SoI from interference and noise, even in very low SNR contexts, but its implementation is computationally very complex [11]. On the other hand, ED is straightforward to deploy but is unreliable in cases of in-band interference and very noisy conditions [11]. As the matched filtering detector is highly dependent on the demodulation of a primary user signal, for its part, it requires a dedicated receiver design, which may not be possible in a practical CR terminal [12]. More significantly, the techniques above can only determine the presence and not the type of interference.

To address this, Artificial Intelligence (AI)-based methods have been employed and found as a superior to traditional approaches since classification techniques may not only determine the different kinds of RFI but also specify the modulation types efficiently. In most of the related studies, the authors focus on either determining the various types of RFI or the modulation types of SoI without considering the presence of the jamming signal.

In this work, we propose a novel approach based on hierarchical classification design to recognize not only the types of the received signal and Automatic Modulation Classification (AMC). To this end, we have considered modulation recognition under two severe conditions such as channel noise and the existence of various jamming signals. Moreover, we have generated a new real-time RFI dataset in which SoI is a video stream transmitted with four modulation types (QPSK, 8APSK, 16APSK, and 32APSK) in a real-time direct satellite-to-ground communication based on DVB-S2 standards. Furthermore, the studied SoI is combined with one of the three major jamming signal types, namely CWI, MCWI, and CI. Notably, we have studied the effect of these three jamming signals on RFI classification and AMC because they are very simple to be deployed by adversaries.

As feature selection is the previous step to classification in supervised learning, we have developed two feature selection approaches, namely the Genetic Algorithm (GA) and Principal Component Analysis (PCA), to select more informative features for an efficient classification process. GA is an evolutionary algorithm that improves the selection process over time. An advantage of GA over other techniques is that it allows the best solution to emerge from the best of prior solutions [13]. PCA, as a robust feature selection approach, is considered according to its low noise sensitivity, fewer memory requirements, and high efficiency in the training process [14]. This technique has been used in a wide range of computer science applications for feature extraction [15]. PCA projects a dataset from many correlated coordinates onto fewer uncorrelated coordinates called principal components, with information preserved as much as possible [15].

The rest of the chapter is organized as follows: Section 2 reviews the previous research in RFI classification and characterization; a methodology is proposed in Section 3; experimental results are provided in Section 4, and finally, the paper is concluded in Section 5.

2. Related Works

The main focus of this study is determining the type of received signal and its modulation in a Satcom scenario based on a supervised ML-based classification using MLP. ML methods have been used in different research studies for RFI classification and AMC regarding their flexibility in data processing and accuracy in the designed models. As in [12] an Artificial Neural Network (ANN) is proposed for a jammer detection in wideband radios. To this end, spectral correlation is used as the feature extraction technique. According to the obtained results, the proposed technique performed efficiently on SNR values down to -3 dB.

In [16], three robust ML-based classifiers such as MLP, Support Vector Machine (SVM) and Random Forest (RF) have been deployed to detect jamming attacks in the 5G wireless network. According to the results RF has slightly more precise classification performance. Reference [17] presents an efficient MLP-based approach to recognize one of the jamming attacks known as Denial of Service (Dos). Moreover, an unsupervised correlation-based feature selection method using the Pearson

Correlation Coefficient (PCC) is used to select more informative features. The authors in [18], present a robust MLP technique combined with a GA-based feature selection to detect various intrusions such as Remote to Local (R2L), User to Root (U2R) and DoS attacks. In [19], an efficient MLP design has been developed to recognize different intrusions. In [20], an efficient MLP is proposed to determine either a DoS attack exists in the wireless network or not.

In our previous work [21], we developed an MLP in different learning modes, namely Stochastic Gradient Descent (SGD), full-batch, and mini-batch. As the results show, the proposed technique can precisely classify four received signals, including SoI and a combination of SoI with three other jamming signals. Moreover, we deployed PCA to select more appropriate features to optimize the classification process.

In this work, we intend to propose a solution to determine not only the type of received signal but also Automatic modulation classification (AMC). AMC is a significant procedure for present and next-generation communication networks and facilitates the demodulation process at the receiver side [22]. Therefore, modulation recognition is an intermediate step between signal detection and demodulation. AMC is a system that automatically identifies the modulation type of the received signal [23], and the technology is widely used in various applications, such as dynamic spectrum management [24] and interference identification in CR [23].

In [25], a robust hierarchical classification based on MLP has been presented to recognize the modulation types of communication signals. As was thoroughly discussed in [14], AMC can be implemented regarding either a likelihood-based or a feature-based scenario. The disadvantages of a likelihood-based approach may include high computational complexity and sensitivity to impairments, such as phase and frequency offsets. The feature-based AMC, maybe comparatively more efficient [23] as it leverages robust extracted features such as an instantaneous amplitude, phase, and frequency [26], cyclostationary features [27], higher-order cumulants [28], and spectral correlation features [22]. In [23,29], the authors investigated different modulation types of recognition in a digital video broadcasting scenario based on higher-order cumulants and MLP.

AMC-based higher-order cumulant features are the primary concern of this study, inspired by [23] with proposing an approach based on higher-order cumulants of wavelet coefficients and MLP. Moreover, this study also applies PCA to select features that are more informative to accelerate the classification process. The results indicate that this method precisely classifies various modulation types with an accuracy above 99% at different SNRs ranges ($-4 \sim 4$) dB [23]. In [30], an AMC-based DT approach and higher-order cumulants are proposed. Additionally, different designs of DT, such as Fine Tree (FT), Medium Tree (MT), and Coarse Tree (CT), are also thoroughly analyzed. In [31], the author has applied SVM as the classifier and 4th-order cyclic cumulant feature to classify three modulation types such as ASK, BPSK, and QPSK by assuming that the channel noise is in the range of $-10 \sim 10$ dB. The experiment result shows that the 4th-order cyclic cumulant feature has an efficient discrimination capability in both non-noisy and noisy channels [31].

3. Proposed Methodology

The focus of this research study is on recognizing and characterizing the received radio frequency signals. To this end, a hierarchical classification design is proposed in which the first level deals with signal type recognition, and the second one automatically classifies the modulation type. Figure 1 presents the overview of the set-up, configuration, and framework of this study. It includes four main steps, namely data acquisition, feature extraction, feature selection, and classification. Each step is presented in detail in the rest of this section.

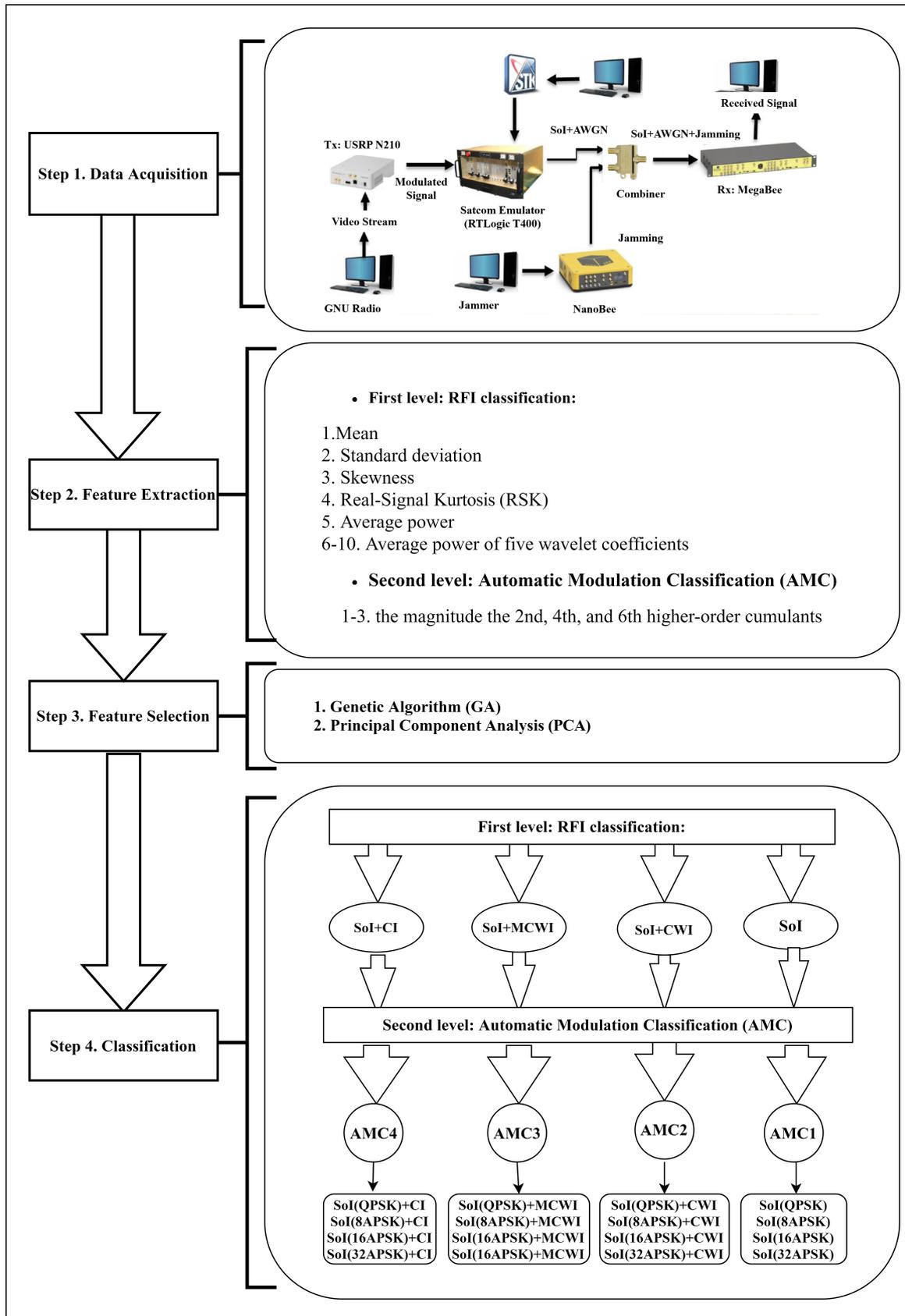


Figure 1. Proposed RFI recognition and characterization methodology.

3.1. Data Acquisition

As thoroughly has been explained in [21], the dataset created and used in this study is extracted from a real-time video stream, which is modulated and processed by GNU radio. GNU is a free and open-source software development toolkit that provides signal processing blocks to implement software radios. The SoI is transmitted using a Universal Software Radio Peripheral (USRP-N210) [32], which is an enhanced version of the USRP that includes a larger FPGA. In GNU radio, the modulation type and amplitude of the transmitted signal can be easily adjusted [33].

A Satcom emulator [34] (RTLogic) is used for modeling a real-time communication channel. The programmatic control of the channel simulator is facilitated over an Ethernet connection using a control protocol or optional plugin to System Tool-Kit (STK) software. The channel simulator produces IF/RF signals with extracting signal behavior for any scenario, such as various satellite's orbit around Earth [35]. The Kratos STK plugin provides real-time, phase-continuous control of the channel simulator when playing STK scenarios [34]. Furthermore, the generated jamming signals are transmitted using a NanoBee modem and are combined with SoI by an additive combiner [21]. Finally, the combined signal is received by a MegaBee modem [21].

A summary of the dataset characteristics is presented in Table 1. The generated dataset consists of 300 samples for each class of modulation type, namely Quadrature phase-shift keying (QPSK) and 8/16/32 asymmetric phase-shift keying (APSK).

Table 1. Real-time dataset characterization.

Characteristic	Value
Total number of samples	4800
Length of generated signals	32,448 (8 ms)
Sampling frequency	40×10^6 Hz
Modulation types	QPSK, 8 APSK, 16 APSK, and 32 APSK
AWGN power	140 dBm ($SNR \cong 9$ dB)
No. of each signal type per modulation type	300

3.2. Feature Extraction

According to distinguishable characteristics of RF signal in each classification phase (RFI classification and characterization), this study considers two feature extraction methods to use in “RFI classification” and “modulation type recognition”, as follows:

3.2.1. Features for RFI Classification:

To address RFI classification phase, by inspiring from [21] the six extracted features from each received signals are mean, standard deviation, skewness, Real-Signal Kurtosis (RSK), average power, and average power of the wavelet coefficients (4th approximation and details from 1st to 4th Order) as follows by assuming each received signal x with size 1 by n :

- Mean:

$$\mu(x) = \frac{1}{n} \sum_{i=0}^{n-1} x(i) \quad (1)$$

- Standard deviation (σ):

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n |x(i) - \mu(x(i))|} \quad (2)$$

- Skewness (I/Q): As in [36], the skewness of a signal is computed as:

$$Skewness = \frac{E(x - \mu(x))^3}{\sigma^3} \tag{3}$$

- The Real-Signal Kurtosis (RSK): This feature has been proposed by the author of [37] in which the kurtosis of In phase and Quadratic (I/Q) components of the signal is computed, and finally the RSK is obtained by averaging the computed kurtosis (I/Q).

$$kurtosis = \frac{E((x - \mu(x))^4)}{\sigma^4} \tag{4}$$

$$RSK = \frac{|Kurtosis_I| + |Kurtosis_Q|}{2} \tag{5}$$

- The average power: according to [38], the average power of each signal over its length is calculated as:

$$AveragePower = \frac{1}{n} \sum_{i=1}^n |x(i)|^2 \tag{6}$$

- Average power of the wavelet coefficients:

First, each observation (x) is decomposed up to 4 levels (in this case, the higher levels than 4 are less informative) using 10th Daubechies wavelet (db 10). Notably, different types of the wavelet have been tested and the best classification result obtained using db 10. Secondly the average power of the 4th approximation and details of 1st to 4th level is computed by applying Equation (6).

Therefore, the size of each feature vector for the RFI classification phase is 1 by 10.

3.2.2. Features for Modulation Type Recognition:

To present the modulation type recognition, the higher-order cumulant-based features have been computed for the AMC process. In mathematics, moments are employed to describe the probability distribution of a function. The p th order and q th conjugation moment for a received signal $x[n]$ in time-step n are defined as [39]:

$$M_{pq} = E[x[n]^{p-q}(x[n]^*)^q] \tag{7}$$

where $E(\cdot)$ and $[\cdot]^*$ are the expected operator and complex conjugate, respectively. Equation (7) can be approximated as [29]:

$$M_{pq} = \frac{1}{n} \sum_{n=0}^{N-1} x[n]^{p-q}(x^*[n])^q \tag{8}$$

In this work, the magnitude of the 2nd, 4th and 6th [21] cumulants are computed to distinguish between QPSK, 8APSK, 16APSK, and 32APSK:

$$C_{21} = M_{21} = E[|x(n)|^2] \tag{9}$$

where C and M , and index, respectively present Cumulant, Moment and their order.

$$C_{42}(x) = M_{42} - 2M_{21} - |M_{21}|^2 \tag{10}$$

$$C_{63} = M_{63} + 18M_{20}M_{21}M_{22} + 12(M_{21})^3 - 9M_{21}M_{42} - 3M_{20}M_{43} - 3M_{22}M_{41} \tag{11}$$

Notably, different orders of cumulant-based features were tested, and the best results were achieved using the above orders ($|C_{21}|$, $|C_{42}|$ and $|C_{63}|$).

3.3. Feature Selection Techniques

In supervised classification, it is required to minimize the number of features to speed up the training and classification processes, [40]. However, fewer characteristic features may come at the expense of classification accuracy degradation [40]. However, the trick in dimensionality reduction is to trade some precision for simplicity [15]. This study considers a comparative study of two feature selection techniques, namely PCA and GA. The main reason for using these techniques in this study is due to its low noise sensitivity, less memory requirements, and high efficiency in the training process.

3.3.1. Principal Component Analysis (PCA)

In [41], PCA has been presented as an efficient feature selection method with lower computational complexity in comparison to other approaches, like colony optimization [41]. PCA as a feature selection approach is implemented as follows [21].

First, the mean of each feature set is calculated then each value is subtracted from the mean. It is a crucial step to ensure that the first principal component describes the direction of the maximum variance [15]. A further step is calculating the covariance matrix, which is a representation of the linear dependency between two values. The third step is the calculation of eigenvectors and eigenvalues of the covariance matrix to determine the principal components of the data. Therefore, the highest eigenvalues are related to the most uncorrelated eigenvectors, which are considered to be the principal components. The main focus of PCA is to put maximum possible information in the direction of the first component, then the maximum remaining data in the second one, and so on. Since a feature component is less significant for feature extraction, it can also be interpreted that this feature is less informative in the original space [15].

3.3.2. Genetic Algorithm (GA)

GA is known as a search-based optimization technique inspired by Genetics and Natural Selection [42]. GA has been widely used not only for feature selection but also for optimizing the hyper-parameters of an Artificial Neural Network (ANN), such as weight [40,43–51].

As shown in Figure 2, a GA algorithm starts with an initial population, which is a subset of all the possible solutions (Also known as individuals) to the given problem [42]. Each individual has a set of genes, represented by a string of zeros and ones [42]. To evaluate the quality of a solution, a fitness value is assigned to each individual. Further various strategies can be applied to select the best individuals known as parents [42]. After the parent selection step, the variation operations such as mutation and crossover [42] are applied to generate new off-springs [42]. Finally, these off-springs replace the existing individuals in the population, and the process is repeated until reaching the stopping criterion.

As shown in Figure 3, to adapt GA technique with this study, the steps below are proposed:

- Binary representation: In the feature selection step, each of the elements in the feature vector may or may not be selected. Therefore, a string is used to represent the selected and discarded features. In the given string, each 1 indicates that the corresponding feature to the index is chosen, while each 0 determines the discarded features [42].
- Population model: In a steady state of GA technique, one or more off-spring is generated in each iteration, and they replace one or more individuals from the population [42].
- Population initialization: This study considers a random initialization, in which k bits of the string is randomly set to 1. In other words, k features are chosen out of the n -dimensional space of features [42].
- Parent selection: The “Rank Selection” approach is used for selecting parents since the individuals in the population have very close fitness values. In this work, the fitness value is the classification

accuracy which is computed by specifying the number of True Detection (N_{TD}) and the number of False Detection (N_{FD}) [21]:

$$Accuracy = \frac{N_{TD}}{N_{TD} + N_{FD}} * 100 \quad (12)$$

Therefore, each individual in the population is ranked according to their fitness, and the parent selection depends on the rank of each individual, not on the fitness value [42].

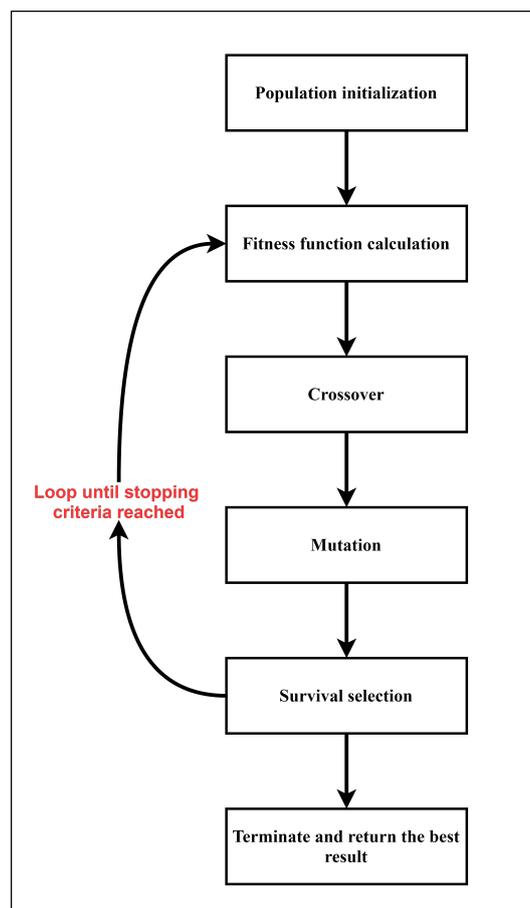


Figure 2. General framework of GA algorithm.

- **Mutation:** The swap mutation is deployed, in which two random positions on the chromosome are selected, and their values interchanged. It should be noted that in this application, the crossover is the same as the mutation. In other words, if two parents are selected, to exchange the elements in the selected indices (single point or other crossover methods), while preserving the number of 1s in the solution, so it accurately mimics the mutation process [49].
- **Survival selection:** The survival policy determines which individuals should be maintained in the next generation. However, this is a crucial step as it must ensure that the fitter individuals are not kicked out of the population while maintaining diversity in the same population. In this work, a fitness-based selection is used, in which the children tend to replace the least fit individuals in the population [49].
- **Termination criterion:** The algorithm is terminated when a set number of generations is reached. Figure 4 shows the variation in the best fitness values during 100 generations for RFI classification using five features. The result presents the improvement in the classification accuracy of over 100 number of generations up to 99.81%. Moreover, the best fitness value corresponds to the chromosome [0010001111], which indicates that the more informative features are the standard

deviation of the received signal and average power of the 4th wavelet coefficients (D_1 to D_4) using Daubechies 10 (dB10) wavelet.

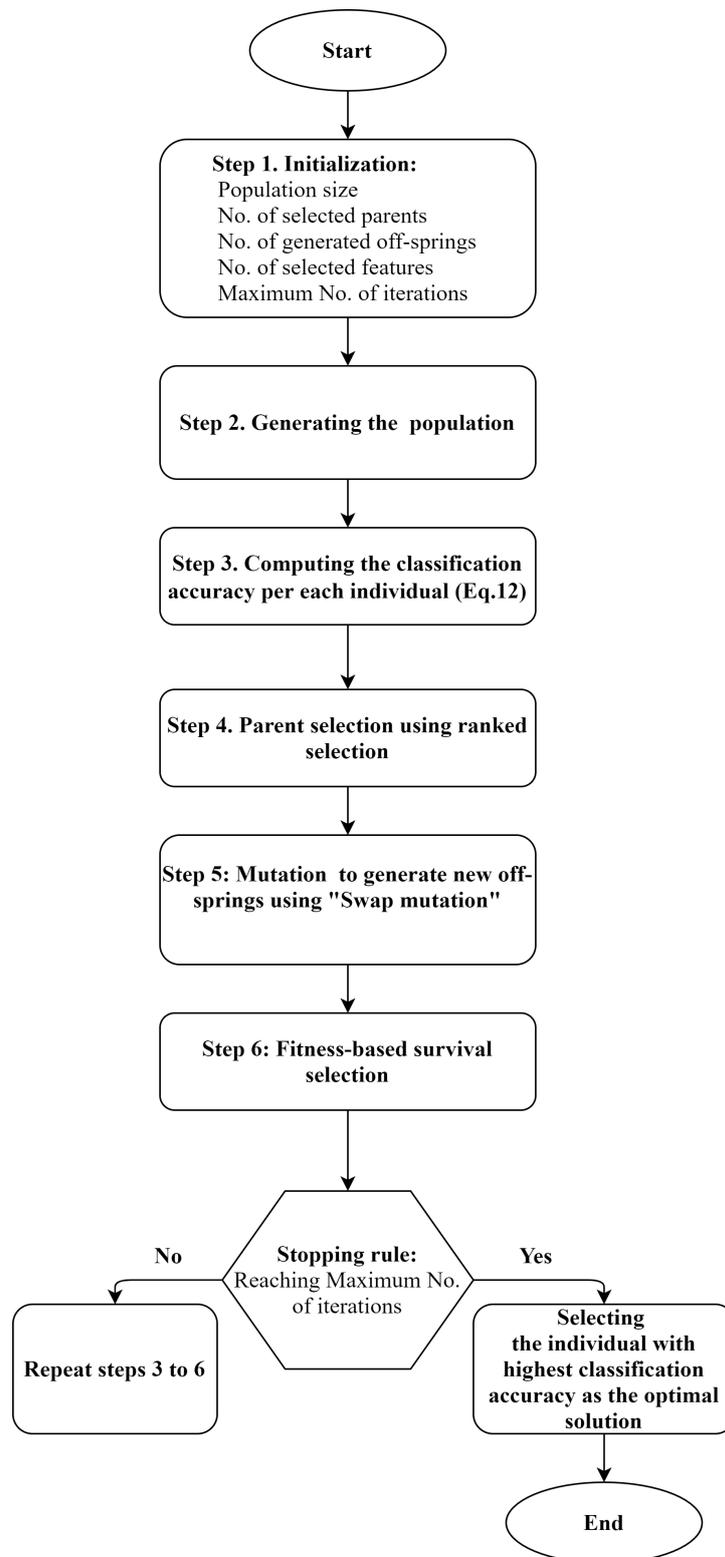


Figure 3. The proposed GA algorithm for feature selection.

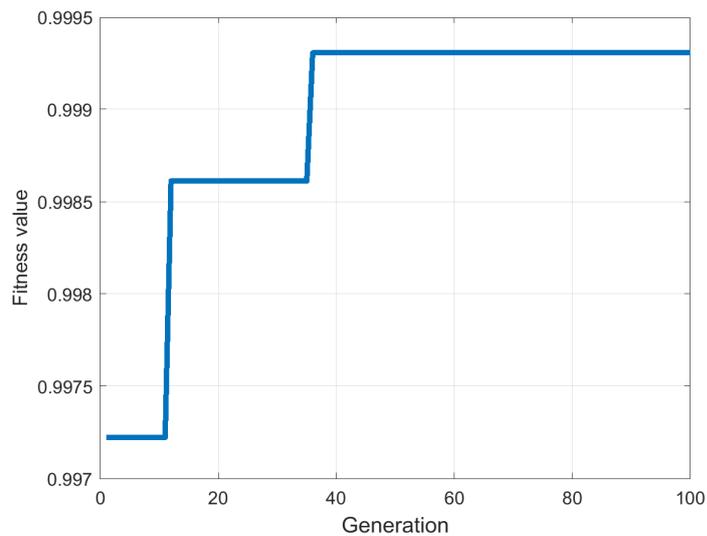


Figure 4. Variation of the best fitness values per generations number for RFI classification application using two features.

3.4. Classification

As shown in Figure 1, the classification phase contains a hierarchical design, including two levels. The first level is a global search using one classifier to detect the type of received signals (SoI, SoI + CWI, SoI + MCWI, and SoI + CI). Furthermore, a localized search is deployed for each type of signal to recognize the modulation type. To this end, we used the same classifier design proposed in our previous work based on MLP [21] trained using 10-fold cross-validation technique.

Cross-validation is a statistical approach used to evaluate the ability of ML-based models on unseen data. Generally, it is deployed in applied ML to compare and select a model for a given predictive modeling problem [52]. As the benefits of this generalization technique can refer to easy to understand, simple implementation, and results in skill estimates that generally have a lower bias than other methods such as a simple train/test split [52]. The procedure has a key parameter called k that refers to the number of groups that a given data sample is to be split into such that the procedure is often called k -fold cross-validation [52]. The implementation steps of k -fold cross-validation are as follows:

Step 1: Shuffle the dataset randomly and split it into k groups.

For each unique group:

Step 2: Take the group as a holdout or test dataset.

Step 3: Take the remaining groups as a training dataset.

Step 4: Fit a model on the training set and evaluate it on the test set.

Step 5: Retain the classification accuracy and discard the model.

Finally, the model's performance is evaluated by the average of all the obtained classification accuracy values.

Therefore, each observation in the dataset is assigned to an individual group and stays in that group during the training procedure. Therefore, each sample has the opportunity to be used in the holdout set once and used to train the model $k - 1$ times [52]. Choosing an optimal value for k is very crucial since a poorly chosen value for k may result in a misrepresentative of the model's skill, such as a score with a high variance or bias [52]. This study considers $k = 10$ due to achieved experiments results with low bias in estimated model skill. The value of $k = 10$ is widespread in the field of applied ML tasks.

4. Results and Discussion

This section presents and analyses the performance of the proposed algorithms for both RFI and modulation classification. All the simulations have been performed with MATLAB (Version R2019b), in a Core i5-5257U CPU computer system, operating at 2.70 GHz with RAM = 8 GB. Moreover, the dataset used for the classification has been generated at AWGN power -140 dBm, and approximately SNR value is 9 dB. As shown in Table 2, the power of the received signals (dBm/Hz) is measured using a signal analyzer in a 3.84 MHz Bandwidth (BW).

Table 2. Measured power of received signals.

SoI Power	CWI Power	MCWI Power	CI Power
-40 dBm	-33 dBm	-35 dBm	-32 dBm

The Jamming to Signal Ratio (JSR) is computed as

$$JSR = Power_{SoI} - Power_{Jamming} \quad (13)$$

Therefore, the measured JSR for CWI, MCWI, and CI is 7 dB, 5 dB, and 8 dB, respectively.

4.1. RFI Classification Results

The performance of the developed MLP approach in [21] is considered for RFI classification. Therefore, a two-layer MLP is used with 10 and 4 neurons in the input and output layers, respectively. Moreover, the logarithmic sigmoid and linear functions are used as the hidden and output layer activation functions [21]. Table 3 presents the details of the designed MLP using different Hidden Layer Neurons (HLN) and Batch Sizes (BS). The training process of this study is performed by the MLP using a 10-fold cross-validation.

Table 3. RFI Classification results based on MLP at different HLN and BS.

HLN	BS = 1	BS = 100	BS = 500	BS = 1500	Full-Batch
10	98.58%	91.56%	56.88%	45.18%	39.29%
20	98.50%	92.75%	61.46%	49.16%	37.83%
30	99.58%	92.88%	59.86%	49.16%	40.62%

As the results show, by increasing the HLN the accuracy can slightly improve. Moreover, due to high variety of classes in the dataset the online learning mode performs better in which the network's key parameters are updated based on each sample one by one. Therefore, the most accurate result has been achieved using $HLN = 30$ and online learning mode ($BS = 1$), with a precision of 99.58%. Table 4 illustrates the results of applying the proposed MLP + GA approach for RFI classification, using the different Numbers of Features (NoF) with $HLN = 30$. According to the results, the highest accuracy, 99.97%, is obtained with $BS = 1$ and $NoF = 8$. Also, the classification precision is increased by dealing with more features. Table 5 demonstrates the result of deploying PCA + MLP for the first classification phase using a various number of features and learning modes such as online learning ($BS = 1$), mini-batch, and Full-Batch. As can be seen, the classification accuracy reached to 97.05% with only five features. Table 5 demonstrates the result of deploying MLP + PCA for the first classification phase. As can be seen, a precise classification was achieved with only five features.

Table 4. Effect of GA-based feature selection on the designed MLP’s performance (*HLN* = 30).

NoF	BS = 1	BS = 100	BS = 500	BS = 1500	Full-Batch
2	94.41%	85.46%	80.04%	52.77%	48.46%
5	99.81%	98.39%	87.84%	72.63%	49.65%
8	99.97%	99.76%	96.45%	81.53%	50.27%

Table 5. Effect of PCA-based feature selection on the designed MLP’s performance (*HLN* = 30).

NoF	BS = 1	BS = 100	BS = 500	BS = 1500	Full-Batch
2	48.95%	44.74%	22.39%	18.21%	20.52%
5	97.05%	96.03%	94.79%	73.26%	54.40%
8	90.75%	96.53%	92.36%	64.05%	50.34%

4.2. AMC Phase Results

Table 6 illustrates the MLP-based classification results for modulation type recognition. The number of neurons for each layer is 3, 30, and 4, respectively. The results show that in case of no jamming, the average AMC accuracy is 87%. In the presence of jamming signals, the classification performance is degraded locally, about 3%, 16%, and 36% for CWI, CI, and MCWI, respectively. Each AMC classifier refers to a specific class of the received signal. AMC1, AMC2, AMC3, and AMC4, are respectively referred to SoI, SoI + CWI, SoI + MCWI, and SoI + CI.

Table 6. AMC results based on MLP at different *HLN* and BS.

Classifier	BS = 1	BS = 100	Full-Batch
AMC1	87%	57%	37.25%
AMC2	84.25%	43%	27%
AMC3	48%	47.50%	49%
AMC4	71%	65%	20%

Table 7 demonstrates the effect of GA-based feature selection on the second phase of classification to recognize the modulation types. As is shown, MLP + GA can precisely classify the four modulation types per each received signals using only one feature with a precision of 90.83%, 86.66%, 61.66% and 76.67% for AMC1 to AMC4 respectively.

Table 7. Effect of GA-based feature selection on the proposed MLP using *HLN* = 30 for AMC phase.

NoF	AMC1	AMC2	AMC3	AMC4
2	88.58%	81.67%	43.33%	70%
1	90.83%	86.66%	61.66%	76.67%

Table 8 illustrates the effect of PCA-based feature selection on classifying the modulation types. As is shown, MLP + PCA performs efficiently using only one feature for AMC1 to 4 with an accuracy of 86%, 89.77%, 51.11%, and 67.66%, respectively. The learning process is accelerated using PCA because the classifier is trained with fewer features.

Table 8. Effect of PCA-based feature selection on the proposed MLP using *HLN* = 30 for AMC phase.

NoF	AMC1	AMC2	AMC3	AMC4
2	81.33%	85.66%	50.33%	66.66%
1	86%	89.77%	51.11%	67.66%

4.3. Comparative Classification Results

This section includes a summary of the results for RFI classification and characterization. Figure 5 demonstrates the comparative analysis of deploying the proposed approaches for the first level classification to recognize the type of received signals. All the three approaches reach performance accuracy more than 90%, while MLP + GA with five features slightly outperforms other techniques.

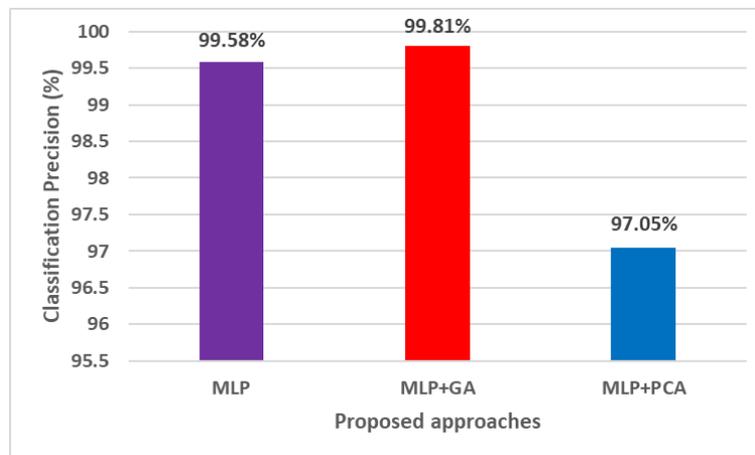


Figure 5. Comparative results of the RFI classification results with considering BS = 1 and HLN = 30 MLP (with all features), MLP + GA ($NoF = 5$) and MLP + PCA ($NoF = 5$).

Figure 6 indicates the results of the proposed classifiers for the second classification process to distinguish four modulation types (QPSK, 8APSK, 16APSK and 32APSK). As the results present, in no jamming case, the highest accuracy is achieved using MLP + GA ($NoF = 1$) with an accuracy of 90.83%. For AMC2, MLP + PCA ($NoF = 1$) outperforms other techniques with a precision of 89.77%. In the presence the two other jammers such as MCWI and CI, MLP + GA performs more precisely.

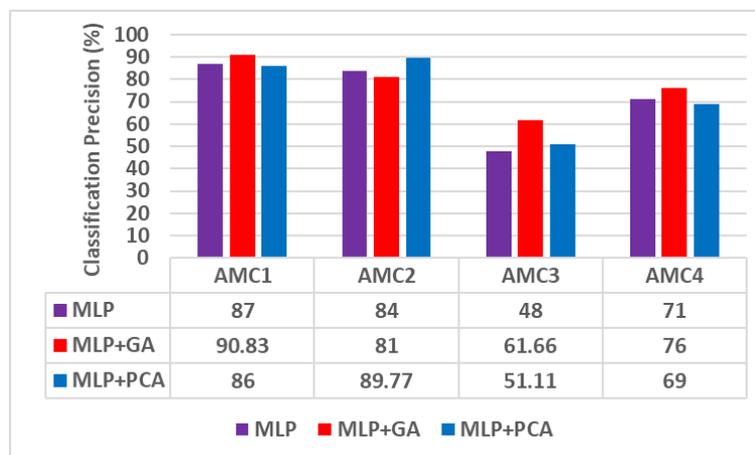


Figure 6. Comparative results of the RFI classification results with considering BS = 1 and HLN = 30 MLP (with all features), MLP + GA ($NoF = 5$) and MLP + PCA ($NoF = 5$).

Figure 7 illustrates the computation times of the presented techniques for the RFI classification. To this end, we used the “tic toc” function in MATLAB to calculate the consumed training time. As the results show, GA + MLP is computationally more expensive, and as a result, the computation time is comparatively longer, while MLP + PCA is the most efficient one.

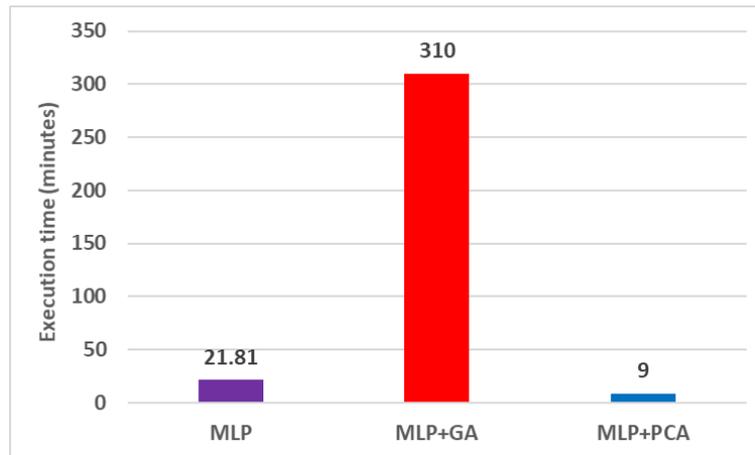


Figure 7. Comparative training time for MLP (21.81 min), MLP + GA (310 min) and MLP + PCA (9 min).

4.4. Prediction Phase Results

In this section, the robustness of the proposed techniques is thoroughly analyzed in detecting unseen data generated at different AWGN power ranging from -140 to -125 dBm. The performance of the classifier varies depending on the noise level and the existence of jammers. To detect the type of the received signal, the trained MLP using 10-fold cross-validation has a higher generalization accuracy in detecting unseen data generated at the same noise level as the training data. Therefore, the prediction accuracy degrades by increasing AWGN power. Table 9 shows the results of evaluating the trained classifier's performance to recognize the type of received signal at different AWGN powers.

Table 9. Evaluating the robustness of the trained networks in predicting unseen data at different noise levels using BS = 1 and HLN = 30.

AWGN Power (dBm)	-140	-135	-130	-125
MLP + 10-fold cross-validation	85.63%	78.70%	60%	55%
MLP + PCA	82%	72.30%	70%	66.90%
MLP + GA	81.56%	73.44%	71%	63.25%

Tables 10–13 show the effect of noise on the trained classifiers in predicting the type of new data. As is shown, two main factors concerning the prediction accuracy are the type of jammer and noise power. Apparently, in the presence of MCWI and CI, the classifiers cannot perform precisely.

Table 10. ML-based prediction results for AMC1.

AWGN Power (dBm)	-140	-135	-130	-125
MLP + 10-fold cross-validation	81%	53%	34.73%	15.30%
MLP + PCA	78%	55%	35.70%	15.70%
MLP + GA	83.50%	54%	36%	15.25%

Table 11. ML-based prediction results for AMC2.

AWGN Power (dBm)	-140	-135	-130	-125
MLP + 10-fold cross-validation	81.33%	75.33%	50.50%	42.50%
MLP + PCA	84%	76%	53%	43%
MLP + GA	82.33%	76.50%	54.23%	44%

Table 12. ML-based prediction results for AMC3.

AWGN Power (dBm)	−140	−135	−130	−125
MLP + 10-fold cross-validation	52.12%	48.50%	48.50%	48%
MLP + PCA	57%	50%	49.50%	48.50%
MLP + GA	58.25%	47%	48.33%	47%

Table 13. ML-based prediction results for AMC4.

AWGN Power (dBm)	−140	−135	−130	−125
MLP + 10-fold cross-validation	61%	57.33%	53.50%	52%
MLP + PCA	55.50%	48.50%	48%	48.50%
MLP + GA	60%	54.33%	47%	50%

5. Conclusions

In this study, a novel hierarchical classifier has been proposed to facilitate the RFI classification and characterization. The proposed classifier not only classifies the type of the received signal but also determines its modulation precisely. Moreover, three robust approaches have been developed, namely MLP + 10-fold cross-validation, MLP + GA, and MLP + PCA. The results confirm that for the RFI classification phase, the classification accuracy reaches 99.81% using the MLP + GA technique, depending on the chosen batch size. In general, MLP + GA performs more precisely in determining the modulation type and RFI classification. Moreover, the results approve that PCA-based feature selection is more efficient in terms of computation time and computational complexity versus GA. In future studies, we intend to look at deep learning-based classification techniques using raw received data to avoid complicated feature extraction and selection steps.

6. Materials

The raw RFI dataset is available at <https://zenodo.org/record/3819586#315.XriaSGhKh3h>. In this dataset, Signal of Interest (SoI) is a real-time video stream that is transmitted using DVB-S2 standards in four modulation types, including (QPSK, 8/16/32 APSK). Furthermore, this SoI combined with three well-known jamming signals, namely Continuous Wave Interference (CWI), Multiple CWI (MCWI), and Chirp Interference (CI). This dataset includes 300 samples per modulation type for each type of signal. Therefore, there are 4800 observations in the dataset, and each sample is a vector of size 1 by 32488 (8 ms) at sample frequency 40 Hz. Also, AWGN power is −140 dBm which is approximately equal to SNR = 9 dB.

Author Contributions: The overall study supervised by R.J.L.; Methodology, Software and preparing the original draft by S.U.; review and editing by N.N.; The results were analyzed and validated by R.J.L. and N.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research is part of the project entitled AVIO-601 in LASSENA Lab (École de Technologie Supérieure) named Interference Mitigation in Satellite Communication. It is supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), Thales, Telesat, VIGILANT GLOBAL, CRIAQ, and Atem Canada. Special thanks to CMC for providing the required equipment to succeed with this project

Conflicts of Interest: The authors declare no conflict of interest.

References

- Zheng, G.; Krikidis, I.; Masouros, C.; Timotheou, S.; Toumpakaris, D.A.; Ding, Z. Rethinking the role of interference in wireless networks. *IEEE Commun. Mag.* **2014**, *52*, 152–158. [[CrossRef](#)]
- Jahanshahi, J.A.; Ghorashi, S.A.; Eslami, M. A support vector machine based algorithm for jamming attacks detection in cellular networks. In Proceedings of the 2011 Wireless Advanced IEEE, London, UK, 20–22 June 2011; pp. 180–184.

3. Rob, R. International Satellite Communication Workshop: The ITU-Challenges in the 21st Century: Preventing Harmful Interference to Satellite Systems. 2013. Available online: <https://www.itu.int/en/ITU-R/space/workshops/2013-interference-geneva/Pages/default.aspx> (accessed on 20 April 2020).
4. Knoblock, E.J.; Madanayake, A. Assessment of Cognitive Communications Interest Areas for NASA Needs and Benefits. In Proceedings of the IEEE Cognitive Communications for Aerospace Applications Workshop, Cleveland, OH, USA, 27–28 June 2017.
5. Lim, A.; Yang, Q. Jamming and anti-jamming techniques in wireless networks: A survey. *Int. J. Ad Hoc Ubiquitous Comput.* **2014**, *17*, 197–215.
6. Yuanyuan, W.; Hao, C. Interference Recognition Scheme Based on Blind Signal Separation in Satellite Communications. *Int. J. Future Comput. Commun.* **2015**, *4*, 395.
7. Haykin, S.; Thomson, D.J.; Reed, J.H. Spectrum sensing for cognitive radio. *Proc. IEEE* **2009**, *97*, 849–877. [[CrossRef](#)]
8. Getu, T.M.; Ajib, W.; Landry, R. Energy-based RFI detection: Theory and results. In Proceedings of the 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Limassol, Cyprus, 15–17 October 2018; pp. 161–168.
9. Salahdine, F.; El Ghazi, H.; Kaabouch, N.; Fihri, W.F. Matched filter detection with dynamic threshold for cognitive radio networks. In Proceedings of the International Conference on Wireless Networks and Mobile Communications (WINCOM), Marrakech, Morocco, 20–23 October 2015; pp. 1–6.
10. Arjoun, Y.; Kaabouch, N. A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions. *Sensors* **2019**, *19*, 126. [[CrossRef](#)]
11. Cabric, D.; Mishra, S.M.; Brodersen, R.W. Implementation issues in spectrum sensing for cognitive radios. In Proceedings of the Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, Pacific Grove, CA, USA, 7–10 November 2004; pp. 772–776.
12. Nawaz, T.; Campo, D.; Mughal, M.O.; Marcenaro, L.; Regazzoni, C.S. Jammer detection algorithm for wide-band radios using spectral correlation and neural networks. In Proceedings of the 13th International Wireless Communications and Mobile Computing Conference (IWCMC), Valencia, Spain, 26–30 June 2017; pp. 246–251.
13. Bavafa, M.; Navidi, N.; Monsef, H. A new approach for Profit-Based Unit Commitment using Lagrangian relaxation combined with ant colony search algorithm. In Proceedings of the 2008 43rd International Universities Power Engineering Conference, Padova, Italy, 1–4 September 2008; pp. 1–6.
14. Tipping, M.E.; Bishop, C.M. Probabilistic principal component analysis. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1999**, *61*, 611–622. [[CrossRef](#)]
15. Jaadi, Z. A Step by Step Explanation of Principal Component Analysis. Available online: <https://builtin.com/data-science/step-step-explanation-principal-component-analysis> (accessed on 20 April 2020).
16. Arjoun, Y.; Salahdine, F.; Islam, M.S.; Ghribi, E.; Kaabouch, N. A Novel Jamming Attacks Detection Approach Based on Machine Learning for Wireless Communication. In Proceedings of the 2020 International Conference on Information Networking (ICOIN), Barcelona, Spain, 7–10 January 2020; pp. 459–464. [[CrossRef](#)]
17. Idhammad, M.; Afdel, K.; Belouch, M. Dos detection method based on artificial neural networks. *Int. J. Adv. Comput. Sci. Appl.* **2017**, *8*, 465–471. [[CrossRef](#)]
18. Jayakumar, K.; Revathi, T.; Karpagam, S. Intrusion Detection using Artificial Neural Networks with Best Set of Features. *Int. Arab J. Inf. Technol. IAJIT* **2015**, *12*. Available online: [https://iajit.org/PDF/Vol%2012,%20No.%207%20\(Special%20Issue\)/6652.pdf](https://iajit.org/PDF/Vol%2012,%20No.%207%20(Special%20Issue)/6652.pdf) (accessed on 20 April 2020).
19. Amato, F.; Mazzocca, N.; Moscato, F.; Vivenzio, E. Multilayer perceptron: An intelligent model for classification and intrusion detection. In Proceedings of the 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA), Taipei, Taiwan, 27–29 March 2017; pp. 1–6.
20. Siaterlis, C.; Maglaris, B. Detecting DDoS attacks using a multilayer Perceptron classifier. In Proceedings of the ICANN'10 20th International Conference on Artificial Neural Networks, Part III, Athens, Greece, March 2004; pp. 118–123.
21. Ujan, S.; Same, M.H.; Landry, R., Jr. A Robust Jamming Signal Classification and Detection Approach Based on Multi-Layer Perceptron Neural Networks. *Int. J. Res. Stud. Comput. Sci. Eng. IJRSCSE* **2020**, *7*, 12. [[CrossRef](#)]

22. Dobre, O.A.; Abdi, A.; Bar-Ness, Y.; Su, W. Survey of automatic modulation classification techniques: Classical approaches and new trends. *Electron. Multidiscip. Digit. Publ. Inst. IET* **2007**, *1*, 137–156. [CrossRef]
23. Farhang, M.; Dehghani, H.; Bahramgiri, H. Multi-receiver modulation classification for satellite communications signals. In Proceedings of the IEEE International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, 16–18 November 2011; pp. 569–573.
24. Shamsi, N.; Mousavinia, A.; Amirpour, H. A channel state prediction for multi-secondary users in a cognitive radio based on neural network. In Proceedings of the IEEE International Conference on Electronics, Computer and Computation (ICECCO), Ankara, Turkey, 7–9 November 2013; pp. 200–203.
25. Iversen, A.; Taylor, N.K.; Brown, K.E.; Karstad, J. *Classification of Communication Signals and Detection of Unknown Formats Using Artificial Neural Networks*; Heriot-Watt University: Edinburgh, UK, 2006. Available online: <https://apps.dtic.mil/sti/pdfs/ADA521144.pdf> (accessed on 20 April 2020).
26. Al-Nuaimi, D.H.; Hashim, I.A.; Zainal, Abidin, I.S.; Salman, L.B.; Mat Isa, N.A. Performance of Feature-Based Techniques for Automatic Digital Modulation Recognition and Classification—A Review. *Electron. Multidiscip. Digit. Publ. Inst.* **2019**, *8*, 1407.
27. Deepa, B.; Iyer, A.P.; Murthy, C.R. Cyclostationary-based architectures for spectrum sensing in IEEE 802.22 WRAN. In Proceedings of the IEEE Global Telecommunications Conference GLOBECOM, Miami, FL, USA, 6–10 December 2010; pp.1–5.
28. Ghauri, S.A.; Qureshi, I.M.; Malik, A.N.; Cheema, T.A. Automatic digital modulation recognition technique using higher order cummulants on faded channels. *J. Basic Appl. Sci. Res.* **2014**, *4*, 1–12.
29. Smith, A.; Evans, M.; Downey, J. Modulation classification of satellite communication signals using cummulants and neural networks. In Proceedings of the Cognitive Communications for Aerospace Applications Workshop (CCAA), Cleveland, OH, USA, 27–28 June 2017; pp. 1–8.
30. Subbarao, M.V.; Samundiswary, P. Automatic Modulation Recognition in Cognitive Radio Receivers using Multi-Order Cumulants and Decision Trees. Available online: <https://www.ijrte.org/wp-content/uploads/papers/v7i4/E1802017519.pdf> (accessed on 20 April 2020).
31. Ran, L.I.; Li, L.O.U.; Ming-Qiang, W.A.N.G. An Approach of modulation recognition using 4th order cyclic cummulants vector of digital signal for cognitive radio. *DEStech Trans. Comput. Sci. Eng.* **2016**. [CrossRef]
32. National Instruments. USRP N210 Kit. Available online: <https://www.ettus.com/all-products/un210-kit/> (accessed on 20 April 2020).
33. Ettus Research. UHD. Available online: <https://kb.ettus.com/UHD> (accessed on 20 April 2020).
34. Solutions, K.D.S. T400CS Channel Simulator. Available online: <https://www.kratosdefense.com/products/space/signals/test-and-simulation/rf-channel-simulator?r=krtl> (accessed on 20 April 2020).
35. The European Space Agency. Types of Orbits. Available online: https://www.esa.int/Enabling_Support/Space_Transportation/Types_of_orbits (accessed on 20 April 2020).
36. Mosiane, O.; Oozer, N. Radio Frequency Interference Detection using Machine Learning. In Proceedings of the IEEE Radio and Antenna Days of the Indian Ocean, St. Gilles-les-Bains, Reunion, 10–13 October 2016; p. 012012.
37. Schoenwald, A.J.; Gholian, A. RFI detection and mitigation using independent component analysis as a pre-processor. In Proceedings of the InRadio Frequency Interference (RFI), Socorro, NM, USA, 17–20 October 2016; pp. 100–104.
38. Llenas, A.M.; Riihijarvi, J.; Petrova, M. Performance Evaluation of Machine Learning Based Signal Classification Using Statistical and Multiscale Entropy Features. In Proceedings of the 2017 IEEE Wireless Communications and Networking Conference (WCNC), San Francisco, CA, USA, 19–22 March 2017.
39. Li, X.; Dong, F.; Zhang, S.; Guo, W. A survey on deep learning techniques in wireless signal recognition. *Wirel. Commun. Mob. Comput.* **2019**, *2019*. [CrossRef]
40. Samanta, B.; Al-Balushi, K.R.; Al-Araimi, S.A. Artificial neural networks and genetic algorithm for bearing fault detection. *Sensors* **2006**, *10*, 264–271.
41. Song, F.; Guo, Z.; Mei, D. Feature selection using principal component analysis. In Proceedings of the International Conference on System Science, Engineering Design and Manufacturing Informatization, Yichang, China, 12–14 November 2010; Volume 1, pp. 27–30.
42. TutorialsPoint. Genetic Algorithms—Quick Guide. Available online: <https://www.tutorialspoint.com/geneticalgorithms/geneticalgorithmsquickguide> (accessed on 20 April 2020).

43. Alsultanny, Y.A.; Aqel, M.M. Pattern recognition using multilayer neural-genetic algorithm. *Sensors* **2003**, *51*, 237–247. [[CrossRef](#)]
44. Hajnayeb, A.; Ghasemloonia, A.; Khadem, S.E.; Moradi, M.H. Application and comparison of an ANN-based feature selection method and the genetic algorithm in gearbox fault diagnosis. *Expert Syst. Appl.* **2011**, *38*, 10205–10209. [[CrossRef](#)]
45. Liu, Z.; Liu, A.; Wang, C.; Niu, Z. Evolving neural network using real coded genetic algorithm (GA) for multispectral image classification. *Future Gener. Comput. Syst.* **2004**, *20*, 1119–1129. [[CrossRef](#)]
46. Guha, R.; Ghosh, M.; Kapri, S.; Shaw, S.; Mutsuddi, S.; Bhateja, V.; Sarkar, R. Deluge based Genetic Algorithm for feature selection. *Evol. Intell.* **2019**, 1–11. [[CrossRef](#)]
47. Chtioui, Y.; Bertr, ; D.; Barba, D. Feature selection by a genetic algorithm. Application to seed discrimination by artificial vision. *J. Sci. Food Agric.* **2019**, *76*, 77–86. [[CrossRef](#)]
48. Guo, H.; Jack, L.B.; Nandi, A.K. Feature generation using genetic programming with application to fault classification. *IEEE Trans. Syst. Man, Cybern. Part B Cybern.* **2005**, *35*, 89–99. [[CrossRef](#)] [[PubMed](#)]
49. Filippidis, A.; Jain, L.C.; Martin, N.M. Using genetic algorithms and neural networks for surface land mine detection. *IEEE Trans. Signal Process.* **1999**, *47*, 176–186. [[CrossRef](#)]
50. Hussein, F.; Kharma, N.; Ward, R. Genetic algorithms for feature selection and weighting, a review and study. In Proceedings of Sixth International Conference on Document Analysis and Recognition, Seattle, WA, USA, 13 September 2001; pp. 1240–1244.
51. Vafaie, H.; De Jong, K.A. Genetic Algorithms as a Tool for Feature Selection in Machine Learning. In Proceedings of the ICTAI, Arlington, VA, USA, 10–13 November 1992; pp. 200–203.
52. Brownlee, J. Gentle Introduction to k-Fold Cross Validation. Available online: <https://fmachinelearningmastery.com/k-fold-cross-validation/> (accessed on 20 April 2020).



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).