

Investigation on Analog and Digital Modulations Recognition Using Machine Learning Algorithms

Jean Ndoumbe^{1*}, Ivan Basile Kabeina¹, Gaelle Patricia Talotsing², Soubiel-Noël Nkomo Biloo¹

¹Laboratory of Computer Engineering, Data Science and Artificial Intelligence, National Higher Polytechnic School of Douala, Douala, Cameroon

²Mobile Computing and Networking Research Laboratory, Department of Computer and Software Engineering, Polytechnique Montreal, Montreal, Canada

Email: *jndoumbe02@gmail.com

How to cite this paper: Ndoumbe, J., Kabeina, I.B., Talotsing, G.P. and Biloo, S.-N.N. (2024) Investigation on Analog and Digital Modulations Recognition Using Machine Learning Algorithms. *World Journal of Engineering and Technology*, 12, 867-884.

<https://doi.org/10.4236/wjet.2024.124053>

Received: August 20, 2024

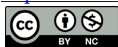
Accepted: September 27, 2024

Published: September 30, 2024

Copyright © 2024 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0).

<http://creativecommons.org/licenses/by-nc/4.0/>



Open Access

Abstract

In the field of radiocommunication, modulation type identification is one of the most important characteristics in signal processing. This study aims to implement a modulation recognition system on two approaches to machine learning techniques, the K-Nearest Neighbors (KNN) and Artificial Neural Networks (ANN). From a statistical and spectral analysis of signals, nine key differentiation features are extracted and used as input vectors for each trained model. The feature extraction is performed by using the Hilbert transform, the forward and inverse Fourier transforms. The experiments with the AMC Master dataset classify ten (10) types of analog and digital modulations. AM_DSB_FC, AM_DSB_SC, AM_USB, AM_LSB, FM, MPSK, 2PSK, MASK, 2ASK, MQAM are put forward in this article. For the simulation of the chosen model, signals are polluted by the Additive White Gaussian Noise (AWGN). The simulation results show that the best identification rate is the MLP neuronal method with 90.5% of accuracy after 10 dB signal-to-noise ratio value, with a shift of more than 15% from the k-nearest neighbors' algorithm.

Keywords

Automatic Recognition, Artificial Neural Networks, K-Nearest Neighbors, Machine Learning, Analog Modulations, Digital Modulations

1. Introduction

Automatic Modulation Recognition or Classification (AMR/AMC) is a technique for classifying or identifying the type of modulation of the detected signal and for

estimating certain signal parameters such as power, symbol rate, carrier frequency, etc...

In the last decade, advanced analog and digital modulation techniques have been used in communications systems with the aim of improving spectrum efficiency and transmission reliability [1]. In analog communication systems, a transmission signal is encoded by using analog modulations like amplitude modulation (AM), phase modulation (PM), and frequency modulation (FM). Typically, an analog modulation technique encodes an analog baseband signal onto a high frequency periodic waveform. In digital communication, the signals are digitized, and then the information is transferred through these digitized signals from source to destination. Digital modulations are more preferable in terms of usage, thanks to better coordination with digital data and stronger robustness against interference compared with analog modulations [2]. Different types of digital modulations are employed, these included amplitude shift keying (ASK), phase-shift keying (PSK), frequency shift keying (FSK), pulse amplitude modulation (PAM), amplitude and phase-shift keying (APSK), and quadrature amplitude modulation (QAM). Due to the increase in analogue and digital modulations related to civil and military applications, modulation recognition is important, as recognition of the modulation type is an intermediate step between signal detection and signal demodulation [2]-[4]. It is a very important step in reception as it allows selection of the corresponding demodulator. The use of machine learning techniques has made the recognition process easier as well as more reliable. Hence, many works have attempted to apply Machine Learning (ML) techniques to AMC [2]. Hong and Ho [5] used Bayes method to classify BPSK and QPSK without a priori information about the received signal level. Wong *et al.* [6] introduced minimum distance classifier to reduce the complexity of Machine learning classifier. They also used blind source separation algorithm. Swami and Sadler [7] proposed a method for digital modulation classification using fourth order cumulants. Their method was robust to the presence of carrier phase and frequency offsets. Author in [8] used higher-order statistics for blind channel estimation and pattern recognition. They presented results of modulation classification in the presence of a fading channel. Headley *et al.* [9] used a two-stage process for AMC. In the first stage, local radios make a decision which is then sent to a fusion center, which makes a global decision about the modulation type. Wong and Nandi [10] used artificial neural networks (ANN) and genetic algorithms for recognition of various digital modulation types. They used two types of ANNs; multi-layer perceptron (MLP) and resilient back propagation (RPROP). They also used Bayes method with higher-order cumulants for classification of BPSK, QPSK, 16QAM and 64QAM.

Hassanpour *et al.* [11] proposed a pattern-based AMC approach. The authors focused on feature extraction blocks and digital modulation pattern identification, assuming AWGN, BASK, BFSK, BPSK, 4-ASK, 4-FSK, QPSK, and 16-QAM will be employed. To extract their attributes, signs are analyzed in time, frequency,

and wavelet domains. A Binary SVM-based hierarchical structure is being studied to resolve the problem of multiple classes. Simulations show how the suggested characteristics enhance digital signal differentiation in a noisy environment with low SNR. They obtained, a 98.15 percent accuracy rate with an SNR of -10 dB, proving that this is the minimum required for flawless identification.

In their studies, Ansari *et al.* [12] and Jajoo *et al.* [13] proposed a modulation recognition and separation method by choosing useful features from the modulated signal for identifying the modulation type using a decision tree and probabilistic neural network methods. They were conducted using MATLAB simulations with a signal-to-noise ratio over an additive white Gaussian noise channel. The results showed that picking useful features and establishing tuning parameters significantly improved modulation type recognition accuracy and speed.

With the development of machine learning, more and more people have begun to use supervised learning to improve the efficiency of signal recognition [14]. Daldal *et al.* used a deep LSTM model to identify six modulation types of signal sequences and achieved significant results [15]. Cheol-Sun Park uses Support Vector Machine (SVM) to improve the recognition effect [16]. Timothy J. O'Shea [17] proposed convolutional neural network to solve the problem of modulation recognition.

In several recent works, many types of datasets have been used to evaluate the performance of machine-learning algorithms for modulation classification. A. Jagannath and J. Jagannath [18] proposed a synthetic wireless waveform dataset suited for modulation recognition and wireless signal (protocol) classification tasks separately as well as jointly. Cheng *et al.* [19] shared a dataset consisting of four different modulation schemes and varying levels of AWGN and fading. They evaluated the performance of several deep learning algorithms on this dataset and found that Convolutional Neural Networks (CNN)s and Recurrent Neural Networks (RNN) achieved the highest classification accuracy. Alam *et al.* [20] used a dataset involving eight modulation schemes and varying levels of Additive White Gaussian Noise (AWGN). The performance of several machine learning and deep learning algorithms was investigated on this dataset and found that deep learning algorithms outperformed traditional machine learning algorithms in high-noise environments. **Table 1** presents the summary of related works.

Table 1. Summary of related works.

Reference	Modulations Studied	Method Used	Limitations
[5] Hong and Ho	BPSK, QPSK	Bayes method	Lacks consideration of varying SNR conditions.
[6] Wong <i>et al.</i>	Digital modulation: ASK, PSK, FSK, QAM	Minimum distance classifier, blind source separation	Complexity reduction impacts recognition accuracy in noisy environments.
[7] Swami and Sadler	Digital modulation: ASK, PSK, FSK, QAM	Fourth-order cumulants	Susceptible to carrier phase and frequency offsets in high noise environments.
[9] Headley <i>et al.</i>	Multiple digital modulations	Two-stage AMC process	Relies heavily on local decisions, potentially reducing robustness in diverse conditions.

Continued

[10] Wong and Nandi	BPSK, QPSK, 16QAM, 64QAM	ANN, Genetic algorithms, Bayes method	They can be computationally intensive, impacting real-time application feasibility.
[11] Hassanpour <i>et al.</i>	BASK, BFSK, BPSK, 4-ASK, 4-FSK, QPSK, 16-QAM	Binary SVM-based hierarchical structure	May struggle with feature extraction in extremely low SNR environments.
[12] Ansari <i>et al.</i>	Various digital modulations	Decision tree, Probabilistic neural network	Performance highly dependent on feature selection and tuning parameters.
[15] Daldal <i>et al.</i>	Six modulation types	Deep LSTM model	Requires extensive training data, and computational resources may limit deployment.
[16] Cheol-Sun Park	Various digital modulations	Support Vector Machine (SVM)	May not generalize well across all modulation types without extensive tuning.
[17] Timothy J. O’Shea	Various digital modulations	Convolutional Neural Network (CNN)	High computational complexity may limit real-time application, especially in noisy channels.

In this paper, we therefore analyze modulation systems and thus experiment their recognition on KNN and ANN automatic approaches to classify ten types of modulation. The studied modulations include: AM_DSB_FC, AM_DSB_SC, AM_USB, AM_LSB, FM, MPSK, 2PSK, MASK, 2ASK and MQAM.

This study is a continuation of the works mentioned above on AMR using Machine Learning algorithms. His objective is to contribute to the development and implementation of a reliable and robust classification algorithms for both digital and analog modulations.

2. Approach Description

The aim of this paper is to design and implement an automatic signal classifier according to the type of modulation. This includes a system compatible with the communication intelligence system where the receiver is able to automatically detect the modulation pattern of the signal it receives using Modulation Recognition Algorithms (MRA), without having any prior knowledge of the transmitted signal. The purpose of automating modulation recognition is to analyze signals more quickly; to extract information that can be used by the algorithm to distinguish the type of modulation. The approach used in this study has two main steps, the extraction of characteristics of the received or intercepted signal representing the input element of the classifier and the classification process, involving a set of data for training, validation and testing of the classifier.

2.1. Features Extraction for AMC

The recognition system is based on pattern recognition, so it requires suitable features for the process. Regarding feature extraction, Nandi and Azzouz use instantaneous features for the classification of both analog and digital signals in [21], which is the most representative work in the field of AMC. In this work nine

features are taken in consideration and which are therefore derived from the instantaneous phase $\varphi(t)$; the instantaneous amplitude $a(t)$ and the instantaneous frequency $f(t)$ of the signal. They are divided in two categories; the spectral characteristics and statistical characteristics. Most of which were proposed by Nandi and Azzouz in [21], are described as follows:

Maximum power spectral density value (γ_{\max}),

$$\gamma_{\max} = \frac{\max |DTF(A_{cn})|^2}{N} \tag{1}$$

$$A_{cn}[n] = A_n[n] - 1 \tag{2}$$

where:

N is the number of samples and $A_{cn}[n]$, the value of the instantaneous centered amplitude normalized to time $t = n/f_c$, defined by:

$$A_n[n] = \frac{A[n]}{\mu_A} \tag{3}$$

μ_A is the average value of the instantaneous amplitude evaluated on a segment.

$$\mu_A = \frac{1}{N} \sum_{n=1}^N a[n] \tag{4}$$

knowing that: $A[n] = a[n]$,

Standard deviation of the absolute value of the instantaneous phase (σ_{ap})

$$\sigma_{ap} = \sqrt{\frac{1}{N_c} \left(\sum_{A_n[n] > A_t} \Phi_{NL}^2[n] \right) - \left(\frac{1}{N_c} \sum_{A_n[n] > A_t} \Phi_{NL}[n] \right)^2} \tag{5}$$

where:

$\Phi_{NL}[n]$: is the component of the instantaneous phase of the n -th sample at the time instants $t = n/f_c$ ($n = 1, 2, 3, \dots, N$);

N_c : is the number of samples meeting the condition: $A_n[n] > A_t$;

$A_n[n]$ is the normal instantaneous amplitude; A_t the threshold value for $A[n]$ below which the phase evaluation is very sensitive to noise. The variable A_t is a threshold value that filters out low amplitude signal samples because of their high sensitivity to noise.

Standard deviation of the instantaneous direct phase value (σ_{dp})

$$\sigma_{dp} = \sqrt{\frac{1}{N_c} \left(\sum_{A_n[n] > A_t} \Phi_{NL}^2[n] \right) - \left(\frac{1}{N_c} \sum_{A_n[n] > A_t} \Phi_{NL}[n] \right)^2} \tag{6}$$

Spectral symmetry (P)

This feature extraction is based on the spectral powers for the lower and upper sidebands of the modulated signal. The characteristic key is defined as:

$$P = \frac{P_L - P_U}{P_L + P_U} \tag{7}$$

With

$$P_L = \sum_{n=1}^{f_{cn}} |X_c[n]|^2, \tag{8}$$

$$P_U = \sum_{n=1}^{f_{cn}} |X_c[n + f_{cn} + 1]|^2 \tag{9}$$

P_U : upper sideband spectral power;
 P_L : lower sideband spectral power;
 $X_c(n)$: is the Fast Fourier Transform of the modulated signal;
 $(f_{cn} + 1)$ is the number of samples corresponding to the carrier frequency f_c and the sampling rate f_s is defined as:

$$f_{cn} = \frac{f_c N}{f_s} - 1 \tag{10}$$

Standard deviation of the absolute value of the centred-normalized instantaneous amplitude (σ_{aa})

$$\sigma_{aa} = \sqrt{\frac{1}{N} \left(\sum_{n=1}^N A_{cn}^2[n] \right) - \left(\frac{1}{N} \sum_{n=1}^N |A_{cn}[n]| \right)^2} \tag{11}$$

Standard deviation of the absolute value centred-normalized of the instantaneous frequency (σ_{af})

$$\sigma_{af} = \sqrt{\frac{1}{N_c} \left(\sum_{A_n[n] > A_t} f_{NL}^2[n] \right) - \left(\frac{1}{N_c} \sum_{A_n[n] > A_t} |f_N[n]| \right)^2} \tag{12}$$

Standard deviation of the value normalized-centered instantaneous amplitude (σ_a)

$$\sigma_a = \sqrt{\frac{1}{N_c} \left(\sum_{A_n[n] > A_t} a_{cn}^2[n] \right) - \left(\frac{1}{N_c} \sum_{A_n[n] > A_t} a_{cn}[n] \right)^2} \tag{13}$$

2.1.1. Statistical Characteristics

Two statistical parameters are used in this study, the kurtosis of the normalized and centralized amplitude and the kurtosis of the normalized and centralized frequency.

Kurtosis of the normalized-centered instantaneous amplitude (μ_{42}^a) is defined

$$\mu_{42}^a = \frac{E\{A_{cn}^4[n]\}}{\{E\{A_{cn}^2[n]\}\}^2} \tag{14}$$

μ_{42}^f : kurtosis of the normalized-centered instantaneous frequency

$$\mu_{42}^f = \frac{E\{f_N^4[n]\}}{\{E\{f_N^2[n]\}\}^2} \tag{15}$$

2.1.2. Process of Feature Extraction

The process of feature extraction is done by using some transformations such as the Hilbert transform, the forward and inverse Fourier transforms. The filtering

of the signal before the feature extraction is done by means of a convolution. The alternative, filtering by multiplication in the Fourier domain, would imply that the phase and magnitude of the output of the IFFT (Inverse Fast Fourier Transform) must be found. The magnitude found here is equivalent to the filtered input of the FFT (Fast Fourier Transform).

2.2. Classification Algorithm Analysis

In this work, two classification algorithms were used: the KNN and ANN. The K-Nearest Neighbor (KNN) is known as a simple but robust classifier and is capable of producing high performance results even for complex applications [22]. In the KNN algorithm, the training phase corresponds to a knowledge induction of several modulation samples, stored in a database called training data.

For the test and validation phase, when a new signal from the test sample is presented, it is compared with all signals from the training phase. It seems logical that the smaller the distance between the features of two modulations, the more similar they are. Thus, the recognition result is the modulation from the training base that is closest to the newly presented modulation. The comparison is done by calculating the Euclidean distance between the test modulation (vector) and the modulations (vectors) from the training base using the Equation 16 below.

$$D(P_x, P_t) = \sqrt{\sum_{j=1}^L [P_x(j) - P_t(j)]^2} \quad (16)$$

The constant j varies from 0 to 9, which corresponds to the positions of the nine features of each vector.

The artificial neural network (ANN) used for this classification problem is a Multi-Layer Perceptron (MLP). The MLP algorithm is a computational system inspired by biological neural networks that use a network of functions to understand and transform a wide range of input data into the desired output [23]. ANN minimizes errors for nonlinear inputs and can obtain relationships between inputs and outputs without complex mathematical equations.

3. Materials and Methods

In this section, we outline the tools, datasets, machine learning model architectures, and hyperparameters used in our investigation of analog and digital modulation recognition.

3.1. Tool

Experiments and simulations of different algorithms were carried out using MATLAB software version R2019a. In addition to the basic part of the software, several tool boxes were used such as, Neural Network Toolbox (nn tool), Neural Pattern Recognition Toolbox (npr tool), nntrainTool, Signal Processing Toolbox.

3.2. AMC_Master Dataset

The AMC master dataset, introduced by J. Visser and A. Farquharson [24] was

used for this study. The distribution of modulations in the AMC_Master learning/assessment/test set is as follows: 70% of the dataset corresponds to learning data, 15% to test data and the remaining 15% to system assessment. The sample distribution is depicted in **Table 2**.

Table 2. Distribution of modulations in the AMC_Master evaluation set.

<i>Modulations</i>	<i>Class</i>	<i>Number</i>	<i>Percentage</i>
AM_DSB_FC	0	3602	6.1781
AM_DSB_SC	1	3733	6.4028
AM_USB	2	4111	7.0512
AM_LSB	3	4785	8.2072
FM	4	7294	12.5107
MPSK	5	5554	9.5262
2PSK	6	7159	12.2779
MASK	7	7284	12.49355
2ASK	8	7374	12.6479
16QAM	9	7406	12.7028
Total	10	58,302	100

3.3. Model Architectures and Hyperparameter Tuning

We explore K-Nearest Neighbors (KNN) and Artificial Neural Networks (ANN) to determine the most effective approach for our task. For KNN, we optimize the number of neighbors and distance metrics to enhance classification accuracy. For ANN, we adjust hyperparameters such as the number of hidden layers, neurons per layer, learning rate, and batch size. Through systematic tuning and cross-validation, we aim to achieve a model configuration that provides both high accuracy and computational efficiency. **Table 3** provides a concise overview of key parameters for both KNN and ANN models, along with example values for each.

Table 3. Summary of model architectures and hyperparameters for KNN and ANN.

Model	Parameter	Description	Values
KNN	Number of Neighbors	The number of neighbors to consider for classification	1, 2, 3, 4, 5
	Distance Metric	The distance metric used to calculate similarity	Euclidean
ANN	Neurons per Layer	The number of neurons in hidden layer	50, 60, 70, 120, 160
	Activation Function	The function used to introduce non-linearity	ReLU, Softmax
	Learning Rate	The rate at which the model learns during training	0.001

Continued

Batch size/Data division	The number of samples processed before updating the model	Random
Epochs	The number of times the entire dataset is passed through the network	1000

For KNN, the number of neighbors is tested from 1 to 5, using Euclidean distance. For the ANN, neurons per layer range from 50 to 160, with activation functions including ReLU and Softmax, and a fixed learning rate of 0.001 over 1000 epochs.

3.4. Performance Metrics

The performance of the ANN and KNN models was evaluated using ROC (Receiver Operating Characteristic), Accuracy, and Precision [25]-[27]. ROC assesses the model's ability to distinguish between classes, providing insights into the trade-off between true positive and false positive rates [28]. Accuracy measures the overall correctness of the model, indicating the proportion of correctly classified instances. Precision focuses on the model's ability to correctly identify positive instances, reflecting its reliability in positive class predictions [29].

To complete our investigation, a number of signals are generated and simulated at different Signal Noise Ratio (SNR) level values. Secondly, the algorithm for extracting key characteristics is represented as vectors corresponding to inputs of KNN and ANN classifiers. Each model is first created and then trained on a set of training data to be finally tested and used with noisy signal segments. Through repeated experiments, we have observed how classifier performance scales with variances of different approaches and even validated the robustness of key characteristics with respect to noise.

4. Results and Discussions

In this section, the performance of both KNN and ANN algorithms is evaluated by presenting the confusion matrix with corresponding plots and analyzing the extracted features from the signal. Finally, the implementation complexity and processing speed of these methods are compared.

Figure 1 displays the confusion matrix, also known as the percentage matrix, for the KNN model. From the confusion matrix, it can be seen that more than 7 classes can be classified with more than 50% of accuracy. The best result gives an accuracy of 73.5% and a precision of 73.17%. By varying the number of neighbors, the result shows that the more the number of neighbors increases, the weaker the performance of the used model: the precision (percentage of good overall detection) and accuracy (percentage of detection for a specific class) decrease, as depicted in **Table 4**. It shows that the best recognition rate obtained with KNN is close to 74%.

54.79	4.38	13.98	11.4	4.19	0.46	1.86	1.86	2.05	4.93
3.79	28.8	9.94	8.68	12.93	0.18	9.94	3.88	9.22	12.56
14.71	6.79	28.5	12.9	8.16	0.72	8.0	6.46	5.49	8.16
8.46	6.45	11.4	39.9	5.38	3.58	6.95	5.52	4.44	7.89
0.45	3.15	1.55	1.64	85.2	0.27	2.64	0.68	1.64	2.73
0.36	0.12	0.66	2.55	0.12	95.5	0.12	0.12	0.24	0.12
0.18	2.32	0.91	1.59	2.87	0	84.5	2.10	2.92	2.51
0.36	0.67	0.98	1.93	1.43	0	2.60	87.0	2.51	2.46
0.27	1.57	1.07	1.39	1.52	0.13	2.92	3.05	86.2	1.75
1.04	2.12	2.35	1.99	2.85	0.09	2.26	1.67	2.35	83.23

Figure 1. Confusion matrix of the KNN model.

Table 4. Variations in the performance of the KNN model as a function of K.

K	1	2	3	4	5
Precision	73.98	60.21	52.01	47.13	45.25
Accuracy	73.17	59.14	51.4	47.18	45.44

Figure 2 displays the confusion matrix of the ANN algorithm. For optimizing the number of neurons in the hidden layer, a neural network with 9 input layer neurons, 60 hidden layer neurons and 10 output layer neurons was used. The training time was 3 minutes 41 seconds.

Output Class	1	2	3	4	5	6	7	8	9	10		
1	3547 6.1%	41 0.1%	30 0.1%	36 0.1%	0 0.0%	0 0.0%	0 0.0%	11 0.0%	10 0.0%	0 0.0%	96.5%	3.5%
2	4 0.0%	3474 6.0%	132 0.2%	154 0.3%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	92.3%	7.7%
3	27 0.0%	113 0.2%	3905 6.7%	77 0.1%	0 0.0%	0 0.0%	0 0.0%	3 0.0%	0 0.0%	0 0.0%	94.7%	5.3%
4	16 0.0%	103 0.2%	40 0.1%	4511 7.7%	0 0.0%	0 0.0%	0 0.0%	2 0.0%	9 0.0%	0 0.0%	96.4%	3.6%
5	0 0.0%	0 0.0%	0 0.0%	0 0.0%	7198 12.3%	0 0.0%	205 0.4%	0 0.0%	0 0.0%	0 0.0%	97.2%	2.8%
6	1 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5554 9.5%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100.0%	0.0%
7	1 0.0%	0 0.0%	0 0.0%	0 0.0%	96 0.2%	0 0.0%	6945 11.9%	0 0.0%	0 0.0%	16 0.0%	98.4%	1.6%
8	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	5701 9.8%	764 1.3%	1223 2.1%	74.2%	25.8%
9	6 0.0%	2 0.0%	4 0.0%	7 0.0%	0 0.0%	0 0.0%	6 0.0%	594 1.0%	6167 10.6%	391 0.7%	85.9%	14.1%
10	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	3 0.0%	973 1.7%	424 0.7%	5776 9.9%	80.5%	19.5%
	98.5%	93.1%	95.0%	94.3%	98.7%	100%	97.0%	78.3%	83.6%	78.0%	90.5%	9.5%
	1.5%	6.9%	5.0%	5.7%	1.3%	0.0%	3.0%	21.7%	16.4%	22.0%		
	1	2	3	4	5	6	7	8	9	10		

Figure 2. Confusion matrix of the ANN model.

After several experiments related to the variation in the number of neurons, we see in **Table 5** that the best rate is 60 neurons with 90.5%.

Table 5. Variations in the performance of the ANN model as a function of the number of N neurons.

N	15	50	60	70	120	160
Precision	89.3	90.2	90.5	89.6	88.7	83.2

The best recognition rate obtained with KNN is 74%. The learning time has always been long, but it is short compared to the neural method alone. On the other hand, the best recognition rate goes to ANN with 90.5%. Therefore, ANN is a preferred method to derive the modulation type from the KNN method.

Classification by the K-nearest neighbor algorithm (KNN) is a simple and efficient approach. However, improvement in classifier performance is always attractive. The learning time is sometimes long, but it is always short compared to the neural method alone. Using Neural Networks has made the recognition process easier and more reliable. Combining multiple classifiers is an effective technique for improving accuracy.

4.1. Using the Model Obtained

As described above, in order to exploit the power of the machine learning operations, some modulated signals were generated by adding white Gaussian noise. Using the parameter extraction algorithm, we have implemented the features representing the input elements of the classifiers. Once the vector of 9 features is obtained, it is injected into the classifiers and the results are checked.

4.1.1. Generation of Modulated Signals

The parameters below can be used to generate an analogue signal:

Fm: Frequency of the modulating signal;

Am: Amplitude of the modulating signal;

Fp: Carrier frequency;

Ap: Amplitude of the carrier;

Fe: Sampling frequency considering the values:

$F_m = 1$; $F_c = 20$; $F_p = 200$; $A_m = 15$; $A_p = 10$; $F_e = 4 * F_p * F_m$;

Figure 3 shows the modulated and noisy FM signal with SNR = 10 dB. Similarly for the single sideband modulated signal as shown in **Figure 4**. QAM modulated signal 3 with $M = 4$ and SNR = 10 dB is displayed in **Figure 5**.

4.1.2. Extraction of Features

Once the signals have been generated, the characteristic parameters are extracted. **Table 6** depicted the extracting features from signals 1, 2 and 3.

From these data we can see that for signal 1 FM, the amplitude variations σ_a and σ_{aa} are zero, while for signal 3 the first element indicates a strong amplitude compactness. Signal 2 AM_LSB is better positioned here, as the value of γ_{\max} indicates a strong amplitude membership.

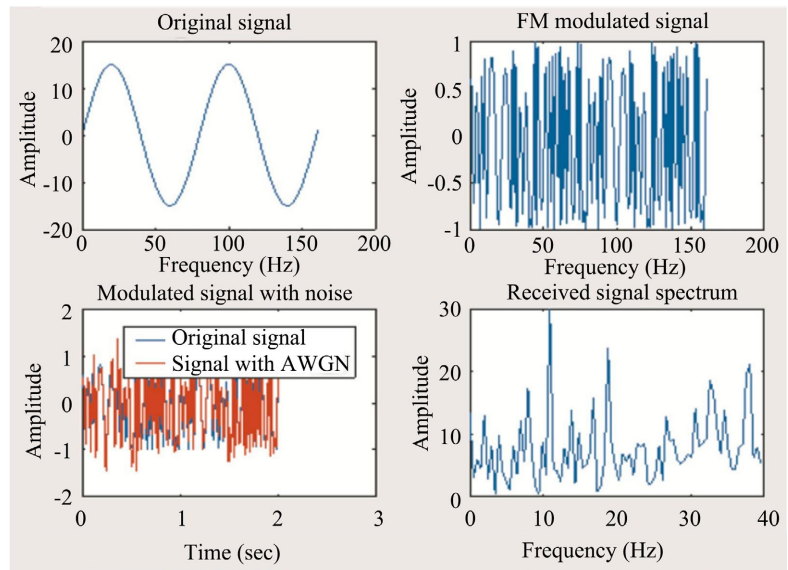


Figure 3. Signal 1 FM modulated signal with SNR = 10 dB.

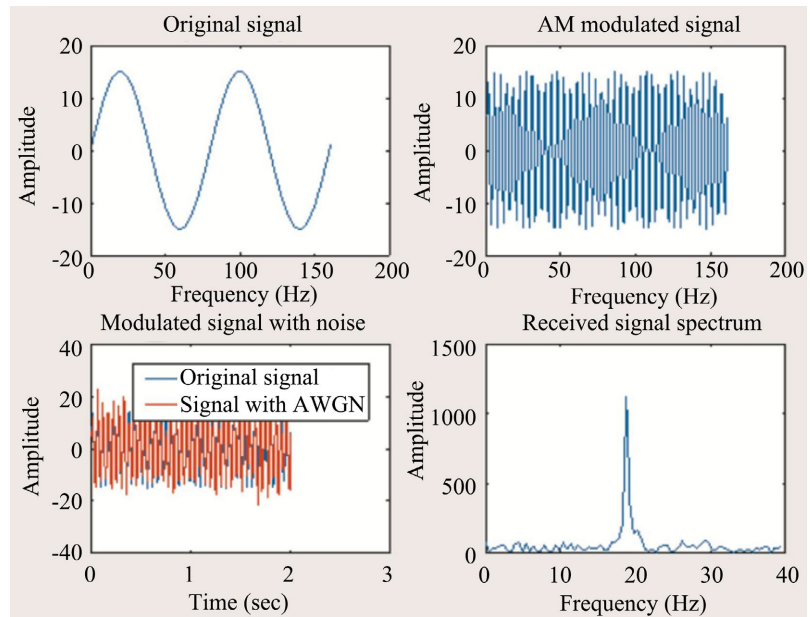


Figure 4. Signal 2 modulated with AM_LSB with SNR = 10 dB.

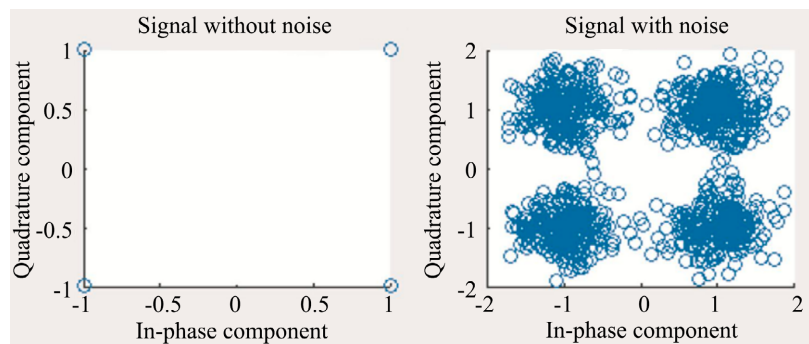


Figure 5. QAM modulated signal 3 with $M = 4$ and SNR = 10 dB.

Table 6. Result of extracting features from signals 1, 2 and 3.

Features	Signal 1	Signal 2	Signal 3
μ_{42}^a	1.94	1.19	1456.97
σ_{af}	$9.7e^{-11}$	$6.7e^{-10}$	0.043
σ_{dp}	3.97	4.29	4.01
σ_{ap}	6.5	7.08	6.656
γ_{max}	508.9	$4.06e^5$	3605.9
P	-0.9	-0.99	-0.998
σ_a	0	0.16	21.80
μ_{42}^f	1.74	1.83	2.403
σ_{aa}	0	0.11	21.808

4.1.3. Classification

The parameters thus extracted are injected into various classifiers and the results obtained are shown in **Figures 6-8** below.

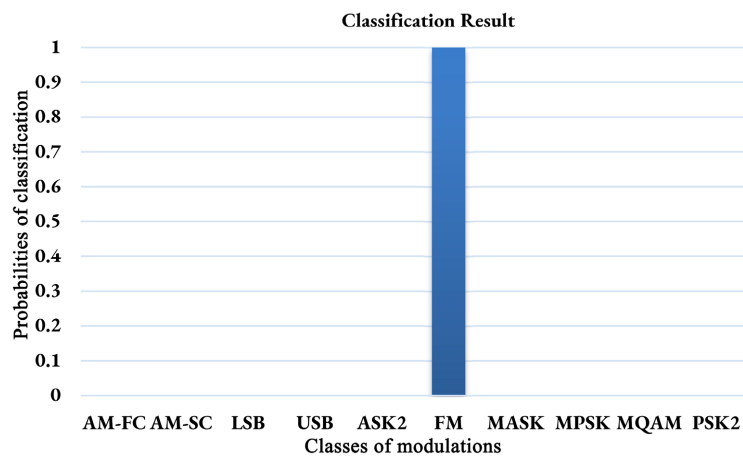


Figure 6. Classification result for signal 1 KNN.

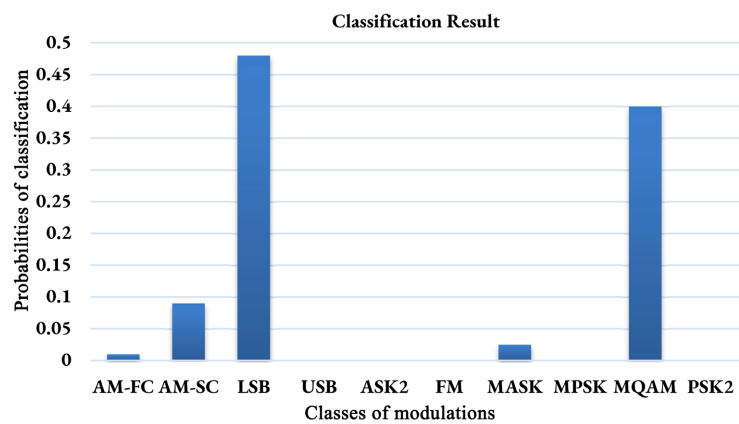


Figure 7. Classification result for signal 2 ANN.

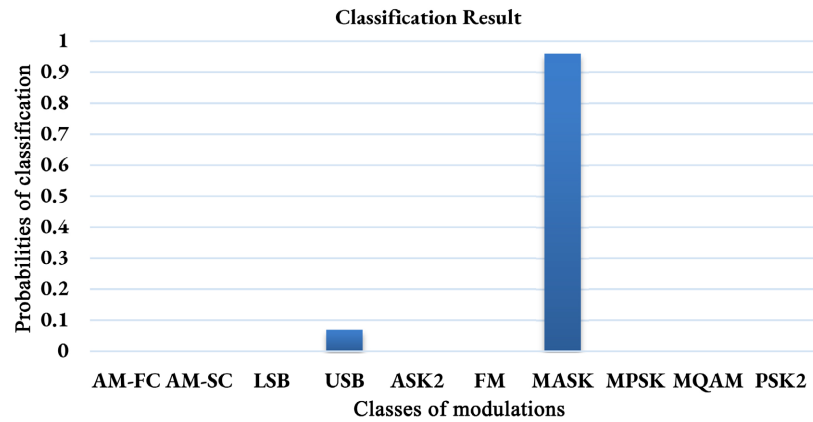


Figure 8. Classification result for the 3 ANN signal.

From these results, we can observe that the 1 FM signal was very well classified at 100%, but note that the results obtained with the KNN classifier are bivalent 0 or 1, so the probability of error is very high.

Signal 2 LSB is classified as such with a dominant percentage of more than 42%.

Signal 3 QAM on the other hand is misclassified as ASK. Going back to the characteristics, it can be said that this result is predictable as its γ_{\max} element is very dominant.

4.1.4. Impact of SNR Conditions on Model Robustness

The current study's findings indicate that added noise levels make modulation detection more challenging, which impacts model performance. This observation aligns with previous research that underscores the detrimental effects of noise on modulation recognition. For instance, Zhou demonstrated that noise significantly degrades the accuracy of modulation classifiers, highlighting the necessity for robust algorithms capable of handling noisy environments [30]. Liu *et al.* [31] showed that the recognition rate is not ideal when the signal-to-noise ratio is low. Similarly, Hu emphasized that effective modulation detection requires strategies to mitigate the impact of noise and interference [32]. Jader [33] examined how various machine learning algorithms, including Decision Trees, Random Forests, SVM, and KNN, perform in recognizing four different modulation types (PSK, QPSK, AM, Morse) under varying Signal-to-Noise Ratios (SNRs) ranging from -10 dB to $+25$ dB. They demonstrated that as the SNR increases, the performance of machine learning models in signal modulation recognition generally improves. This trend highlights the models' ability to better distinguish signal features when the noise is reduced. However, in low SNR conditions, where noise dominates, model performance declines, particularly for those models less robust to noise, such as Random Forest. When the SNR is lower than 10 dB, particularly around 5 dB, model performance tends to stabilize, showing little improvement as the noise remains dominant.

4.1.5. Comparison of Models Obtained

To evaluate the performance of the proposed system, a training base of 58,200

signal samples were used, each characterized by nine attributes with two training techniques: KNN and ANN. The hold-out method is performed during the process of evaluating the machine learning model, where the training database corresponds to 70% of the global dataset.

In this work, KNN model is trained using the same data under different K values, and the K value giving the best performance with the validation data was retained. The best recognition rate obtained with KNN is **74%**. The learning time was still long, but short compared with the neural method alone. On the other hand, the best recognition rate was achieved by ANN with **90.5%**. ANN is therefore a reliable method for predicting the type of modulation compared with the KNN method. The analysis of the performance of each approach is depicted in **Table 7**.

Table 7. Analysis of the performance of the different approaches/the best recognition results obtained.

Models	Settings/ Features	Recognition rate (%)	Learning time	Testing time	Complexity
KNN	9	74.7	Long (seconds)	Very short (a few m.seconds)	Low
ANN	9	90.5	Very long (minutes/hours)	Very short (a few m.seconds)	High

Classification using the K-nearest Neighbor (KNN) algorithm is a simple and effective approach. But the improvement in classifier performance is still attractive. The learning time is sometimes long, but it is always short compared with the neural method alone. The use of Neural Networks has made the recognition process easier and more reliable. Combining several classifiers is an effective technique for improving accuracy.

5. Conclusion

The contribution of this work is on the development of analog and digital based on ANNs and KNNs, which can automatically recognize all forms of modulation schemes. Using the AMC_Master dataset, we chose a characteristic-based automatic modulation classification scheme for ten analogue and digital modulation signals of; AM_DSB_FC, AM_DSB_SC, AM_USB, AM_LSB, FM, MPSK, 2PSK, MASK, 2ASK and MQAM. From a statistical and spectral analysis of signals, nine key differentiation characteristics are extracted and used as input vectors for each model formed. Simulation results including the recognition rate of 90.5% from the ANN method with a hidden layer (sixteen neurons) highlight this algorithm effectiveness in accurately predicting the modulation type for testing the dataset. These results indicate that the proposed ANN algorithm has successfully captured the distinguishing features and patterns of different modulation schemes, enabling it to make precise predictions. In order to improve this result, future work

should explore more advanced techniques, such as convolutional neural networks and reinforcement learning, as potential methods to address the current study's limitations and improve model performance across a wider range of real-world scenarios of modulations. Our next step is to develop and evaluate methods that enhance recognition accuracy under high noise conditions to improve overall robustness. Also, various channel types, including Rayleigh and Rician fading, can be investigated in the future to extend the work.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Zhu, Z. and Nandi, A.K. (2014) Automatic Modulation Classification: Principles, Algorithms and Applications. Wiley. <https://doi.org/10.1002/9781118906507>
- [2] Dobre, O.A. and Inkol, R. (2012) Blind Signal Identification: Achievements, Trends, and Challenges. 2012 9th International Conference on Communications (COMM), Bucharest, 21-23 June 2012, 349-352. <https://doi.org/10.1109/icomm.2012.6262564>
- [3] Nandi, A.K. and Azzouz, E.E. (1998) Algorithms for Automatic Modulation Recognition of Communication Signals. *IEEE Transactions on Communications*, **46**, 431-436. <https://doi.org/10.1109/26.664294>
- [4] Wong, M.L.D. and Nandi, A.K. (2008) Semi-Blind Algorithms for Automatic Classification of Digital Modulation Schemes. *Digital Signal Processing*, **18**, 209-227. <https://doi.org/10.1016/j.dsp.2007.02.007>
- [5] Liang, H. and Ho, K.C. (2003) Classification of BPSK and QPSK Signals with Unknown Signal Level Using the Bayes Technique. *Proceedings of the 2003 International Symposium on Circuits and Systems*, Bangkok, 25-28 May 2003, 4. <https://doi.org/10.1109/iscas.2003.1205758>
- [6] Wong, M.L.D. and Nandi, A.K. (2004) Automatic Digital Modulation Recognition Using Artificial Neural Network and Genetic Algorithm. *Signal Processing*, **84**, 351-365. <https://doi.org/10.1016/j.sigpro.2003.10.019>
- [7] Swami, A. and Sadler, B.M. (2000) Hierarchical Digital Modulation Classification Using Cumulants. *IEEE Transactions on Communications*, **48**, 416-429. <https://doi.org/10.1109/26.837045>
- [8] Çoruk, R.B., Yalçinkaya, B., Benzaghta, M. and Kara, A. (2021) On the Classification of Modulation Schemes using Higher Order Statistics and Support Vector Machines.
- [9] Headley, W.C., Reed, J.D. and da Silva, C.R.C.M. (2008) Distributed Cyclic Spectrum Feature-Based Modulation Classification. 2008 *IEEE Wireless Communications and Networking Conference*, Las Vegas, 31 March-3 April 2008, 1200-1204. <https://doi.org/10.1109/wcnc.2008.216>
- [10] Wong, M.L.D., Ting, S.K. and Nandi, A.K. (2008) Naïve Bayes Classification of Adaptive Broadband Wireless Modulation Schemes with Higher Order Cumulants. 2008 *2nd International Conference on Signal Processing and Communication Systems*, Gold Coast, 15-17 December 2008, 1-5. <https://doi.org/10.1109/icspcs.2008.4813755>
- [11] Hassanpour, S., Pezeshk, A.M. and Behnia, F. (2016) Automatic Digital Modulation Recognition Based on Novel Features and Support Vector Machine. 2016 *12th*

International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), Naples, 28 November-1 December 2016, 172-177.

<https://doi.org/10.1109/sitis.2016.35>

- [12] Ansari, S., Alnajjar, K.A., Abdallah, S. and Saad, M. (2020) Automatic Digital Modulation Recognition Based on Machine Learning Algorithms. 2020 *International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI)*, Sharjah, 3-5 November 2020, 1-6. <https://doi.org/10.1109/ccci49893.2020.9256809>
- [13] Jajoo, G., Kumar, Y., Yadav, S.K., Adhikari, B. and Kumar, A. (2017) Blind Signal Modulation Recognition through Clustering Analysis of Constellation Signature. *Expert Systems with Applications*, **90**, 13-22. <https://doi.org/10.1016/j.eswa.2017.07.053>
- [14] Pouyanfar, S., Sadiq, S., Yan, Y., Tian, H., Tao, Y., Reyes, M.P., et al. (2018) A Survey on Deep Learning. *ACM Computing Surveys*, **51**, 1-36. <https://doi.org/10.1145/3234150>
- [15] Daldal, N., Yıldırım, Ö. and Polat, K. (2019) Deep Long Short-Term Memory Networks-Based Automatic Recognition of Six Different Digital Modulation Types under Varying Noise Conditions. *Neural Computing and Applications*, **31**, 1967-1981. <https://doi.org/10.1007/s00521-019-04261-2>
- [16] Park, C., Choi, J., Nah, S., Jang, W. and Kim, D.Y. (2008) Automatic Modulation Recognition of Digital Signals Using Wavelet Features and SVM. 2008 *10th International Conference on Advanced Communication Technology*, Gangwon, 17-20 February 2008, 387-390. <https://doi.org/10.1109/icact.2008.4493784>
- [17] O'Shea, T.J., Corgan, J. and Clancy, T.C. (2016) Convolutional Radio Modulation Recognition Networks. In: Jayne, C. and Iliadis, L., Eds., *Engineering Applications of Neural Networks*, Springer International Publishing, 213-226. https://doi.org/10.1007/978-3-319-44188-7_16
- [18] Jagannath, A. and Jagannath, J. (2021) Dataset for Modulation Classification and Signal Type Classification for Multi-Task and Single Task Learning. *Computer Networks*, **199**, Article ID: 108441. <https://doi.org/10.1016/j.comnet.2021.108441>
- [19] Cheng, X., Zhang, Y., Chen, J. and Hu, J. (2020) Modulation Classification Based on Deep Learning with Real-World Noisy Dataset. *IET Communications*, **14**, 2834-2841.
- [20] Alam, M.M., Islam, M.T. and Hossain, M.M. (2021) A Novel Deep Learning Approach for Automatic Modulation Classification in Additive White Gaussian Noise. *Journal of Communications and Networks*, **23**, 22-32.
- [21] Nandi, A.K. and Azzouz, E.E. (1995) Automatic Analogue Modulation Recognition. *Signal Processing*, **46**, 211-222. [https://doi.org/10.1016/0165-1684\(95\)00083-p](https://doi.org/10.1016/0165-1684(95)00083-p)
- [22] Cover, T. and Hart, P. (1967) Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory*, **13**, 21-27. <https://doi.org/10.1109/tit.1967.1053964>
- [23] Hagan, M.T., Demuth, H.B., Beale, M.H. and De Jesús, O. (2014) *Neural Network Design*. 2nd Edition, Martin Hagan.
- [24] Visser, J. and Farquharson, A. (2015) *Automatic Detection, Classification and Demodulation of Radio Signals*. School of Electrical & Information Engineering, University of the Witwatersrand.
- [25] Fawcett, T. (2006) An Introduction to ROC Analysis. *Pattern Recognition Letters*, **27**, 861-874. <https://doi.org/10.1016/j.patrec.2005.10.010>
- [26] Powers, D.M.W. (2011) Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness, and Correlation. *Journal of Machine Learning Technologies*, **2**, 37-63.
- [27] Davis, J. and Goadrich, M. (2006) The Relationship between Precision-Recall and

- ROC Curves. *Proceedings of the 23rd International Conference on Machine Learning*, Pittsburgh, 25-29 June 2006, 233-240. <https://doi.org/10.1145/1143844.1143874>
- [28] Liu, M., Liu, Z., Lu, W., Chen, Y., Gao, X. and Zhao, N. (2022) Distributed Few-Shot Learning for Intelligent Recognition of Communication Jamming. *IEEE Journal of Selected Topics in Signal Processing*, **16**, 395-405. <https://doi.org/10.1109/jstsp.2021.3137028>
- [29] Powers, D.M.W. (2011) Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation. *Journal of Machine Learning Technologies*, **2**, 37-63.
- [30] Zhou, S., Yin, Z., Wu, Z., Chen, Y., Zhao, N. and Yang, Z. (2019) A Robust Modulation Classification Method Using Convolutional Neural Networks. *EURASIP Journal on Advances in Signal Processing*, **2019**, Article No. 21. <https://doi.org/10.1186/s13634-019-0616-6>
- [31] Li, J., Qi, L. and Lin, Y. (2016) Research on Modulation Identification of Digital Signals Based on Deep Learning. 2016 *IEEE International Conference on Electronic Information and Communication Technology (ICEICT)*, Harbin, 20-22 August 2016, 402-405. <https://doi.org/10.1109/iceict.2016.7879726>
- [32] Hu, S., Pei, Y., Liang, P.P. and Liang, Y. (2020) Deep Neural Network for Robust Modulation Classification under Uncertain Noise Conditions. *IEEE Transactions on Vehicular Technology*, **69**, 564-577. <https://doi.org/10.1109/tvt.2019.2951594>
- [33] Jader, R.F., Abd, M.H.M. and Jumaa, I.H. (2022) Signal Modulation Recognition System Based on Different Signal Noise Rate Using Artificial Intelligent Approach. *Journal of Studies in Science and Engineering*, **2**, 37-49. <https://doi.org/10.53898/josse2022244>