Article

Hierarchical Classification Method for Radio Frequency Interference Recognition and Characterization

Sahar Ujan *, Neda Navidi 🆔 **and Rene Jr Landry**

1   LASSENA Laboratory, Ecole de Technologie Superieure (ETS), 1100 Notre-Dame Street West, Montreal, QC, Canada H3C 1K3; sahar.ujan@lassena.etsmtl.ca, neda.nav3@gmail.com, renejr.landry@etsmtl.ca
*   Correspondence: sahar.ujan@lassena.etsmtl.ca

Abstract: Satellite communication (Satcom) is an artificial geostationary satellite that facilitates a wide range of telecommunications. Considering its quality of service (QoS) and security is crucial in government/military applications. The most challenging situation for efficient Satcom is radio frequency interference (RFI) environment. Thus, it is necessary to ensure that transmissions are incorruptible or at least sense the quality of its spectrum. This paper presents a new method to recognize received signal characteristics using a hierarchical classification in a multi-layer perceptron neural network. We consider signal modulation and the type of RFI as the characteristics of a real-time video stream transmitted in the direct broadcast satellite. Four different modulation types are investigated in this study. Moreover, the combination of the communication signal with various kinds of interference and their effects on the classification method widely have been analyzed. Besides, two robust feature selection techniques have been developed to reduce the data-set dimensional, which leads to optimizing the classification process. The results show that the Genetic Algorithm (GA) slightly outperforms Principal Component Analysis (PCA) for feature selection. Furthermore, the robustness of the proposed techniques is assessed to detect unknown signals at different signal to noise ratios.

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 anti-jamming methods in both civil and military applications. Anti-jamming 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 [5]. 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 Signal-to-Noise Ratio (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.

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 [12]. 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 [13]-[30].

This paper presents a hierarchical classification approach using the multi-layer perceptron (MLP), which includes a combination of signal detection and Automatic Modulation Classification (AMC) to facilitate the RFI signal processing. For the design, the first layer is used in detecting the type of the received signal, and the second one is used in recognizing its modulation type.

The data used in this study 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. Quadrature Phase-Shift Keying (QPSK) and Asymmetric Phase-Shift Keying (APSK) with a finite number (8, 16, and 32 in this paper) of distinct signals are used to represent digital data. Further, the studied SoI is combined with one of the three major jamming signal types, namely, CWI, MCWI, and CI. This study considers the Genetic Algorithm (GA) and Principal Component Analysis (PCA) to select informative features for an efficient classification process. GA can perform a global search and explore the search spaces using a different kind of crossover. Also, it can be combined with a local search method to increase the exploitation search [31]. Also, 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 [32].

The rest of the paper is organized as follows: Section 2 reviews the previous researches 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 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. In [5], the authors present the Random Forest classifier (RFC) for jamming detection in IEEE 802.11 wireless networks and compare the RFC performance to other classifiers, including Decision Tree (DT), Adaptive Boosting (AdaBoost), the Support Vector Machine (SVM) and Expectation-Maximization (EM). As the results illustrate, the RFC and SVM as classifiers slightly outperform the other techniques. In [15], different feature selection techniques are analyzed, and finally, SVM is deployed for GNSS Spoofing Jamming Recognition at different SNRs in the range of 1 to 10 dB. An Artificial Neural Network (ANN) is proposed in [11] 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 $-3\,dB$.

Recently, in a few related works [14], [16]-[21], [30], MLP with different designs have been presented to recognize different kinds of jamming attacks. The focus of all these studies is on recognizing either the type of received signal or its modulation type. However, in this paper, a new methodology is proposed to detect and characterize the received signal based on a hierarchical classification.

Automatic modulation classification (AMC) is a significant procedure for present and next-generation communication networks and facilitates the demodulation process at the receiver side [31]. 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 [22], and the technology is widely used in various applications, such as dynamic spectrum management [23] and interference identification in CR [22].

As was thoroughly discussed in [32], 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 [22] as it leverages robust extracted features such as an instantaneous amplitude, phase, and frequency [33], cyclostationary features [25], higher-order cumulants [26], and spectral correlation features [34]. In [22] and [27], 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 [22] 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 in order 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 $(-4 \sim 4)\ dB$. In [28], 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 [29], 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.

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.

3.1. Data acquisition

As thoroughly has been explained in [30], 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 tool-kit that provides signal processing blocks to implement software radios. The SoI is transmitted using a Universal Software Radio Peripheral (USRP-N210) [37], which is an enhanced version of the USRP that includes a larger FPGA. This facility allows users to move additional functionality into the FPGA, increasing the maximum processing capability in both communication directions while offering potential improvements in processing latency. In GNU radio, the modulation type and amplitude of the transmitted signal can be easily adjusted [38].

A Satcom emulator [39] (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. The Kratos STK plugin provides real-time, phase-continuous control of the channel simulator when playing STK scenarios [39].

Further, the generated jamming signals are transmitted using a NanoBee modem and are combined with SoI by an additive combiner. Finally, the combined signal is received by a MegaBee modem.

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
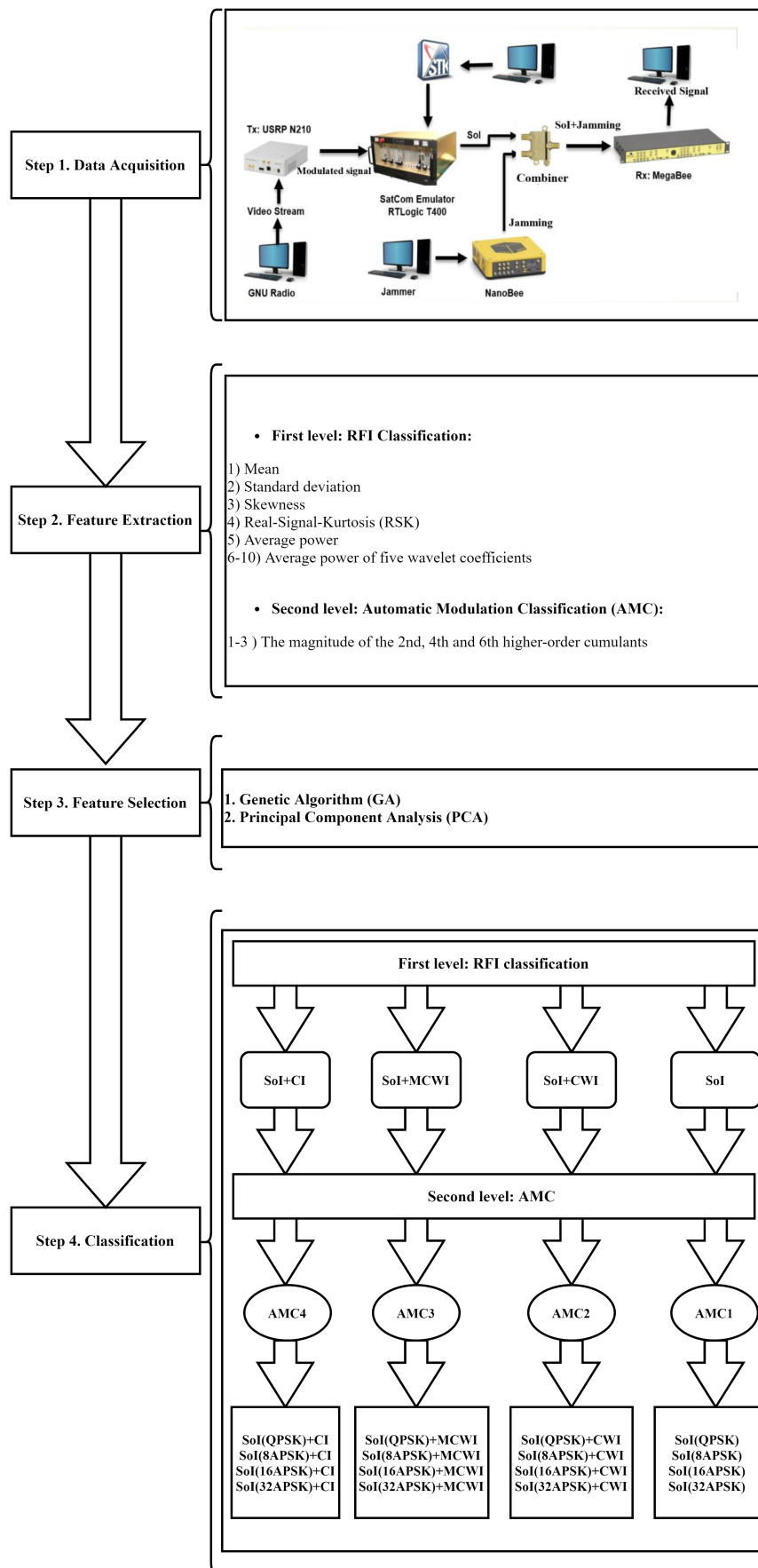
Figure 1. Proposed RFI recognition and characterization methodology

8/16/32 asymmetric phase-shift keying (APSK). Any digital modulation scheme uses a finite number (8, 16, and 32 in this paper) of distinct signals to represent digital data. QPSK and APSK use a finite number of phases, each assigned a unique pattern of binary digits. The length of $8ms$ is considered to generate continuous time-series signals.

Table 1. Real-time data-set characterization

characteristic	Value
Total number of samples	4800
Length of generated signals	32448 (8 ms)
Sampling frequency	$40x10^6 Hz$
Modulation types	QPSK, 8APSK, 16APSK, and 32APSK
AWGN power	$140\ dBm\ (SNR \cong 9dB)$
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 utilize in "RFI classification" and "modulation type recognition," as follows:

• Features for RFI classification:

To address RFI classification phase, by inspiring from [30] the six extracted features 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). Therefore, the size of each feature vector for the RFI classification phase is 1 by 10. The importance of those features is explained in [30] in detail.

• 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 pth order and q^{th} conjugation moment for a received signal x[n] in time-step n are defined as [41]:

$$M_{pq} = E[x[n]^{p-q}(x[n]^*)^q] \tag{1}$$

where $E(.)$ and $[]^*$ are the expected operator and complex conjugate, respectively. Eq. 1 can be approximated as [27]:

$$M_{pq} = \frac{1}{n}\sum_{n=0}^{N-1} x[n]^{p-q}(x^*[n])^q \tag{2}$$

In this work, the magnitude of the 2nd, 4th and 6th [27] cumulants are computed to distinguish between QPSK, 8APSK, 16APSK, and 32APSK:

$$C_{21} = M_{21} = E[|x(n)|^2] \tag{3}$$

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

$$C_{42}(x) = M_{42} - 2M_{21} - |M_{21}|^2 \tag{4}$$

$$C_{63} = M_{63} + 18M_{20}M_{21}M_{22} + 12(M_{21})^3 - 9M_{21}M_{42} - 3M_{20}M_{43} - 3M_{22}M_{41} \tag{5}$$

Notably, different orders of cumulant-based features were tested and the best results were achieved using the above orders ($|C21|$, $|C42|$ and $|C63|$).

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, [42]. However, less number of characteristic features may come at the expense of classification accuracy degradation [42]. However, the trick in dimensionality reduction is to trade some precision for simplicity [43]. 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)

This technique has been used in a wide range of computer science applications for feature extraction [43]. PCA projects a dataset from many correlated coordinates onto fewer uncorrelated coordinates called principal components, with information preserved as much as possible [43]. Moreover, in [44], PCA is presented as an efficient feature selection method with lower computational complexity in comparison to other approaches, like colony optimization [44]. The implementation steps of PCA have been fully explained in [30]. As shown in Algorithm 1, PCA for feature extraction is implemented in four steps.

Firstly, 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 [46]. 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 as 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 [46]. The implementation steps of PCA have been fully explained in [30].

Algorithm 1 PCA for feature selection

Require: The generated dataset
    Compute mean of each column
    Subtracting each value in each column from the mean
    Computation of the covariance matrix
    calculation of eigenvectors and eigenvalues of the covariance matrix

---

3.3.2. Genetic Algorithm (GA)

GA is known as a search-based optimization technique inspired by Genetics and Natural Selection [56]. GA has been widely used not only for feature selection but also for optimizing the hyperparameters of an Artificial Neural Network (ANN), such as weight [42],[47]-[55]. 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 [56]. Each individual has a set of genes, represented by a string of zeros and ones [56]. 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 [56]. After the parent selection step, the variation operations such as mutation and crossover [56] are applied to generate new off-springs [56]. Finally, these off-springs replace the existing individuals in the population, and the process is repeated until reaching the stopping criterion.
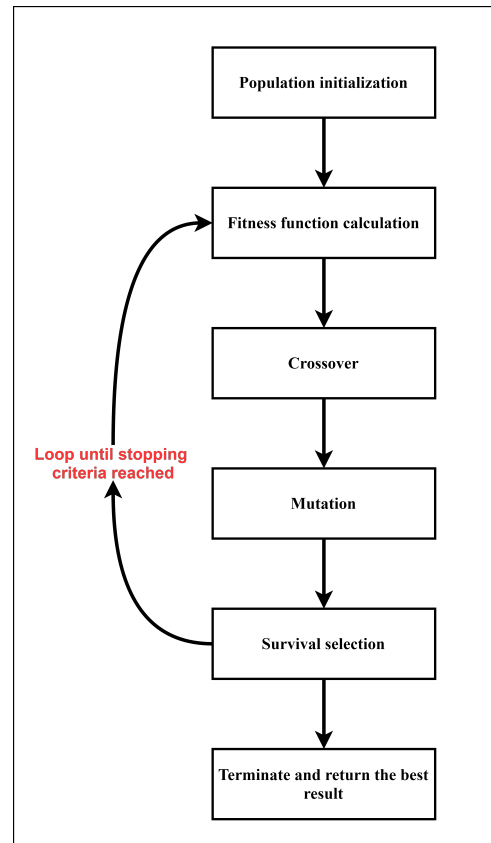
Figure 2. General framework of GA algorithm, modified from [56]

To adapt GA technique with this study, presented in pseudo-code of Algorithm 2, 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 [56].

• 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[56].

• Population initialization: This study considers a random initialization, in which k bits of the string are randomly set to 1. In other words, k features are chosen out of the n-dimensional space of features [56].

• Parent selection: The "Rank Selection" approach is utilized for selecting parents since the individuals in the population have very close fitness values. In this study, the fitness function is the classification accuracy, which is defined as the number of correct detections divided by the total number of detections. 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 [56].

• 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 [53].

• 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,

---

Algorithm 2 Proposed GA for feature selection

Require:
   Initialization:
   Population size
   Mutation rate
   Number of selected parents
   Number of generated children at each iteration
   Maximum number of generations
   Defining number of selected features (k)
   Mutation rate
   for $i = 1$ to $No.\ \ of\ \ population$ **do**

      Generate a vector (with zeros) with the same size of the feature vector
      Randomly set k values to 1.
   end for
   for $J = 1$ to population size **do**

      Compute the fitness value of each individual
   end for
   for $n = 1$ to generation size **do**

      Calculate the sum of the fitness values at each generation
      for $l = 1$ to children size **do**

         selecting the chromosome with highest fitness value from the selected parents
      end for
      for $h = 1$ to Mutation rate **do**

         Sweep two random indices
      end for
      Calculate the the fitness value of the new off-spring
      if new fitness > minimum fitness value **then**

         Replace the individual with the minimum fitness value with the new off-spring
      end if
      Save the individual with the maximum fitness value
   end for
   return the individual with the maximum fitness value as the best feature set

---

a fitness-based selection is used, in which the children tend to replace the least fit individuals in the population [53].

    • Termination criterion: The algorithm is terminated when a set number of generations is reached. Figure 3 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 10 dB mother wavelet.

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, CWI, MCWI, and CI). Further, 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 [30] 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 [40]. 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 [40]. 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 [40]. Algorithm 3 illustrates the implementation steps of k-fold cross-validation as follows
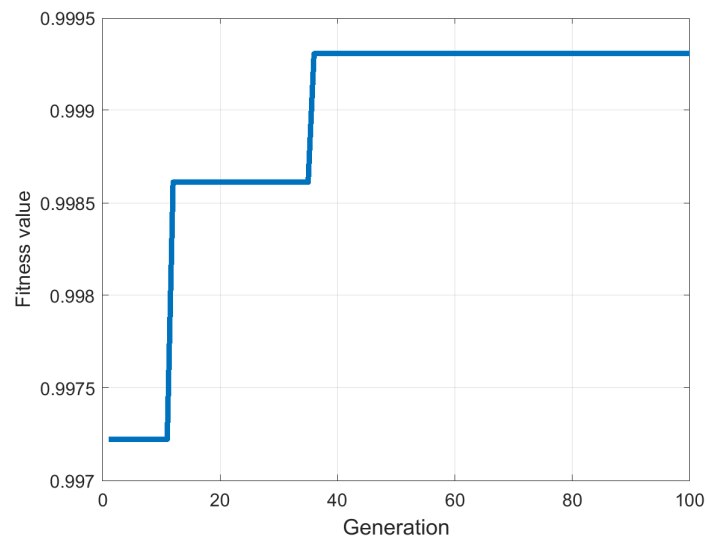
Figure 3. Variation of the best fitness values per generations number for RFI classification application using two features

---
Algorithm 3 Training using $k - fold$ cross-validation

Input: The generated dataset
  Shuffle the dataset randomly
  Split the dataset into k groups
  for each unique group **do**

  Take the group as a holdout or test data set
  Take the remaining groups as a training data set
  Fit a model on the training set and evaluate it on the test set
  Retain the evaluation score and discard the model
  Summarize the skill of the model using the sample model evaluation scores
  end for
  return Evaluation score

As shown in Algorithm 3, 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 [40]. 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 [40]. 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. EXPERIMENTAL RESULTS

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 $RAM = 8GHz$. Moreover, the dataset used for the classification has been generated at AWGN power $-140dBm$, and approximately SNR value is $9dB$. 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

$$JSR = Power_{SoI} - Power_{Jamming} \tag{6}$$

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 [30] 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 utilized as the hidden and output layer activation functions. 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, 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 $GA + MLP$ 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 $PCA + MLP$ 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 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

NF	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

NF	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 4 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 accuracies more than 90%, while $MLP + GA$ with five features slightly outperforms other techniques.

Figure 5 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 illustrates the computation times of the presented techniques for the RFI classification. $GA + MLP$ is computationally more expensive, and as a result, the computation time is comparatively longer, while $MLP + PCA$ is the most efficient one.

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
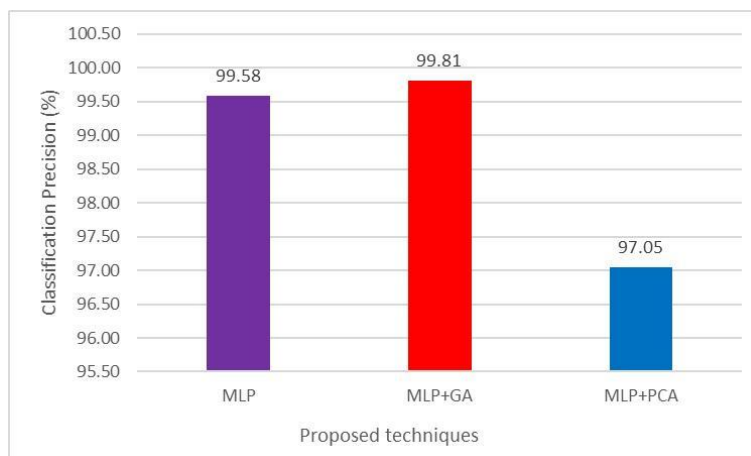
Figure 4. Comparative results of the RFI classification results with considering BS=1 and HLN=30 MLP (with all features), $MLP + GA$ $(NoF = 5)$ and $MLP + PCA$ $(NoF = 5)$

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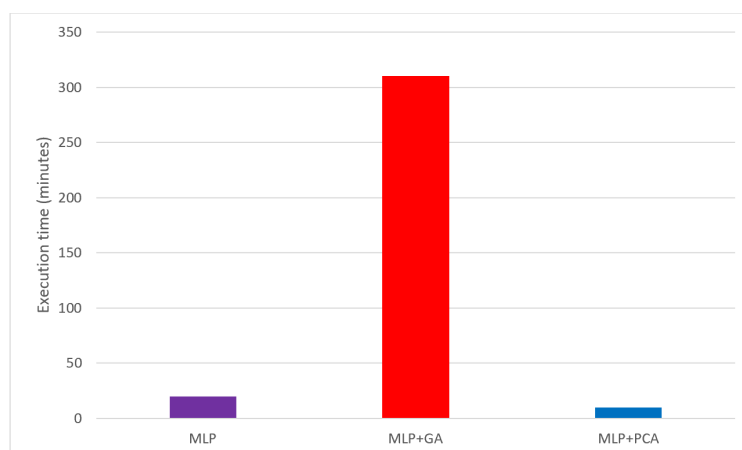

Figure 6. Comparative training time for $MLP (21.81 min)$, $MLP + GA (310 min)$ and $MLP + PCA (9 min)$

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. 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 to 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. Conclusion

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. Besides, the results approve that PCA 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 in order to avoid complicated feature extraction and selection steps.

6. Materials and Methods

The raw RFI dataset including 4800 samples is available at https://zenodo.org/record/3819586#.XriaSGhKh3h.

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.

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

References

1. Zheng, Gan and Krikidis, Ioannis and Masouros, Christos and Timotheou, Stelios and Toumpakaris, Dimitris-Alexandros and Ding, Zhiguo, Rethinking the role of interference in wireless networks, IEEE Communications Magazine,2014, pp. 152–158.
2. Jahanshahi, Javad Afshar and Ghorashi, Seyed Ali and Eslami, Mohammad, A support vector machine based algorithm for jamming attacks detection in cellular networks, 2011 Wireless Advanced IEEE,2011, pp. 180–184.
3. Rideout, Rob, International Satellite Communication Workshop: The ITU-challenges in the 21st Century: Preventing Harmful Interference to Satellite Systems, 2013.
4. Knoblock, Eric J and Madanayake, Arjuna, Assessment of Cognitive Communications Interest Areas for NASA Needs and Benefits, 2017.
5. Puñal, Oscar and Aktaş, Ismet and Schnelke, Caj-Julian and Abidin, Gloria and Wehrle, Klaus and Gross, Jame, Machine learning-based jamming detection for IEEE 802.11: Design and experimental evaluation, Proceeding of IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks,2014, pp. 1–10.
6. Yuanyuan, Wen and Hao, Chen, Interference Recognition Scheme Based on Blind Signal Separation in Satellite Communications, International Journal of Future Computer and Communication, IACSIT Press, Vol: 4 (6), 2015, pp. 395.
7. Haykin, Simon and Thomson, David J and Reed, Jeffrey H, Spectrum sensing for cognitive radio, Proceedings of the IEEE, IACSIT Press, Vol: 97 (5), 2009, pp. 849–877.
8. Getu, Tilahun M and Ajib, Wessam and Landry, René, Energy-based RFI detection: theory and results, 14th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob) IEEE, 2018, pp. 161–168.
9. Salahdine, Fatima and El Ghazi, Hassan and Kaabouch, Naima and Fihri, Wassim Fassi, Matched filter detection with dynamic threshold for cognitive radio networks, international conference on wireless networks and mobile communications (WINCOM) IEEE, 2015, pp. 1–6.
10. Arjoune, Youness and Kaabouch, Naima, A comprehensive survey on spectrum sensing in cognitive radio networks: Recent advances, new challenges, and future research directions, international conference on wireless networks and mobile communications (WINCOM) Sensors,Vol: 19(1), 2019, pp. 126.
11. Nawaz, Tassadaq and Campo, Damian and Mughal, Muhammad Ozair and Marcenaro, Lucio and Regazzoni, Carlo S, Jammer detection algorithm for wide-band radios using spectral correlation and neural networks, 13th International Wireless Communications and Mobile Computing Conference (IWCMC) , 2017, pp. 246–251.
12. Cabric, Danijela and Mishra, Shridhar Mubaraq and Brodersen, Robert W, Implementation issues in spectrum sensing for cognitive radios, Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers IEEE,Vol:1, 2004, pp. 772–776.
13. Karagiannis, Dimitrios and Argyriou, Antonios, Jamming attack detection in a pair of RF communicating vehicles using unsupervised machine learning, Vehicular Communications, Elsevier,Vol:13, 2018, pp. 56–63.
14. Idhammad, Mohamed and Afdel, Karim and Belouch, Mustapha, Dos detection method based on artificial neural networks, International Journal of Advanced Computer Science and Applications,Vol:8(4), 2017, pp. 465–471.

15. Gao, Pan and Sun, Songlin and Zeng, Zhen and Wang, Chenwei, GNSS Spoofing Jamming Recognition Based on Machine Learning, International Conference On Signal And Information Processing, Networking And Computers, Springer, 2017, pp.221–228

16. Jayakumar, Kaliappan and Revathi, Thiagarajan and Karpagam, Sundararajan, Intrusion Detection using Artificial Neural Networks with Best Set of Features., International Arab Journal of Information Technology (IAJIT), 2015, pp.12.

17. Sawant, Anuja A and Game, Pravin S, Approaches for Anomaly Detection in Network: A Survey., 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA),IEEE, 2018, pp.1–6.

18. Amato, Flora and Mazzocca, Nicola and Moscato, Francesco and Vivenzio, Emilio, Multilayer perceptron: An intelligent model for classification and intrusion detection, 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA),IEEE, 2017, pp.1–6.

19. Iversen, Alexander and Taylor, Nicholas K and Brown, Keith E and Karstad, Jorn, Classification of Communication Signals and Detection of Unknown Formats Using Artificial Neural Networks, HERIOT-WATT UNIV EDINBURGH (UNITED KINGDOM), 2006.

20. Siaterlis, Christos and Maglaris, Basil, Detecting DDoS attacks using a multilayer Perceptron classifier, proceedings of ICANN'10 Proceedings of the 20th international conference on Artificial neural networks: Part III, 2004, pp.118–123.

21. Haddadi, Fariba and Khanchi, Sara and Shetabi, Mehran and Derhami, Vali, Intrusion detection and attack classification using feed-forward neural network, 2010 Second International Conference on Computer and Network Technology,2010, pp.262–266.

22. Farhang, Mohsen and Dehghani, Hamid and Bahramgiri, Hossein, Multi-receiver modulation classification for satellite communications signals, IEEE International Conference on Signal and Image Processing Applications (ICSIPA),2011, pp.569–573.

23. Shamsi, Nakisa and Mousavinia, Amir and Amirpour, Hadi, A channel state prediction for multi-secondary users in a cognitive radio based on neural network, IEEE International Conference on Electronics, Computer and Computation (ICECCO),2013, pp.200–203.

24. Mühlhaus, Michael S and Öner, Mengüc and Dobre, Octavia A and Jaekel, Holger U and Jondral, Friedrich K, A novel algorithm for MIMO signal classification using higher-order cumulants, IEEE Radio and Wireless Symposium,2013, pp.7–9.

25. Deepa, B and Iyer, Anand Padmanabha and Murthy, Chandra R, Cyclostationary-based architectures for spectrum sensing in IEEE 802.22 WRAN, IEEE Global Telecommunications Conference GLOBECOM, 2010, pp.1–5.

26. Ghauri, Sajjad Ahmed and Qureshi, IM and Malik, AN and Cheema, TA, Automatic digital modulation recognition technique using higher order cummulants on faded channels, Journal of Basic and Applied Scientific Research, Vol: 4(3), pp.1–12, 2014.

27. Smith, Aaron and Evans, Michael and Downey, Joseph, Modulation classification of satellite communication signals using cumulants and neural networks, Cognitive Communications for Aerospace Applications Workshop (CCAA) IEEE, pp.1–8, 2017.

28. Subbarao, M Venkata and Samundiswary, P, Automatic Modulation Recognition in Cognitive Radio Receivers using Multi-Order Cumulants and Decision Trees.

29. Ran, LI and Li, LOU and Ming-Qiang, WANG, An Approach of modulation recognition using 4th order cyclic cumulants vector of digital signal for cognitive radio, DEStech Transactions on Computer Science and Engineering, 2016.

30. Sahar Ujan and M.h Same and Rene Jr Landry, A Robust Jamming Signal Classification and Detection Approach Based on Multi-Layer Perceptron Neural Networks, International Journal of Research Studies in Computer Science and Engineering (IJRSCSE), Vol: 7(1) pp.12, 2020, http://dx.doi.org/10.20431/2349-4859.0701001.

31. Bavafa, M and Navidi, N and Monsef, H, A new approach for Profit-Based Unit Commitment using Lagrangian relaxation combined with ant colony search algorithm, 2008 43rd International Universities Power Engineering Conference IEEE, pp.1–6, 2008.

32. Tipping, Michael E and Bishop, Christopher M, Probabilistic principal component analysis, Journal of the Royal Statistical Society: Series B (Statistical Methodology), Vol: 61 (3) pp.611–622, 1999.

33. Al-Nuaimi, Dhamyaa H and Hashim, Ivan A and Zainal Abidin, Intan S and Salman, Laith B and Mat Isa, Nor Ashidi, Performance of Feature-Based Techniques for Automatic Digital Modulation Recognition and Classification—A Review, Electronics, Multidisciplinary Digital Publishing Institute, Vol:8 (12) pp.1407, 2019.

34. Dobre, Octavia A and Abdi, Ali and Bar-Ness, Yeheskel and Su, Wei, Survey of automatic modulation classification techniques: classical approaches and new trends, Electronics, Multidisciplinary Digital Publishing Institute IET, Vol:1 (2) pp.137–156, 2007.

35. Fehske, A and Gaeddert, J and Reed, Jeffrey H, A new approach to signal classification using spectral correlation and neural networks, First IEEE International Symposium on New Frontiers in Dynamic Spectrum Access Networks IEEE, DySPAN T, Vol:1 (2) pp. 144–150, 2005.

36. Yang, Zhutian and Wu, Zhilu and Yin, Zhendong and Quan, Taifan and Sun, Hongjian, Hybrid radar emitter recognition based on rough k-means classifier and relevance vector machine, Electronics, Multidisciplinary Digital Publishing Institute, Sensors, Vol:13 (1) pp. 848–864, 2013.

37. National Instruments,USRP N210 Kit[online], https://www.ettus.com/all-products/un210-kit/ [Accessed 20 April. 2020]

38. Ettus Research,UHD[online], https://kb.ettus.com/UHD [Accessed 20 April. 2020]

39. K. D. S. Solutions,T400CS Channel Simulator.[online],https://www.kratosdefense.com/products/space/signals\/test-and-simulation/rf-channel-simulator?r=krtl[Accessed 20 April. 2020]

40. J. Brownlee,Gentle Introduction to k-fold Cross validation[online], https://www.kratosdefense.com/products/space/signals/test-and-simulation/rf-channel-simulator?r=krtl [Accessed 20 April. 2020]

41. Li, Xiaofan and Dong, Fangwei and Zhang, Sha and Guo, Weibin, A survey on deep learning techniques in wireless signal recognition, Wireless Communications and Mobile Computings,Hindawi, 2019.

42. Samanta, Biswanath and Al-Balushi, Khamis R and Al-Araimi, Saeed A, Artificial neural networks and genetic algorithm for bearing fault detection, Electronics, Multidisciplinary Digital Publishing Institute, Sensors, Vol:10 (3) pp.264–271, 2006.

43. Z. Jaadi,A STEP BY STEP EXPLANATION OF PRINCIPAL COMPONENT ANALYSIS[online], https://builtin.com/data-science/step-step-explanation-principal-component-analysis [Accessed 20 April. 2020]

44. F. Song, Z. Guo, and D. Mei, Feature selection using principal component analysis, in international conference on system science, engineering design and manufacturing informatizatio, Vol:1, pp.27–30, 2010.

45. A. PERUNICIC,HOW ARE PRINCIPAL COMPONENT ANALYSIS AND SINGULAR VALUE DECOMPOSITION RELATED?[online], https://intoli.com/blog/pca-and-svd/ [Accessed 20 April. 2020]

46. Zakaria Jaadi,A STEP BY STEP EXPLANATION OF PRINCIPAL COMPONENT ANALYSIS[online], https://builtin.com/data-science/step-step-explanation-principal-component-analysis[Accessed 20 April. 2020]

47. Alsultanny, Yas Abbas and Aqel, Musbah M, Pattern recognition using multilayer neural-genetic algorithm, Neurocomputing, Multidisciplinary Digital Publishing Institute Elsevier, Sensors, Vol:51, pp. 237–247, 2003.

48. Hajnayeb, Ali and Ghasemloonia, Ahmad and Khadem, SE and Moradi, MH, Application and comparison of an ANN-based feature selection method and the genetic algorithm in gearbox fault diagnosis, Expert systems with Applications,Elsevier, Vol: 38(8), pp. 10205–10209, 2011.

49. Liu, Zhengjun and Liu, Aixia and Wang, Changyao and Niu, Zheng, Evolving neural network using real coded genetic algorithm (GA) for multispectral image classification, Future Generation Computer Systems,Elsevier, Vol: 20(7), pp. 1119–1129, 2004.

50. Guha, Ritam and Ghosh, Manosij and Kapri, Souvik and Shaw, Sushant and Mutsuddi, Shyok and Bhateja, Vikrant and Sarkar, Ram, Deluge based Genetic Algorithm for feature selection, Evolutionary intelligence,Springer, pp.1–11 , 2019.

51. Chtioui, Younes and Bertrand, Dominique and Barba, Dominique, Feature selection by a genetic algorithm. Application to seed discrimination by artificial vision, Journal of the Science of Food and Agriculture, Vol:76(1), pp.77–86 , 2019.

52. Guo, Hong and Jack, Lindsay B and Nandi, Asoke K, Feature generation using genetic programming with application to fault classification, IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), Vol:35(1), pp.89–99 , 2005.

53. Filippidis, Arthur and Jain, Lakhmi C and Martin, Noel M, Using genetic algorithms and neural networks for surface land mine detection, IEEE Transactions on Signal Processing, Vol:47(1), pp.176–186 , 1999.

54. Hussein, Faten and Kharma, Nawwaf and Ward, Rabab, Genetic algorithms for feature selection and weighting, a review and study, Proceedings of Sixth International Conference on Document Analysis and Recognition IEEE, pp.1240–1244 , 2001.

55. Vafaie, Haleh and De Jong, Kenneth A, Genetic Algorithms as a Tool for Feature Selection in Machine Learning, ICTAI, pp.200–203 , 1992.

56. TutorialsPoint,Genetic Algorithms - Quick Guide[online], https://www.tutorialspoint.com/geneticalgorithms/geneticalgorithmsquickguide [Accessed 20 April. 2020]