IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received August 15, 2021, accepted September 16, 2021, date of publication September 27, 2021,
date of current version October 13, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3115888

AMC2-Pyramid: Intelligent Pyramidal Feature Engineering and Multi-Distance Decision Making for Automatic Multi-Carrier Modulation Classification

DHAMYAA HUSAM AL-NUAIMI[1,2], NOR ASHIDI MAT ISA[1], MUHAMMAD FIRDAUS AKBAR[1], (Member, IEEE), AND INTAN SORFINA ZAINAL ABIDIN[1]

[1]School of Electrical and Electronic Engineering, Engineering Campus, Universiti Sains Malaysia, Nibong Tebal 14300, Penang, Malaysia
[2]Communication Engineering Department, Al-Mansour University College, Baghdad 10068, Iraq

Corresponding authors: Dhamyaa Husam Al-Nuaimi (dhamyaa.husam@muc.edu.iq) and Nor Ashidi Mat Isa (ashidi@usm.my)

ABSTRACT Automatic modulation classification (AMC) is a method that supported different wireless communication systems for modulation type classification. Currently, orthogonal frequency division multiplexing, multiple-input, multiple-output systems are widely using this technique. Recent AMC methods are designed for a single-carrier system identifying a few modulation types. To motivate the AMC for the current communication systems, we present an intelligent pyramid model for automatic multi-carrier modulation classification (AMC2-pyramid) which alleviates the existing works challenges such as high degradation of accuracy for higher order modulation schemes, inefficient feature extraction and lack of effectiveness in low SNR environments. The proposed work contains three significant operations, namely, signal fortification, feature engineering and modulation classification. First, signal quality is estimated to reduce the complexity in classification because some signals are affected by noise and other environmental or channel artefacts. Hence, before pre-processing the signal, the quality is assessed according to the channel state information, signal to inference plus noise ratio, received signal strength indicator and spectral efficiency. For low quality, quality augmentation is applied. Then, quality augmentation is implemented with noise elimination, equalisation, quantisation and channel frequency offset compensation. In the feature engineering step, feature extraction and clustering are presented using the Gated Feature Response Pyramid Network (GaFP), and a twin-functioned human mental search algorithm is used. The modulation classification is implemented using a multi-distance-based nearest centroid classifier, and improved Q-learning is used to identify signals as any of 16QAM, 32QAM, 64QAM, 128QAM, QPSK, BPSK, DPSK, ASK and FSK. The performance of the proposed AMC2-pyramid is implemented using MatlabR2017b, where accuracy (6.8% - 23.15%) high when compared to sample size and (14% - 46%) high when compared to SNR at −10 dB, precision (4.96% - 29.5%) high when compared to sample size and (16.5% - 48.5%) high when compared to SNR at −10 dB, recall (2- 29.76%) high when compared to sample size and (14% - 45%) high when compared to SNR at −10dB, F-score (2- 30%) high when compared to sample size and (15.5%- 46.5%) high when compared to SNR at −10 dB, error rate (0.7% - 11.5%) low when compared to sample size and (4.5%- 17%) low when compared to SNR at −10 dB, computational time (170ms - 400ms) low when compared to sample size is computed for the proposed work including previous well-known methods. The proposed work proves that this method outperforms the previous ones.

INDEX TERMS Automatic modulation classification, multi-carrier system, signal quality assessment, gated feature pyramid network, distance-based classifier, multi-modulation type detection.

I. INTRODUCTION

Automatic modulation classification (AMC) is an emerging field that has great attention from various signal

The associate editor coordinating the review of this manuscript and approving it for publication was Tianhua Xu.

processing applications. However, the transmitter can choose any modulation type of signal freely, but the receiver must know the modulation type for demodulating it. Thus, the transmission between the transmitter and receiver is successful. The modulation process is essential for the communication system, which is used to transmit the high-frequency carrier wave to a low-frequency signal. In this way, the transmitted signal contains all the information of the original message signal [1]. The communication system needs modulation because the low-frequency signals cannot be transmitted over long distances as such signals need to modulate high-frequency signals. Multi-carrier modulation is used to transmit the data by dividing it into multiple components, which send to the individual carrier signals [2]. This signal has a narrow bandwidth, whereas the multiple signals have a wide bandwidth. The benefits of multi-carrier modulation include relative protection for fading affected by transmission over multiple paths at a particular time, with less susceptibility compared with a single-carrier system. In traditional modulation classification, the signal features are classified manually through time-frequency features, high-order statistics features and others, which are not suitable for all conditions. As the signal characteristics are not captured manually through various modulations, the inappropriate feature selection leads to poor classification accuracy [3]. In addition, the manual modulation is complex and inaccurate when the number of modulation types increased in the environment. These issues are addressed by AMC, which is more effective than manual modulation. AMC is vital in recognising the modulation algorithm used by the transmitter to the receiver. AMC has wide applications in the military and civilian fields [4]. Typically, AMC is performed for single- and multi-carrier modulation techniques with no spectrum efficiency. Moreover, AMC uses two techniques, that is, multiple-input multiple-output (MIMO) and orthogonal frequency division multiplexing (OFDM). The former provides multiple channels with the independent transmission, which is under a definite condition that increases the reliability and limited frequency resources of the systems. Then, the latter is used to reduce inter-carrier interference (ICI) and inter-symbol inference (ISI) because of the use of multi-carrier modulation in the frequency-selective channel. These techniques are used to achieve a high data rate and reliability in AMC. The digital modulation recognition schemes often come under the feature-based technique [5] or likelihood ratio [6]. The maximum likelihood technique provides optimal solutions in AMC. The threshold-based classification scheme is presented under an AMC architecture [7].

Figure. 1 depicts the process of AMC, which includes two blocks, namely, a transmitter and receiver. In the transmitter, the input data are collected and send to the automated modulator and then fed into a conversion. Then, the data are transmitted to the receiver through the additive white Gaussian noise (AWGN) channel. The modulated data are demodulated using a demodulator. Finally, the modulation type is classified and sends into the automatic modulator for feedbacks.

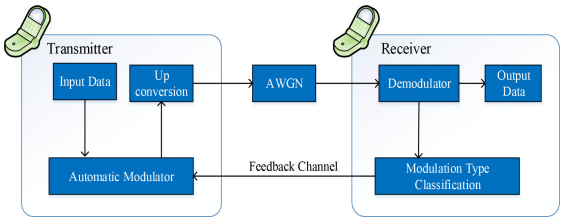

FIGURE 1. Automatic modulation classification.

Initially, machine learning (ML) algorithms are used for feature engineering in AMC, which needs essential expert experience. However, during complex functions learning, ML is highly complex. Hence, a deep learning (DL) algorithm is introduced in AMC. The main AMC works focus on the use of growing DL approaches [7]. Many works developed CNN for feature extraction and classification [8]. In CNN, the extracted features may belong to statistical, spectral, transform-based and others as CNN has learned the features for modulation classification, which helps in addressing the dimensionality issues [8]–[10]. Furthermore, the decision may be made through direct voting, weighty voting, direct averaging or weighty averaging. The CNN structure can be modified (e.g. CNN is combined with a multi-stream network) [11], [12]. Each approach is dedicated to different works, and the final classification is performed by CNN. However, the CNN has a high signal loss in the max-pooling layer, which reduces the accuracy of AMC [13], [14]. However, the communication channel may be corrupted by additive noise, which cannot be handled by many AMC works, and the low-order features are not suitable for classifying higher-order modulation schemes [15]–[17]. Table 1 presents a comparison of single- and multi-carrier modulations parameters. For instance, single-carrier modulation has no subcarriers, whereas multi-carrier modulation has 2048 subcarriers.

TABLE 1. Comparison between single- and multi-carrier modulations.

Parameter	Single-carrier modulation	Multi-carrier modulation
Frequency selectivity	Yes, $10\mu s > \frac{1\mu s}{10}$	No, $10\mu s > \frac{128\mu s}{10}$
Data rate	2.5 Mbps	2.5 Mbps
Modulation	16-ary QAM	16-ary QAM
Sampling frequency	5 MHz	5 MHz
No of subcarriers	–	2048
CP length	3	3
FFT size	2048	2048

A. RESEARCH MOTIVATION

In recent times, AMC has received attention from researchers. Many research works presented single-carrier modulation systems. However, the growing communication systems, such as LTE, use OFDM, which is a multi-carrier system [18]. That is, AMC for a multi-carrier system is an emerging topic. Most of the research works focused on multi-carrier AMC in the presence of white noise or no noise. However, in practice, the channel may be corrupted by many more issues. All AMC works have been developed and tested in the ideal channel condition, which is not possible in real-time applications. With the low-quality modulated signal, the classification accuracy is too low even for single-carrier AMC. Multi-carrier OFDM, which is an emerging candidate for future generation networks (like 5G), receives limited attention. Considering limited features often decreases accuracy, and calculating these features in each signal relatively increases time consumption. The AMC must handle all types of modulation schemes. In classification, time consumption is high because of sequential feature learning and predictions. Motivated by these issues, we design a novel AMC system for multi-carrier AMC for a more realistic environment [19]. Hence, we designed some of the research objectives as follows:

- To develop AMC for multi-carrier OFDM systems in a non-ideal channel environment
- To perform feature engineering to support multiple modulation schemes with high discrimination
- To achieve high classification accuracy through the feedback-based decision making

B. CONTRIBUTIONS

To meet the issues in previous works of single carrier modulation classification, in this study we focus on the multi carrier modulation classification scheme using deep learning algorithms which can identify many signals and our contributions are as follows,

- We propose an automatic multi-carrier modulation classification (AMC^2) model for modulation classification using an intelligent pyramid model, which is shortly defined as AMC^2-pyramid. In this model, three operations of work are deployed, namely, signal fortification, feature engineering and modulation classification.
- Firstly, the (AMC^2) model enhances the signal quality, and such process is known as a fortification, which includes two processes, that is, quality evaluation and quality augmentation, using the Bi-Fold signal fortification Bfsf) approach. The fairness score is calculated for quality enhancement. If the score is low, then quality augmentation is performed by considering noise removal, equalisation, quantisation and channel frequency offset (CFO) compensation, to improve the AMC accuracy.
- Secondly, we propose the gated feature pyramid network (GaFP-Net) for feature extraction that extracts all features (spectral, statistical, transformation and constellation domains) from the signal. Through this way of effective feature extraction, the proposed AMC method achieves high classification accuracy.
- Thirdly, we clustered the extracted features using an intelligent twin-functioned human mental search (TF-HMS) optimiser to improve classifier accuracy and also minimise classification complexity. TF uses cluster purity and cluster entropy functions to update the clusters.
- Finally, we propose the multi-distance-based nearest centroid classifier ($MdNC^2$) algorithm, and improved Q-learning (IQL) is presented to classify the exact modulation type for the received signal. Furthermore, GaFP-Net outperforms CNN, and the proposed decision-making by Q-learning increases the accuracy of the overall system.

The performance of the proposed work is compared in terms of accuracy, precision, recall, f-score, error rate and computational time with respect to two different scenarios, namely, number of samples and SNR variations. The analysis indicates that the proposed work has reached a peak performance than the existing methods.

The rest of this paper is structured as follows. Section II describes the literature review of AMC for single-carrier and multi-carrier systems. Moreover, this section further discusses the classification methods, modulation types identified and in detail. Section VI is dedicated to the investigation of deficiencies for each work. Section III summarises the significant problems with the multi-carrier-based AMC methods. Section IV discusses the signal model for multi-carrier systems. Section V shows the proposed AMC^2-pyramid model the performance analysis of the proposed and previous works in terms of numerical evaluation. Section VII concludes the paper and discusses future remarks.

II. RELATED WORK

This section reviews the most significant works established for AMC and includes two subsections, namely, AMC with DL and AMC without DL. Research gaps are also determined.

A. AMC WITHOUT USING DL

The Kalman filter is integrated with adaptive interacting multiple models (IMM) [20] to estimate the channel state information (CSI). The channel is decomposed using singular-vector decomposition (SVD) by adding up square root singular values. For AMC classification, a quasi-likelihood ratio test and expectation-maximization (EM) algorithm is proposed. All proposed algorithms are working together for CSI estimation and AMC. However, this work increases complexity, and the accuracy is also low because of the presence of noise. AMC is presented in [21] on 5G unmanned aerial vehicle systems. This work aims at AMC for the signals without prior knowledge of multipath channels. This work uses mean, variance, skewness and kurtosis of wavelet transform as optimal features. To select the optimal

feature subset, principal component analysis (PCA) is proposed. In the end, a neural network classifier is employed to identify the modulation scheme. Before feature extraction, the signals are pre-classified based on the threshold value. The lack of important features and pre-processing minimises the classification accuracy. To improve accuracy, this research proposed the K-nearest neighbor (KNN) algorithm and genetic programming (GP) [22]. This work uses eighth-order cumulants to classify the M-QAM and M-PSK modulation schemes. From the cumulant features, a single-stage super-feature set is generated by using GP and then classified by the KNN algorithm. The overall work is tested under AWGN channel conditions. In KNN, the Euclidean distance measure is employed for M-PSK and M-QAM classification. Considering instantaneous features alone is not suitable to achieve better accuracy. To improve classification accuracy, this research completely focuses on the M-QAM signal classification [23]. In particular, this work aims to classify the higher-order QAM techniques, such as 128-QAM, 256-QAM, 512-QAM and 1024-QAM. To achieve this, this work extracts the higher-order HOC features from the input data. The entire work is tested under the Gaussian noise channel. The extracted features and the signal are fed into a logarithmic classifier to determine the modulation scheme. The Gaussian noise is a common noise that frequently occurs. In addition, the received signal may be corrupted by various factors. However, this work is only suitable for the AWGN noise (i.e. not suitable to handle other noises).

Reference [24] uses the fourth-order cumulants for modulation classification. In this work, ASK, PSK and QAM signals are analysed and classified. Firstly, the constellation characteristics of each signal are analysed to reduce the distance between two constellation points (i.e. constellation size). This work considers the weighted fractional Fourier transform (WF-RFT) system. From the first-order derivatives of fourth-order cumulants, the optimal WFRFT order is determined. Then, classification is performed. This work is only limited to PSK, ASK and QAM schemes. The lower-order cumulant is not suitable for accurate classification. However, to increase classification accuracy, authors [25] presented a revised architecture for AMC. This work considers analogue and digital modulation modes. The analogue modulation, such as amplitude modulation, frequency modulation, phase modulation, single-sideband modulation and double-side band-suppressed carrier techniques, are considered. Similarly, digital modulation techniques, such as ASK, FSK, QAM and VSB, are considered. For classification, instantaneous amplitude and phase features are extracted and analysed. For each modulation scheme, a separate threshold value is computed. Upon the threshold value, the modulation scheme is classified. However, static threshold-based classification is not suitable to handle real-time noisy data because it has large variations. However, the noises are removed by the cepstrum-based pre-processing algorithm, which eliminates the multipath channel coefficient and noise [26]. A logarithmic functional fitting method is used to classify the modulated signals. The cepstrum is a multipath signal processing methodology that estimates the multipath parameters effectively. The logarithmic functional fitting method classifies the signals by computing Euclidean distance between features and logarithmic function values. Furthermore, this work estimates the SNR value using the SVD method. This work is tested under ideal channel conditions, which is not practical.

Authors [27] jointly focused on radio-frequency identification (RFID), tag recognition and AMC. The hierarchical RFID recognition consists of blind source separation (BSS), graph-based AMC and direct sequence spread spectrum (DSSS). Firstly, the multiplexed signal is separated by BSS, which is initiated in the AMC for modulation identification. In the graph-based AMC, the graph is constructed for all received signals. Then, the features are extracted from the graph. Finally, graph-domain classification is performed for AMC. The initial graph construction is a complex and time-consuming process.

B. AMC WITH USING DL

To overcome the issues in ML, most of the researchers use a convolutional neural network (CNN) for AMC, which is under the DL model that provides high accuracy and also solves the large-scale environment issues [28]–[34]. Author [28] a performed signal analysis by using the DL method. In wireless signal processing, the following processes are considered: signal pre-processing, feature learning, classification and decision-making. In particular, AMC becomes an interesting area of wireless signal classification. The CNN is utilised for modulation classification. In the hidden layers, high-level features are extracted. In recent times, multi-carrier modulation schemes are used in many communication systems. Thus, multi-carrier modulation detection will be a better future research direction. Here, the InceptionResNetV2 is combined with the transfer adaptation [29]. Initially, the received signal is pre-processed, and the constellation diagram is generated. The integrated InceptionResNet2-TA is used to extract the features, and the SVM classifier is used to identify the modulation mode. Here, PSK signals (QPSK, BPSK and 8PSK) are considered for identification. However, SVM is a slow classifier, which takes a huge time to classify the signals by matching the features individually. In addition, the pre-processing step is unable to fully eliminate the noise, which degrades the classification accuracy. However, CNN is also used for cooperative AMC in the MIMO system [30]. The receiver is equipped with multiple antennas, and the modulation classification decision is made cooperatively. Each receiving antenna classifies the signal by using CNN and provides the sub-results to the decision-maker. On receiving results from all antennas, the decision is made through direct voting, weighty voting, direct averaging and weighty averaging methods. Furthermore, various decision-making rules are deployed to accurate identification. In general, CNN has a high loss function,

which degrades performance. The loss functions are reduced by the proposed lightweight AMC (LightAMC), which is proposed by [31] based on the DL approach. Here, the CNN algorithm is used for feature extraction and classification. Firstly, a scaling factor is introduced at each neuron. To provide the scaling factor, the compressive sensing concept is introduced. The CNN parameters are tuned as per the SNR estimated from the signals. The stochastic gradient descent is preferred as the optimiser to train the CNN. Here, the fourth-order cumulants are extracted as the features. Typically, CNN has a high-loss function, which degrades the classification accuracy. The considered features are limited, which are insufficient for AMC. To improve the classification accuracy, this work used combined features of CNN and handcraft [32]. In this work, the signal is converted into the images through smooth pseudo-Wigner–Ville distribution and Born-Jordan distribution. Then, a fine-tuned CNN model extracts the features from the signals. Next, joint features are formed by combining CNN and handcrafted features. To fuse these features, a multimodality fusion model is proposed. The received signal is analysed in both distribution methods, and then, the features are fused for classification. This work has the ability to classify the lower-order modulation schemes but is not suitable for higher-order modulation schemes.

A DL-based multi-stream structure is proposed for modulation classification [33]. The CNN architecture is enriched by introducing the multi-stream structure, which increases the classification performance. The multi-stream network uses the superposition of a small convolution kernel with fewer parameters. The involvement of a multi-stream enriches each feature. The multi-stream structure in the CNN also handles the varying signal lengths. However, in the presence of noise, the classification accuracy is poor. Nevertheless, to increase the classification accuracy, this research designs a cooperative classifier for AMC [34]. For cooperative classification, CNN, recurrent neural network (RNN), generative adversarial network (GAN), cyclic-connected CNN (CCNN) and bidirectional RNN (BRNN) are working cooperatively to detect the modulation scheme. GAN is used to extend the training data. CCNN is used to obtain the spatial features, and BRNN is used to obtain the temporal features. To fuse the features, global average and max pooling are proposed. In addition, the attention mechanism is presented for recalibration. The consideration of multiple deep structures increases the overall complexity of the system.

Synthesis of Related Work: From the prior research work analysis, we have identified that the following research issues still need to be resolved in the AMC: (1) AMC under a noisy environment often leads to lower accuracy, (2) the lack of effective feature extraction progress leads to higher time consumption and complexity, (3) existing solutions are unable to handle the noises other than AWGN, (3) CFO and I/Q imbalance, which affects the classification severely, are less focused, (4) overall, the accurate identification of the modulation scheme is still challenging because of noise, ineffective features and the poor classification algorithm.

III. PROBLEM STATEMENT

The current AMC classification techniques only use a few modulation steps and generally focus on the single-carrier signal-based AMC. However, recent research applications demand multi-carrier modulations (e.g. OFDM). For that, a classifier should be designed to classify a vast number of modulations. In earlier works, the following four steps are frequently used for multi-carrier signal modulation classification:

- Pre-processing
- Feature extraction
- Feature clustering
- Modulation type classification

In the following, the research gaps determined in the previous works for the above steps are given, and the corresponding research solutions taken in the proposed model are highlighted. The feature extraction, clustering and classification have been developed by [35], [36], and [37]. The multi-carrier modulation schemes, such as ASK, FSK and PSK, are classified [35]. For modulation classification, the time and frequency domain features are extracted. The research problems identified in this work are as follows:

- In the presence of noise, the classification accuracy is too low. However, communication channels often have high noise, CFO and I/Q imbalance, which severely affect the performance of the classifier [35]–[39].
- The spectral features (time and frequency domain) work well only in the high SNR scenario, and the discrimination is also poor even for three modulation schemes.
- The Neutrosophic can handle only liner signals that can be suitable for single-carrier AMC. However, Neutrosophic produces overlapped clusters for multi-carrier AMC with multiple features. Thus, the feature processing is ineffectual.
- This work summarises that the RF tree outperforms in the classification process. However, the RF tree consumes higher time for decision tree construction and has large complexity.

Authors [36] proposed the OFDM modulation classification algorithm based on the deep neural network. A blind digital modulation classification method is presented for OFDM systems [37]. The research problems are defined as follows:

- Although HOS features work in a noisy environment, such features fail when dealing with higher-order modulation techniques. Thus, the signals modulated with higher-order modulation techniques will have a large false-positive rate.
- In particular, lower-order HOS and HOC features are extracted and are unable to classify even lower-order modulations schemes. Furthermore, the computation of all features from each signal increases the training and testing time (as the system uses multi-carrier modulation).

The pre-processing-based modulation type is identified by [38] and [39]. An AMC method is proposed for

TABLE 2. Research summary.

Types of AMC	Algorithms Methods	Strengths	Limitations
AMC without using deep learning	SVD [20]	Achieves target optimal performance	High complexity due to presence of noise Less classification accuracy due to lack of pre-processing
	PCA [21]	Better performance during feature selection	Less accuracy due to lack of significant features and pre-processing
	KNN [22]	Low SNR rate for all modulations	Less accuracy due to consider only instantaneous features
	M-QAM model [23]	Reliable recognition rate even at low SNR rate (SNR<0)	Only suitable for remove AWGN noise not suitable for other noise
	WF-RFT [24]	Realistic recognition and high accuracy	Only limited to PSK, ASK and QAM method when lower order cumulant is not suitable
	DAM [25]	Rapid performance during modulation recognition for varying noise levels and modulations	Not suitable for handling real time noisy data due to static threshold during classification
	Cepstrum model [26]	Less computational complexity	Inefficient feature extraction
	DSSS [27]	Rapid performance during Tag recognition	High information loss due to not select the number of PCA
AMC with using deep learning	CNN[28]	Improved performance at low SNR rate (-4dB)	Low performance in feature extraction
	CNN[29]	High performance for modulation recognition with medium for high SNR rate	Less accuracy due to lack of pre-processing
	CNN[30]	High accuracy for digital signal modulation at low SNR (4dB)	Not suitable for large datasets because SVM takes high amount of time for selecting modulation type
	CNN[31]	Effective feature extraction	Less accuracy at low SNR rate
	CNN[32]	Reduce model size and accelerate computation	Not suitable for all types modulation such as FSK, ASK, DPSK and etc.
	CNN[33]	Rapid performance during classification	Low feature extraction at low SNR rate
	CNN[34]	Higher recognition accuracy at high SNR rate	Less accuracy lack of quality assessment such as equalization, quantization and etc.

multi-carrier OFDM transmission systems. This work mainly focuses on future 5G networks [38]. The major problems of this research are discussed as follows:

- In general, PCA leads to large signal loss, which affects classification accuracy. Here, PCA only suppresses the Gaussian noise and is not able to fully eliminate the noise from signals. Other noises, such as CFO and impulse noise, are still presented in the signal even after pre-processing.
- The moment-based features help only for low-order modulation schemes. Such features fail for higher-order modulation schemes, such as 64-QAM and 128-QAM.
- Using CNN also tends to information loss because of the presence of a max pool. Thus, signal loss by PCA and CNN degrades the detection accuracy.

Authors [39] aimed to reduce the complexity of AMC systems. Three different classification algorithms are presented for modulation classification. The problems in this work are as follows:

- Here, the complexity is avoided by using simple DNN structures. However, the feature learning and classification process are still complex because the classification has to learn multi-carrier signals.
- The extracted features are insufficient to classify the various modulation schemes.

Figure 2 shows the research work plant at a low SNR environment. Recent researchers focused on addressing the problems of the low SNR environment. Our work focuses on four problems, such as CFO/IQ imbalance, inefficient feature extraction, high complexity and low classification accuracy, which are solved by our proposed work. Signal fortification solves the problem of CFO/IQ imbalance, which increases classification accuracy and reduces complexity because of the balance of the CFO/IQ. Inefficient feature extraction is solved through feature engineering, which extracts the important features that increase classification accuracy and reduce complexity. Here, signal fortification provides equalise and noise-removed signals, which

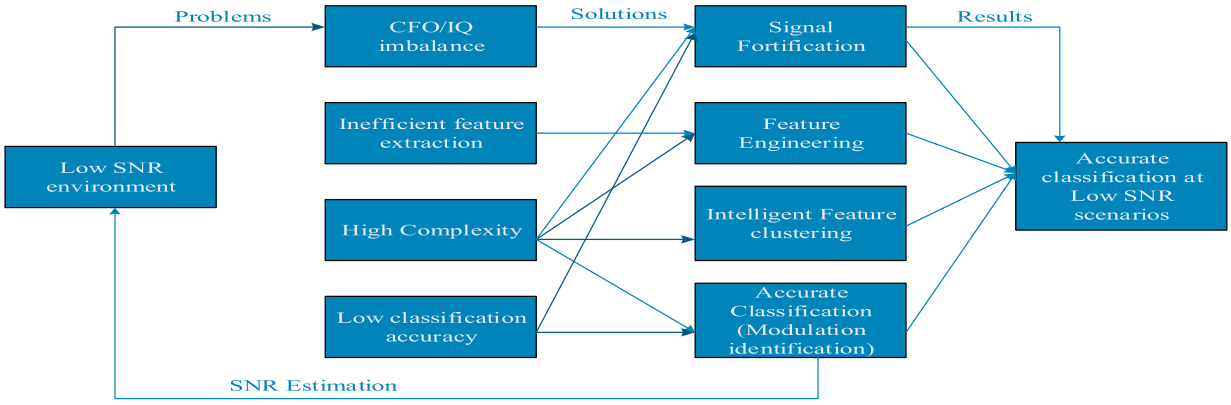

FIGURE 2. Research work plan.

reduce complexity. Moreover, the features are clustered by intelligent feature clustering, which also reduces the complexity of feature extraction from individual signals. Low classification accuracy is solved by proposing an accurate classification model and feature engineering.

Research Solutions: (1) We presented a signal fortification phase with multiple pre-processing steps that eliminate noise and compensate CFO and I/Q. This phase enhances the quality of the received signal based on the channel condition. (2) We extract instantaneous, statistical, transform and constellation features to handle a noisy environment and improve the classification accuracy. All these features are learned by GaFP-Net, which is fast and accurate. (3) For feature clustering, a novel twin-functioned-based human mental search optimiser is proposed. The proposed algorithm considers the clustering problem as an optimisation problem and solves it by considering bi-factors. This algorithm can handle non-linear signals and resolves the overlapping problem through cluster purity and entropy factors. (4) A dual distance-based classifier is presented with two distance measures, which is accurate and has less complexity. For modulation identification, a Q-learning-based decision-making procedure is introduced. Fig. 2 depicts the work plan of research work

IV. SIGNAL MODEL

Multi-carrier modulation is a type of modulation technique that aims to perform data transmission by the source to the destination through multiple carriers. Compared with single-carrier modulation, multi-carrier modulation provides significant advantages, such as high aggregated bit rate, improved bandwidth rate, cost minimised power efficiency, resilience to interference and narrowband fading.

Theorem: For a given measurement $\mathcal{X}(\mathcal{T}), 0 \leq \mathcal{T} \leq \tau$, a modulation classifier is a system to classify the type of modulation for $\mathcal{X}(\mathcal{T})$ from the set of m possible modulation as $\{\dot{\iota}_1, \dot{\iota}_2, \ldots \dot{\iota}_m\}$.

Proof: The received signal $r(\mathcal{T})$ is considered a modulated signal transmitted through the communication signal and corrupted by AWGN, which consists of intrinsic information about the signal.

$$r(\mathcal{T}) = \mathcal{X}(\mathcal{T}) * H(\mathcal{T}) + \mathcal{N}(\mathcal{T}) \tag{1}$$

where $\mathcal{X}(\mathcal{T})$ represents the originally transmitted signal, $*$ represents the convolution and $H(\mathcal{T})$ represents the impulse response for the overall signal path, which contains Transmit Pulse Shaping, Communication Channel and Receiver Antenna before the demodulation steps. Moreover, $\mathcal{N}(\mathcal{T})$ is the AWGN. The $\mathcal{X}(\mathcal{T})$ is expressed by follows:

$$\mathcal{X}(\mathcal{T}) = \mathbb{M}_{\mathfrak{A}} \sum\nolimits_{N} S_{SE} G(\mathcal{T} - N\mathcal{T}_S) \cos[2\pi (f_c + f_m)\mathcal{T} + \varphi_0 + \varphi_m] \tag{2}$$

where $\mathbb{M}_{\mathfrak{A}}$ is the modulation amplitude, S_{SE} is the symbol sequence, $\mathcal{T}_S$ is the symbol time period, f_c is the carrier frequency at IF, $f_{\mathbb{M}}$ is the modulation frequency, φ_0 represents the initial phase and $\varphi_{\mathbb{M}}$ denotes the modulation phase. $G(\mathcal{T})$ represents the gate function, which is computed as follows:

$$G(\mathcal{T}) = \begin{cases} 1 & if\ 1 \leq \mathcal{T} \leq \mathcal{T}_S \\ 0 & Other \end{cases} \tag{3}$$

For example, if M-QAM can be written as,

$$\mathcal{X}(\mathcal{T}) = \mathcal{A}_{(1)} + \mathcal{A}_{(2)} \tag{4}$$

$$\mathcal{A}_{(1)} = \mathbb{M}_{\mathfrak{A}} \sum\nolimits_{N} S_{SE} G(\mathcal{T} - N\mathcal{T}_S) \cos[2\pi f_c \mathcal{T} + \varphi_0] \tag{5}$$

$$\mathcal{A}_{(2)} = \mathbb{M}_{\mathfrak{A}} \sum\nolimits_{N} P_{SE} G(\mathcal{T} - N\mathcal{T}_S) \cos[2\pi f_c \mathcal{T} + \varphi_0] \tag{6}$$

where S_{SE} and P_{SE} are the carriers modulated, which denotes the $\left[2\mathbb{M} - 1 - \sqrt{\mathbb{M}}\right]$ and M represents the signal values of QAM, that is, 16, 32, 64 and 128. Fig. 3 illustrates the multi-carrier modulation.

V. PROPOSED AMC^2 Pyramid MODEL

A. SYSTEM OVERVIEW

The proposed work aims to classify the modulated signals in multi-carrier systems. Thus, we design a novel AMC^2 system by using a new signal fortification, pyramidal-based feature engineering and multi-distance decision-making phases.

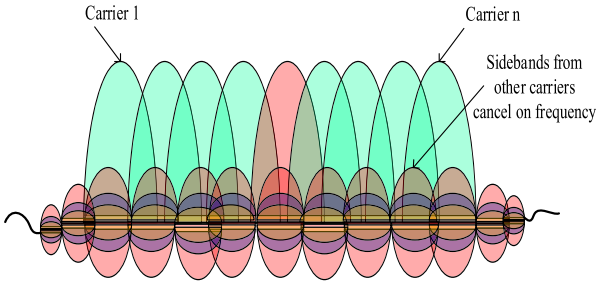

FIGURE 3. Multi-carrier modulation (e.g. OFDM).

Figure 4 shows the overall proposed work, in which we intend to classify the received signal into the following classes: 16QAM, 32QAM, 64QAM, 128QAM, QPSK, BPSK, DPSK, ASK and FSK.

B. SIGNAL FORTIFICATION

As we stated earlier, the signal received by the receiver may be corrupted by different noises and CFO, I/Q imbalance. Sometimes, the communication channel may be in ideal conditions. However, the presence of noise often affects classification accuracy. Therefore, we first fortify the received signal based on the signal quality by introducing the BfSf approach. We perform Fairness Score $\mathbb{f}_{\mathbb{s}}$-based quality evaluation at a first fold, which is described in follows:

For the transmitted signal, $\mathbb{f}_{\mathbb{s}}$ is computed by three input parameters, namely, signal to inference plus noise ratio (*SINR*), spectral efficiency δ_ε and received signal strength indicator (RSSI). Each input parameter is defined by follows:

SINR is defined as the ratio of signal power and the presence of noise in the received signal. Owing to interference to other signals, interference is presented, which is expressed as follows:

$$SINR = \frac{\mathcal{P}_S}{\mathcal{P}_N + \sum_{i=0}^{I} \mathcal{P}_i} \tag{7}$$

The SINR is computed in terms of signal power ($\mathcal{P}_S$), interference power because of i^{th} channel ($\mathcal{P}_i$) and noise power ($\mathcal{P}_N$). The core idea behind the consideration of the

SINR metric is that if the code is selected by the RSS, then the code must have a better SINR range.

The performance of δ_ε measures for the obtained signal is forwarded through transmission medium successfully. Hence, the estimation of δ_ε is performed as follows:

$$\mathrm{E} = \frac{\mathscr{r}}{\mathcal{B}} \tag{8}$$

where $\mathscr{r}$ is the information rate and $\mathcal{B}$ is the channel bandwidth. The value of RSSI is measured by the power of the received signal, and its value is computed by the amount of power transmitted and received through a particular channel, which is computed as follows:

$$RSSI = \frac{\mathcal{P}_{Tx}}{\mathcal{P}_{Rx}} \tag{9}$$

where $\mathcal{P}_{Tx}$ represents the transmitted power and $\mathcal{P}_{Rx}$ represents the received power. A significant factor in the received signal quality prediction is CSI. Most of the existing AMC works have assumed that CSI is known at the receiver, which results in less effective real-time wireless communication systems. The EM estimator is used to estimate the channel conditions with the Gaussian mixture model (GMM). The number of signal samples, $\delta_{(1)}, \delta_{(2)}, \delta_{(3)}, \ldots \delta_{(n)}$, in which the nth sample belongs to the m component of GMM model, is computed as follows:

$$\mathrm{K}(\delta[n], m) = \frac{1}{2\pi\sigma_m^2} exp^{\frac{-(\delta[n]-\theta_m)^2}{2\sigma_m^2}} \tag{10}$$

where σ_m represents the variance and the corresponding soft membership, computed as follows:

$$z(n, m) = \frac{\mathrm{K}(\delta[n], m)}{\sum_{m=1}^{M} \delta[n], m} \tag{11}$$

where θ_m represents the mean component that contains a structure, which is expressed as the transmitted symbols and channel gain. This component is computed by channel coefficient H and modulation symbol $\mathbb{M}_s$ and expressed as follows:

$$\theta_m = H * \mathbb{M}_s \tag{12}$$

For the above derivative functions, the channel coefficient and noise variance are computed as follows:

$$\frac{\partial \log \mathrm{K}(\delta, k)}{\partial H} = \sum_{m=1}^{M} \sum_{n=1}^{N} z(\delta(n), m)\gamma_m \left(\frac{-2\delta[n] \cdot \mathbb{M}_s + 2H \cdot \mathbb{M}_s^2}{\sigma} \right) \tag{13}$$

$$\frac{\partial \log \mathrm{K}(\delta, k)}{\partial \sigma^2} = \sum_{m=1}^{M} \sum_{n=1}^{N} z(\delta(n), m)\gamma_m \left(\frac{-1}{\sigma^2} + \frac{(\delta[n] - H.\mathbb{M}_s)^2}{\sigma^4} \right) \tag{14}$$

where γ_m represents the mixture proportion of the *mth* component. By setting the above equation to zero, the functions that received the channel gain and variance are computed as:

$$H_{i+1} = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} z(\delta(n), m)\gamma_m \delta[n]\mathbb{M}_s}{\sum_{m=1}^{M} \sum_{n=1}^{N} z(\delta(n), m)\gamma_m \delta[n]\mathbb{M}_s^2} \tag{15}$$

$$\sigma_{i+1} = \frac{\sum_{m=1}^{M} \sum_{n=1}^{N} z(\delta(n), m)\gamma_m \delta[n] - H_i\mathbb{M}_s}{\sum_{m=1}^{M} \sum_{n=1}^{N} z(\delta(n), m)\gamma_m} \tag{16}$$

where H_{i+1} and σ_{i+1} are the estimation of the updated parameters for $i+1$ iterations. These functions update the expectation/condition maximisation.

$\mathbb{f}_{\mathbb{s}}$ is the linear combination of the four components, namely, SINR, δ_ε, RSSI and CSI. Based on the SINR, δ_ε, RSSI and CSI, $\mathbb{f}_{\mathbb{s}}$ is computed as follows:

$$\mathbb{f}_{\mathbb{s}} = a \cdot \mathrm{SINR} + \mathrm{b} \cdot \delta_\varepsilon + c \cdot \mathrm{RSSI} + \mathrm{d} \cdot \mathrm{CSI} \tag{17}$$

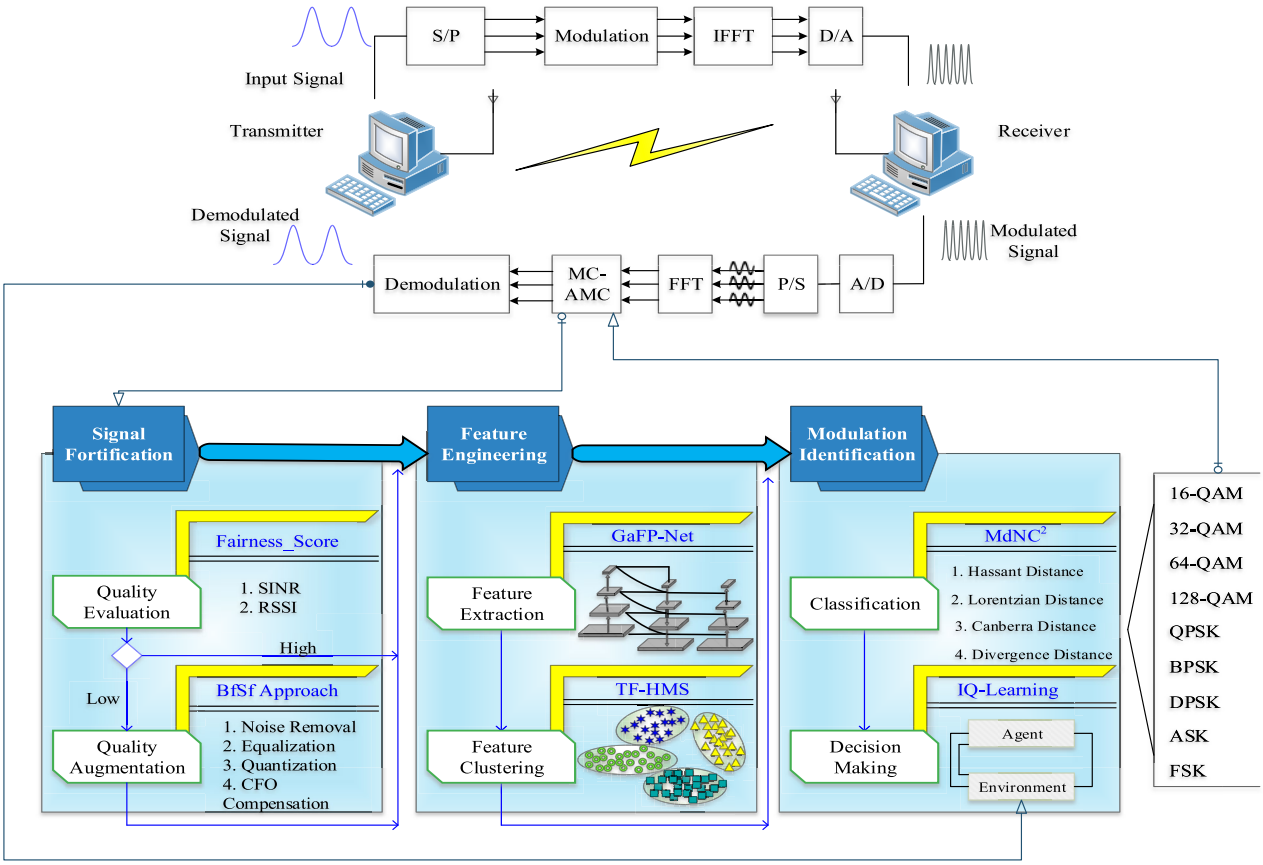

FIGURE 4. System model.

where a, b, c and d are the weights of SINR, δ_ε, RSSI and CSI, respectively. When SINR, δ_ε, CSI and RSSI of the received signal is high, $\mathbb{f}_\mathbb{s}$ is high. If the quality of the signal is too low, then the signal is augmented in the next fold. For the signals not satisfied by the $\mathbb{f}_\mathbb{s}$, augmentation operations are used. Three kinds of operations are used. In the next set of the fold, we present the *L1-Norm Filter (for noise removal), Multi-Rate Equaliser (for equalisation), One-bit Quantiser (for quantisation) and Fast Iterative Method (for CFO compensation).* At the end of this fold, the signal is enriched. If the signal quality in the first fold is high, then the signal is fed into the next phase. Otherwise, the signal is fortified and then fed into the next phase.

1) NOISE REMOVAL

If the transmitted signal consists of $\mathcal{N}(\mathcal{T})$, then the probability density function (PDF) of $\mathcal{N}(\mathcal{T})$ is computed by follows:

$$\mathbb{f}_w(x) = \frac{1}{2\pi\sqrt{\Sigma}} e^{\frac{|x|^2}{2\sqrt{\Sigma}}} \tag{18}$$

where Σ is the covariance of noise. The IQ component, that is, in-phase and quadrature components, is considered separately. Hence, the noise level of both components should be discussed [40]. The covariance σ^2 for the IQ component is derived as follows:

$$\Sigma = \begin{bmatrix} \sigma_I^2 & p\sigma_I\sigma_Q \\ p\sigma_I\sigma_Q & \sigma_Q^2 \end{bmatrix} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma^2 \end{bmatrix} \tag{19}$$

where σ_I and σ_Q are the variance of in-phase and quadrature components, respectively. For the in-phase and out-phase components, the PDF is obtained.

$$I_Q(PDF) = \frac{1}{2\pi\sqrt{\Sigma}} e^{\frac{|x|^2}{2\sqrt{\Sigma}}} \tag{20}$$

L1-norm is a simple noise filter that defines the sum of absolute values of each element in the vector [41], [42]. For a given arbitrary vector of sample n, the L1-norm filter applied for n is as follows:

$$\|L\|_1 = \sum_{j=1}^{n} v_{j(n)}, \tag{21}$$

where v_j is the jth element vectors of n. Based on the above equation, the L1-norm filter removes the noises in the signal.

2) EQUALISATION

The most significant property of equalisation is choosing the receiver's filter for compensating the radio channel frequency

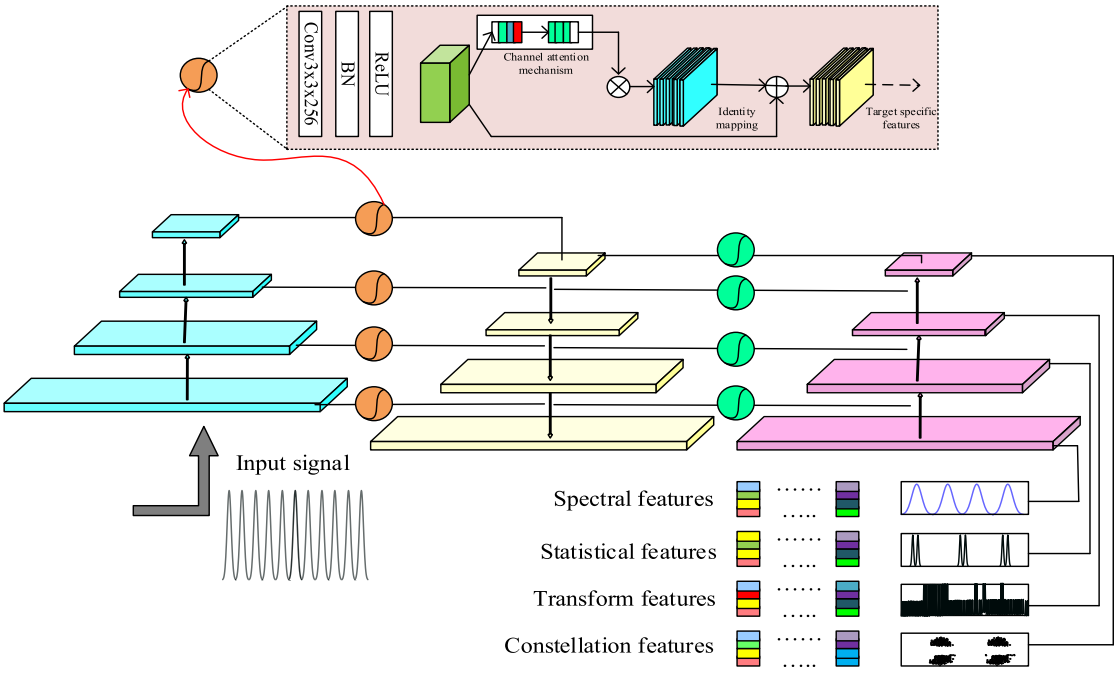

FIGURE 5. Feature extraction using GaFPNet.

selectivity completely. This is achieved by the receiver's filter impulse response, which must satisfy the following constraints:

$$W \otimes \mathcal{S} = 1 \tag{22}$$

where W is the equaliser impulse response, $\mathcal{S}$ is the channel impulse response and $\otimes$ is the linear convolution. The objective of the receiver is to reduce the mean square error (MSE) for equalising the transmitted signal, which is derived as follows:

$$MSE = \varepsilon \left\{ |\mathcal{P}_t - \mathcal{P}_t|^2 \right\} \tag{23}$$

where $\mathcal{P}_t$ and $\mathcal{P}_t$ are the estimated and actual transmitted signals, respectively.

3) QUANTISATION

In this study, we use one-bit quantiser, which is also known as an asymptotically adaptive optimiser for performing quantisation. The cumulative distribution function is defined as the $f\ (w)$ and is computed as follows:

$$f\ (w) = {}^{D * \mathcal{C}}\!/_{2\Gamma\ (1/\mathcal{C})} e^{-(D/w)^{\mathcal{C}}} \tag{24}$$

where $\Gamma(.)$ is the Gamma Function, and $D^{-1} > 0$ and $\mathcal{C} > 0$ represent the scale and shape parameters [43]. In this case, notably, the value of $\mathcal{C}$ ranges from 1, 2 to ∞.

4) CFO COMPENSATION

Assuming that the $\mathbb{F}_C$ and $\mathbb{F}'_C$ are the carrier frequency of the transmitter and receiver respectively, which is expressed as,

$$CFO = \mathbb{F}_C - \mathbb{F}'_C \tag{25}$$

The normalised $CFON(CFO)$ can be computed by,

$$N\ (CFO) = \frac{CFO}{\Delta SS} \tag{26}$$

where ΔSS represents the subcarrier spacing and $N\ (CFO)$ is adaptively address the loss of signal amplitude information at the frequency domain. Therefore, the transmitter and receiver have obtained better performance through CFO compensation [44].

C. FEATURE ENGINEERING

In this phase, we extract the important features from the enriched signals, and then, we group the features to minimise the complexity of classification. The following processes are performed in this phase.

1) PYRAMID-BASED FEATURE EXTRACTION

We present a novel GaFP-Net to learn features from the received signals. The GaFP-Net is the new Pyramid network model that learns the features rapidly without degradation in efficiency [45], [46]. In GaFP-Net, we learn *Spectral* (like time and frequency domain), *Statistical* (e.g. HOS, HOC), *Transform* (like FFT) and *Constellation* features. Recently, signals are complex in nature and consist of background noises caused by channel environments and fluctuation in signals. For instance, statistical features represent the higher-order properties of the signal.

In this study, GaFP-Net is presented, which adaptively extracts features from the signal at any frequency and amplitude. Four components are used in GaFP-Net to extract features, such as gate, channel level and global level attention, identity mapping and feature reuse. Figure 5 present the

working of GaFP-Net. Each component used in GaFP-Net is as follows:

(i) Gate: Gate is a significant function in the feature pyramid network, which is denoted as $\mathfrak{H}_f$ that converts input feature maps to output, that is, $o_t(f_m \rightarrow o_t)$ represents the set of filter maps as $f_m(1), f_m(2), f_m(3) \ldots f_m(n)$. Thus, the output function for feature maps in gate is computed as follows:

$$o_t = \mathfrak{H}_f(f_m) \tag{27}$$

(ii). Global and Channel Level Attention: The goal of channel level attention is to model the relationship between the global level feature to enrich the network to scale well for different signals. In the channel level attention, squeeze and excitation block consist SQUEEZE stage $\mathfrak{H}_{SQ}$ for global feature embedding operation, and $\mathfrak{H}_{EX}$ is used for channel feature excitation. Finally, f_m by $\mathfrak{H}_{SQ}$ and $\mathfrak{H}_{EX}$ is utilised as follows:

$$\grave{f}_m = \mathfrak{H}_{EX}(\mathfrak{H}_{SQ}(f_m)) \tag{28}$$

The $\mathfrak{H}_{SQ}$ and $\mathfrak{H}_{EX}$ stages are used for all feature maps, and the global and channel level attention is performed in the backbone network (BN).

(iii) Identity Mapping: In this step, the Element-Wise addition operation is used for the final output, which is formulated as follows:

$$o_t = \mathfrak{H}_f \oplus V \tag{29}$$

where $\oplus$ is the elementwise addition operation.

(iv). Feature Reuse: In GaFP-Net, this step plays a significant role before and after completing the gating operation. A dense feature connection is evaluated in feature reuse for accurate feature extraction.

GaFP-Net automatically generates higher-order features by computing the weights and thresholds in training. Firstly, the modulated signals are sent to the GaFP-Net for training. In this step, feature vectors are obtained, and then, the loss function is computed by the target vectors. The loss function L is computed as follows:

$$\mathrm{L} = \frac{1}{2}\sum_{k=0}^{n-1}(d_k - y_k)^2 \tag{30}$$

where L derives the standard derivation, y_k represents the output vector and d_k represents target vector. The threshold and weight values of feature vectors are updated simultaneously by L, and the update function is derived by follows:

$$\Delta\mathfrak{w}_{jk}(n) = \frac{\dot{\alpha}}{1+l} * (\Delta\mathfrak{w}_{jk}(n-1)+1) * BN_j * \mathfrak{I}_k \tag{31}$$

where $\dot{\alpha}$ is the learning rate, j is the neural unit of BN layer units, k is the output layer unit, n is the number of output neuron units, BN_j is the output vector of the BN, $\mathfrak{w}$ is the weight value which updated by L and $\mathfrak{I}$ represents the threshold value, which is computed as follows:

$$\mathfrak{I}_k = BN_j\left(1 - BN_j\right)\sum_{k=1}^{n-1}\mathfrak{I}_k\mathfrak{w}_{jk} \tag{32}$$

Each signal varies in terms of wavelength and frequency, and hence, $\mathfrak{I}_k$ and $\Delta\mathfrak{w}_{jk}$ is updated accordingly.

Figure 5 describes the work of the proposed feature extraction method. This method is constructed and guided by various fields of a gated feature pyramid model that exploits the discriminative information for different types of input signals. The proposed feature extraction network is widening and in depth instead of maximising the length of network construction. At the end of this step, the features are extracted from the signals. Considering all significant features improves the classification accuracy and provides better discrimination between modulation techniques. Table 3 presents the list of extracted features.

2) INTELLIGENT FEATURE CLUSTERING

After feature extraction, we perform a feature clustering process to minimise the complexity during classification. To form feature clusters, we propose a novel TF-HMS algorithm, which provides a better-clustered result than the Neutrosophic C-Means clustering algorithm [1]. HMS is a Meta-heuristic population-based optimisation algorithm that works by bid space search. Similar to other population-based approaches, HMS is performed by creating a set of candidate solutions randomly. A form of each solution is called "Bid." To search a bid over the $\mathfrak{q}$ dimensional space, the population is defined as follows:

$$Bid = \{x_1, x_2 \ldots x_{\mathfrak{q}}\} \tag{33}$$

Each bid in equation (33) is accessed by the specific objective function $OF()$, which is evaluated by the candidate solution quality:

$$OF(bid) = OF(x_1, x_2 \ldots x_{\mathfrak{q}}) \tag{34}$$

Based on the mental search operation, bids in $\mathfrak{q}$ are grouped for a maximum number of iterations. In the proposed TF-HMS, a form of a bid is feature vectors $\mathfrak{J}_{\mathbb{V}}$, which encode the cluster centre. An array length is represented as K, which is a number of clusters [48]. The upper and lower bound values are computed for each bid, which is represented as $L = \min(\mathfrak{J}_{\mathbb{V}})$, $U = \max(\mathfrak{J}_{\mathbb{V}})$, where the upper and lower bounds are the minimum- and maximum-based feature vectors. For that, our objective function is computed by three parameters: (1) expressing error, (2) intra-cluster distance and (3) inter-cluster distance. The weight value of these parameters is as follows:

$$\begin{aligned} OF(Bid, \mathfrak{J}_{\mathbb{V}}) = {} & \mathsf{C}_1 d_{max}(Z, Bid) \\ & + \mathsf{C}_2\left(Z_{max} - d_{min}(\mathfrak{J}_{\mathbb{V}}, Bid) + \mathsf{C}_3 I_e\right) \end{aligned} \tag{35}$$

where Bid represents the $[x_1, x_2 \ldots x_k]$ with x_i representing the ith cluster centre, Z_{max} is the maximum data value and C_1, C_2 and C_3 are the weight values of $\mathfrak{J}_{\mathbb{V}}$.

The algorithm procedure for TF-HMS is as follows:

Step 1 (Initialise Parameters): A set of input parameters are initialised in this step, and the parameters are number of clusters K, number of bids N, the number of clusters for bid

TABLE 3. List of extracted features.

Features	Subfeatures	Formula [1], [47]	Subfeatures	Formula [1]
Spectral Domain features	Variance	$F_{var} = [X(t) - F_m]$	Kurtosis	$F_{kur} = \frac{\sum_{i=1}^{n}(X(t) - F_m)^4}{(n-1)F_R^4}$
	Standard Deviation	$F_{std} = \sqrt{\frac{\sum_{i=1}^{n}(X(t) - F_m)^2}{n-1}}$	Skewness	$F_{skew} = \frac{\sum_{i=1}^{n}(X(t) - F_m)^3}{(n-1)F_R^3}$
	Mean Frequency	$F_m = \frac{\sum_{i=1}^{n}(X(t)}{n}$	Median Frequency	$F_{me} = \frac{\sum_{i=1}^{n} X(t)}{2}$
	Power Bandwidth	$F_p = \frac{\sum_{i=1}^{n}(pX(t)}{n}$	Zero Cross Rate	$F_z = \sum_{i=1}^{n} sgn\, X(t).x(t+1)$
	Amplitude	$X(t) = a(t)\cos(\omega_0 t)$	Phase	$X(t) = a(t)\cos(\emptyset t)$
	Frequency	$F(t) = \frac{d}{dt}\emptyset t$		
Statistical features	Cumulant	$F_{cu} = logE[e^{X(t)}]$	Moment	$F_{mo} = \sum_{i=1}^{n}\frac{X(t)}{t}$
Transformation features	Mean Absolute Value	$F_{ma} = \frac{1}{n}\sum_{i=1}^{n}\lvert X(t)\rvert$	Root Mean Square	$F_R = \frac{\sum_{i=1}^{n}(X(t)^2}{n}$
Constellation features	Constellation Shape	$F_C = \sum_{i,j=1}^{N} P_i, L_i$	Symmetric Structure	$F_{SS} = \sum_{i,j=1}^{N}\theta_i$

grouping K_{BG} and the mental searches (Min and Max) M_{min} and M_{max}

Step 2 (Population Initialisation): For the set of bids N, population P is defined as follows:

$$P = \begin{bmatrix} Bid_1 \\ .... \\ Bid_N \end{bmatrix} \tag{36}$$

Step 3 (Objective Function Calculation): For each Bid_i, the objective function is computed.

Step 4 (Choose the Best Bid): Choose the best bid that consists of better objective values.

Step 5 (Select Random Number): For every Bid_i, select the random value between M_{min} and M_{max}.

Step 6 (Mental Search Operator): Make the new bids in the vicinity of existing bids by levy function, which is carried out as follows:

$$N_{POS} = Bid_i + S \tag{37}$$

where S is the number of steps, which is computed by the maximum number of iterations and random values and is expressed as follows:

$$S = \left(2 - I * \left(\frac{2}{Max_I}\right)\right) * \lambda \oplus Levy \tag{38}$$

where Max_I represents the maximum number of iterations, I represents the current iteration, λ is the random variable and the symbol $\oplus$ is the element-wise multiplication. Then, the step size S is written as follows:

$$S = \left(2 - I * \left(\frac{2}{Max_I}\right)\right) * 0.01 * \frac{U}{V^{\frac{1}{\beta}}} * \left(x^i - x_*\right) \tag{39}$$

where x_* represents the best position and U and V are the two random variables, which represent the normal distribution.

Step 7 (Replacement Operator): In this step, the new bid is replaced with its previous bid when it is better.

Step 8 (Bids Grouping): This step groups the bid population into a number of clusters.

Step 9 (Compute Objective Function): This step computes the objective function value for each cluster member. A cluster within the minimum mean objective function is chosen as the final cluster. Then, the remaining bids move towards the best bid available in the winner group.

Step 10 (Terminate Conditions): Do until the stop criterion is satisfied; otherwise, go back to step 2.

For $\mathfrak{J}_{\mathbb{V}}$, the k number of feature clusters are formed, such as $C_1, C_2 \ldots C_k$. Each cluster must be related to the two functions: Cluster Purity $\varsigma_{\mathfrak{p}}$ and Cluster Entropy ς_{ε}. Their computations are illustrated below:

$$\varsigma_{\mathfrak{p}} = \frac{1}{N_i}\sum_{x_j \in c_j}^{n} d(x_j, m_j) \tag{40}$$

$$\varsigma_{\varepsilon} = \frac{x_i + x_j}{d(m_i, m_j)} \tag{41}$$

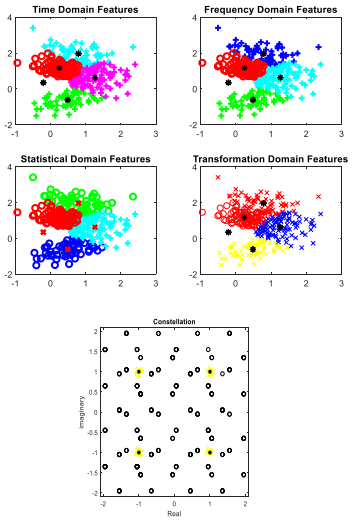

FIGURE 6. Feature clustered result.

where N_i is the number of samples in the *ith* cluster, and $d(x_j, m_j)$ is the distance between sample x_j and center value m_j. The TF-HMS optimiser improves the clustering performance based on $\varsigma_{\mathfrak{p}}$ and ς_{ε} functions.

D. MODULATION IDENTIFICATION

In this phase, we propose classification and decision-making processes to achieve accurate modulation identification. Both processes are explained as follows:

1) MULTI-CARRIER MULTI-LEVEL CLASSIFICATION

The received multi-carrier signals are fed into the classification module, in which we present a multi-level classification. Thus, we design the MdNC^2 algorithm. In the MdNC^2 algorithm, the received multi-carrier signals (clustered feature vectors) are classified in parallel by using efficient distance measures, such as Hassanat Distance (*HasD*), Lorentzian Distance (*LoD*), Canberra Distance (*CanD*) and Divergence Distance (*DivD*). The distance between feature clusters are formulated as follows:

$$HasD(X, Y) = \sum_{i=1}^{N} D(X_i, Y_i) \tag{42}$$

where,

$$D(X_i, Y_i) = \begin{cases} 1 - \dfrac{1 + Min(X_i, Y_i)}{1 + Max(X_i, Y_i)} Min(X_i, Y_i) \geq 0 \\ 1 - \dfrac{1 + Min(X_i, Y_i) + |Min(X_i, Y_i)|}{1 + Max(X_i, Y_i) + |Min(X_i, Y_i)|} Min(X_i, Y_i) < 0 \end{cases} \tag{43}$$

The distance value will be in the range of 0 to 1. *LoD* represents the natural log of the absolute difference between two feature vectors, and when the distance value is smaller, the result will be sensitive. Owing the log scale expands the low range and compresses to a higher range. The *LoD* is computed as follows:

$$LoD(X, Y) = \sum_{i=1}^{N} ln(1 + |X_i - Y_i|) \tag{44}$$

where *ln* is the natural log value and ensures the non-negativity property and that log value zero is avoided by adding the distance with 1. The *CanD* is adopted for high-dimensional and non-linear space, which is another version of Manhattan Distance, where the absolute difference between two feature clusters is computed in this study. The *CanD* is also highly sensitive to small changes. Therefore, the *CanD* is defined as follows:

$$CanD(X, Y) = \sum_{i=1}^{n} \frac{|X_i - Y_i|}{|X_i| + |Y_i|} \tag{45}$$

Finally, the *DivD* is defined as follows:

$$DivD(X, Y) = 2 \sum_{i=1}^{N} \frac{(X_i - Y_i)^2}{(X_i + Y_i)^2} \tag{46}$$

Before computing the weight by all distance measures, the range of the distance value for each metric is normalised. However, the range of the distance metric is obtained by different measurements, which leads to the imbalance in the weight computation and nearest centroid assignment. Hence, we normalise the distance value as follows:

$$Normalize_{D(xy)} = \frac{\frac{[\Im_{k(\mu)} - \delta_{(x,y)}]}{2}}{3\sigma_{(x,y)}} \tag{47}$$

where $\Im_k$ represents the feature clusters and $\delta_{(x,y)}$ and $3\sigma_{(x,y)}$ are the mean and standard deviation, respectively, of the similarity between the distance values. The $Normalize_{D(xy)}$ function ensures that all distances have an equal amount of importance. The weight value w_i is computed by all distance values as follows, which is the sum of all distance measures.

$$w_i = \frac{\mu_1(x, y) + \mu_2(x, y) + \mu_3(x, y) + \mu_4(x, y)}{4} \quad 0 \leq \mu_i \leq 1 \tag{48}$$

where μ_1, μ_2, μ_3 and μ_4 are the distance values of *HasD*, *LoD*, *CanD* and *DivD* between x and y clusters, respectively. Based on w_i, the nearest centroid $\mathcal{N}_{\mathcal{C}}$ is computed for the particular signal. Subsequently, $\mathcal{N}_{\mathcal{C}}$ is forward to the next step for the type of modulation identification.

2) REINFORCED DECISION-MAKING

Typically, the accuracy of an AMC system completely depends on the demodulation process. If the signal is demodulated successfully, then the AMC works accurately. Otherwise, the performance of the AMC system is poor. By considering this fact, we present an IQ-L procedure for accurate MC-AMC. The IQ-L algorithm takes the classification results and identifies the modulation schemes as QAM (16, 32, 64, 128), PSK (QPSK, BPSK, DPSK), ASK and FSK. The IQ-L algorithm takes an action by learning the environment (demodulation error).

Algorithm 1 Algorithm for IQL

Initialize: Q-table, State s_t, s_{init}, control variables a and b, number of feature clusters F_c, learning rate α, discount factor γ and transmitted signal

1: **Loop** for each episode
2: Initialize $F_c = 1$ and $s_t = s_{init}$
3: **While** $(\eta < \eta_c)$ // $\eta_c = threshold$
4: Choose action a_t according to ε greedy policy
5: Take action a_t and observe r_{t+1} and s_{t+1}
6: Compute Q value

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma max_a Q(S_{t+1}, a) - Q(s_t, a_t)]$$

7: $s_t = s_{t+1}$;
8: **End While**

In the proposed AMC^2-pyramid method, to classify modulation types, the IQL elements are defined as follows:

- **Environment:**The perceived environment considered in this study is transmitted signal, feature vector clusters and distance between true and estimated feature vectors.
- **Agents:** The modulation type decision-making step is denoted as an agent. Concurrently, the agent interacts with the perceived environment over time.
- **States:** A list of states for an agent consists of the following observations:
  - A vector of feature values and assigned cluster
  - Transmitted signal
  - Distance between feature vector clusters
- **Actions:** A list of actions for the corresponding transmitted and trained signals 16-QAM, 32-QAM, 64-QAM, 128-QAM, QPSK, BPSK, DPSK, ASK and FSK.
- **Reward:** Reward represents the reciprocal classification error. If the reward function has a positive value, then the classification result is less than the threshold value. Otherwise, the agent received a negative reward. Whenever the agent is close to the target, the agent gains additional rewards. Therefore, the reward function is computed as follows:

$$R_t = \begin{cases} \dfrac{1}{\|o_t - \rho_t\|} & if\ 0 < |\,\|o_t - \rho_t\| \leq \delta \\ -\|o_t - \rho_t\| & Otherwise \end{cases} \quad (49)$$

where o_t and ρ_t are the observed and target results, respectively. The procedure of IQL is given as follows:

VI. EXPERIMENTAL RESULTS

This section investigates the experimental results of the proposed AMC^2-pyramid method. To provide a brief analysis of the proposed AMC^2-pyramid method, this section is further divided into three subsections, namely, simulation setup, comparative analysis and results and discussion.

A. SIMULATION SETUP

The simulation setup for the proposed AMC^2-pyramid method is presented in this section. The simulation environment is created and analysed using MATLAB R2017b. The total number of training and testing signals is 1024 samples.

TABLE 4. Modulated signal settings.

Parameter	*Value*
Carrier frequency	30 MHz
Sampling rate	1.25 MHz
Symbol rate	300 kHz
Number of symbol	1024
Sampling frequency	320MHz
Symbol transmission rate	1.5Mbps
Number of paths	3
Fading of each path	Obey Rayleigh Fading
Time delay	1-2 seconds
Phase offset	0 – 2pie
Number of modulations	9
Types of modulation	16QAM, 32QAM, 64QAM, 128QAM, QPSK, BPSK, DPSK, ASK and FSK
Temporal resolution	128microseconds

TABLE 5. Algorithm parameter setting.

Algorithm	*Parameters*	
GFPN	Epochs	12
	Schedule Rate	1× 1
	No. of Iterations	500
	Learning Rate	0.01-0.001
	Weight Decay	0.00001
	Batch Size	2 per GPU
	Optimizer	SGD
	Momentum Size	0.9
IQL	Learning Rate	0.0001
	Batch Size	20
	Discount Rate	0.98
	Initial Exploration	1
	Exploration Decay Rate	0.0007
	Replay Memory Size	10,000
	Final Exploration	0.2

Table 4 shows each sample simulation setting, and the signals are classified into 16QAM, 32QAM, 64QAM, 128QAM, QPSK, BPSK, DPSK, ASK and FSK. Fig 8 illustrates some samples generated for evaluation.

To test and train the number of signals, AWGN is added with the SNR rate between −10 and 10 dB. Proving a better performance is significant when modulated signals consist of noisy values.

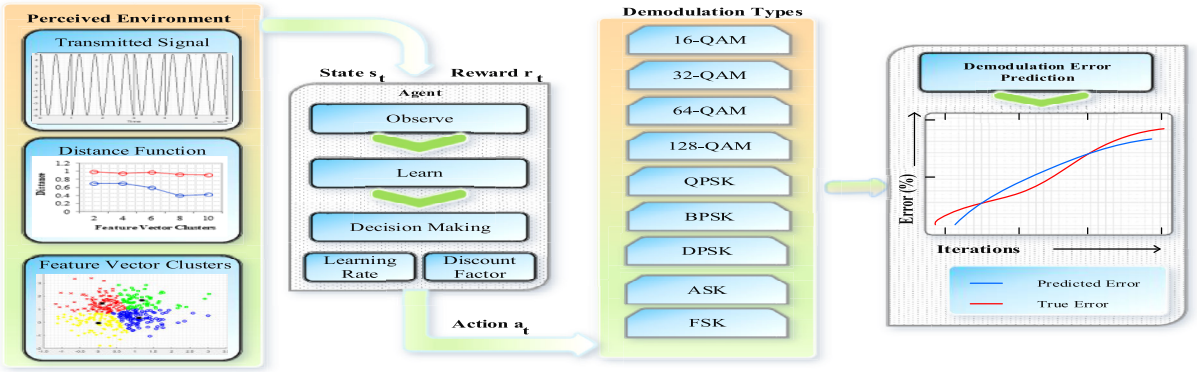

FIGURE 7. Working of IQL.

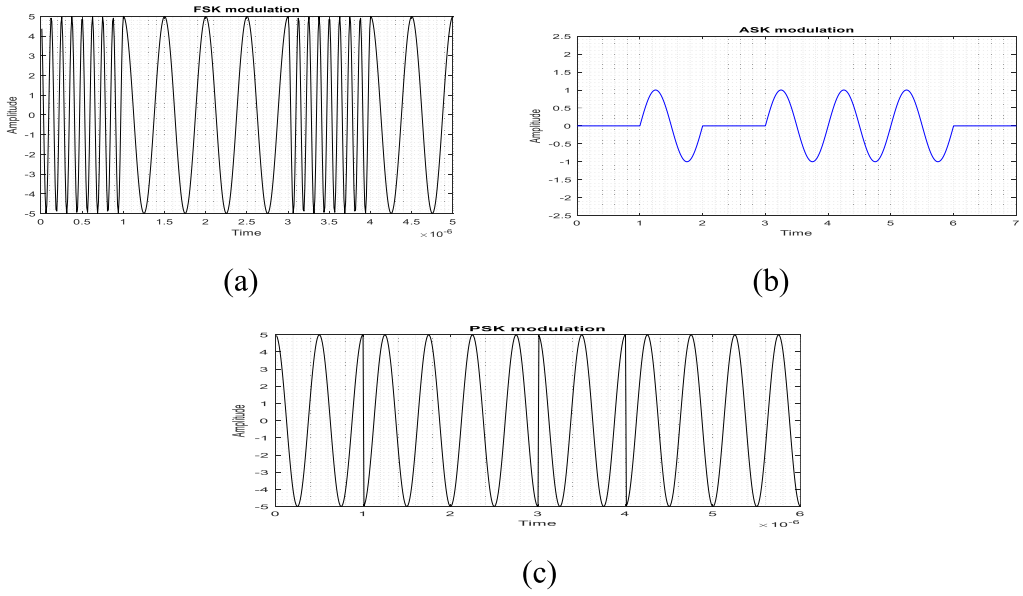

FIGURE 8. (a) FSK modulation, (b) ASK modulation, and (c) PSK modulation.

B. COMPARATIVE ANALYSIS

To compare the proposed AMC^2-pyramid method, four sets of performance metrics are used, that is, accuracy, precision, recall and F-score. The simulation results of the proposed AMC^2-pyramid method is compared with the existing methods, namely, NCA [35], PSO-DNN [36], OFDM [37] and RBF-DNN [39]. The concepts of those previous works are related to the proposed AMC^2-pyramid method with the core intention of modulation type identification for the multi-carrier system. The processes undertaken by the previous works are preprocessing, feature extraction, clustering and classification. Furthermore, Table 6 shows a qualitative comparison of the proposed AMC^2-pyramid method and the previous methods. The performance metrics are evaluated for two different cases, namely, the sample size and SNR variations from low to high. The definition, mathematical formulations and the comparison of the proposed AMC^2-pyramid method are defined as follows:

1) ACCURACY

Accuracy is a significant metric for measuring the modulated signal performance. Moreover, accuracy is estimated by successfully predicting the modulation types to the sum of true values of the received signal at the receiver, which isr computed as follows:

$$Accuracy = \frac{\#\ of\ correct\ predictions}{Total\ number\ of\ predictions} \tag{50}$$

$$(or) = \frac{TP + TN}{TP + TN + FP + FN} \tag{51}$$

where TP is the true positive, TN is the true negative, FP is the false positive and FN is the false negative.

Accuracy is computed for the proposed AMC^2-pyramid method by different distance measures, such as Lorentzian, Canberra, Divergence, Hassanat and Multi-Distance, which is shown in Figure 9(a) and (b). A distance-based classifier is a non-parameter method, where its performance is higher

TABLE 6. Comparative analysis.

Referece ns	*Objectives*	*Measurements*	*Downsides*
AMC2N [1]	To design a modulation type classification system for single carrier modulations	*Modulation Type:* QPSK, BPSK, ASK, FSK, 16-QAM, and 64-QAM *System Type:* Single- carrier *SNR:* Low and high SNR rates *Method Used:* TL-CapsNet *Processes:* Double layer feature extraction based on capsules	• Not efficient for multicarrier modulating schemes • For low and high SNR it is less efficient.
NCA [35]	To design multi-carrier system under higher (positive) and worst case (zero) SNR variations	*Modulation Type:* ASK, FSK, and PSK *System Type:* Multi-Carrier *SNR:* High (positive) *Method Used:* Neutrosophic C-Means with Feature Weighting *Processes:* Feature Extraction and Classification	• Less effective under low SNR variations • Error rate is high for different modulation schemes
PSO-DNN [36]	To create modulation type classification system for wireless communication systems	*Modulation Type:* BPSK, QPSK, 8PSK, 16 QAM, and 256 QAM *System Type:* Multi-Carrier *SNR:* High *Method Used:* PSO and DNN *Processes:* Feature Engineering and Classification	• Less accurate under higher order modulation type • Lack of effective feature consideration and not suited under noisy environment
OFDM [37]	To design of blind digital modulation type identification for OFDM and MIMO systems	*Modulation Type:* QPSK and 16QAM *System Type:* Multi-Carrier *SNR:* Low to High *Method Used:* Machine Learning Algorithms *Processes:* CFO Compensation, Channels Estimation, STBC Decoding, Feature Estimation and Classification	• Not applicable in higher order modulation schemes • Computation time is higher in processing all extracted features to the machine learning algorithms (Multi-Layer Feed Forward ANN, RF, KNN, and SVM)
RBF-DNN[39]	To explore STBC-MIMO based modulation classification system	*Modulation Type:* 2PSK, 4PSK, 8PSK and 16QAM *System Type:* Multi-carrier *SNR:* High *Method Used:* RBF, and DNN *Processes:* Channel State Information Estimation and Classification	• High degradation of accuracy under low SNR variations • Not suited for practical cases • Features are not effective to classify the modulation
AMC^2-Pyramid	To implement AMC for multi-carrier system using accurate feature extraction and signal quality estimation metrics	*Modulation Type:* 16QAM, 32QAM, 64QAM, 128QAM, QPSK, BPSK, DPSK, ASK, and FSK *System Type:* Multi-carrier *SNR:* Low to high *Method Used:* GaFP-Net, TF-HMS, MdNC^2, and IQL *Processes:* Preprocessing, Feature Extraction, Clustering and Classification	Nil

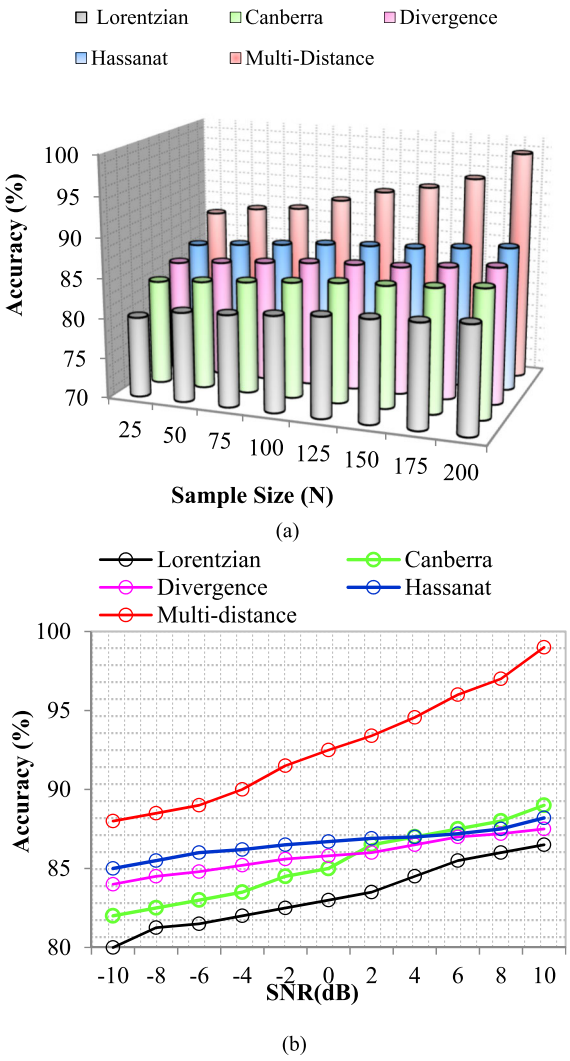

FIGURE 9. (a) Accuracy for different distance measure vs. sample size. (b) Accuracy for different distance measure vs. SNR.

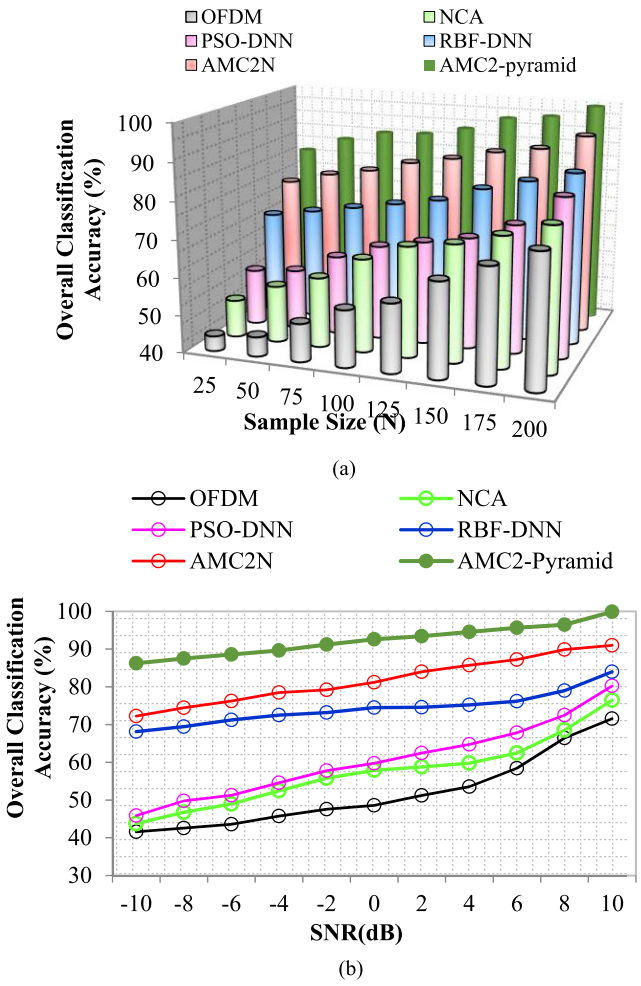

FIGURE 10. (a) Sample size vs. accuracy. (b) SNR vs. accuracy.

than parameter-based methods. A similarity between feature vectors is measured by distance functions. In multi-distance, the mean value gives a better modulation result.

It is computed by the distance weighting in all distance measures. For each test feature vector, the non-parametric method computes the distance to all trained feature vectors. Compared with multi-distance, other distance metrics produces less accuracy. When the feature dimension is very large, then a single-distance function is not sufficient to predict the true values. Hence, the mean value computation by multiple distance functions provides an accurate modulation type for the obtained signal.

For different SNR values (−10 to 10 dB), the performance of accuracy is computed, which is shown in Figure. 9 (b). Computing a single-distance function using feature vectors result in a problem in a modulation type classification, that is, it leads to misclassification. From the graphical results, the multi-distance function represents the best outcome than other distance functions. Furthermore, the performance of extracted feature vectors under different SNR values produces a higher accuracy.

The overall classification accuracy is computed in terms of sample size and SNR variations, which are given in Figure. 10(a) and (b), respectively. The sample size represents the sum of instances considered for classification. The overall classification accuracy is defined as the probability of correct classification (PCC) rate. The most significant open issue in AMC is how to improve the overall classification accuracy for different modulation types in multi-carrier/single-carrier systems. This study is based on the above statement. Thus, the proposed AMC^2-pyramid method fix this open issue by proposing novel contributions, such as noise level reduction, signal quality quantisation, CFO compensation, multiple feature extraction, and distance-based classification. These contributions help to improve the accuracy. As a result of optimised parameters (noiseless environment, adequate feature vectors for different SNR ranges), the overall classification accuracy for the proposed AMC^2-pyramid method has obtained a better outcome.

To conclude the performance of accuracy with respect to the size, the proposed model has used more effective

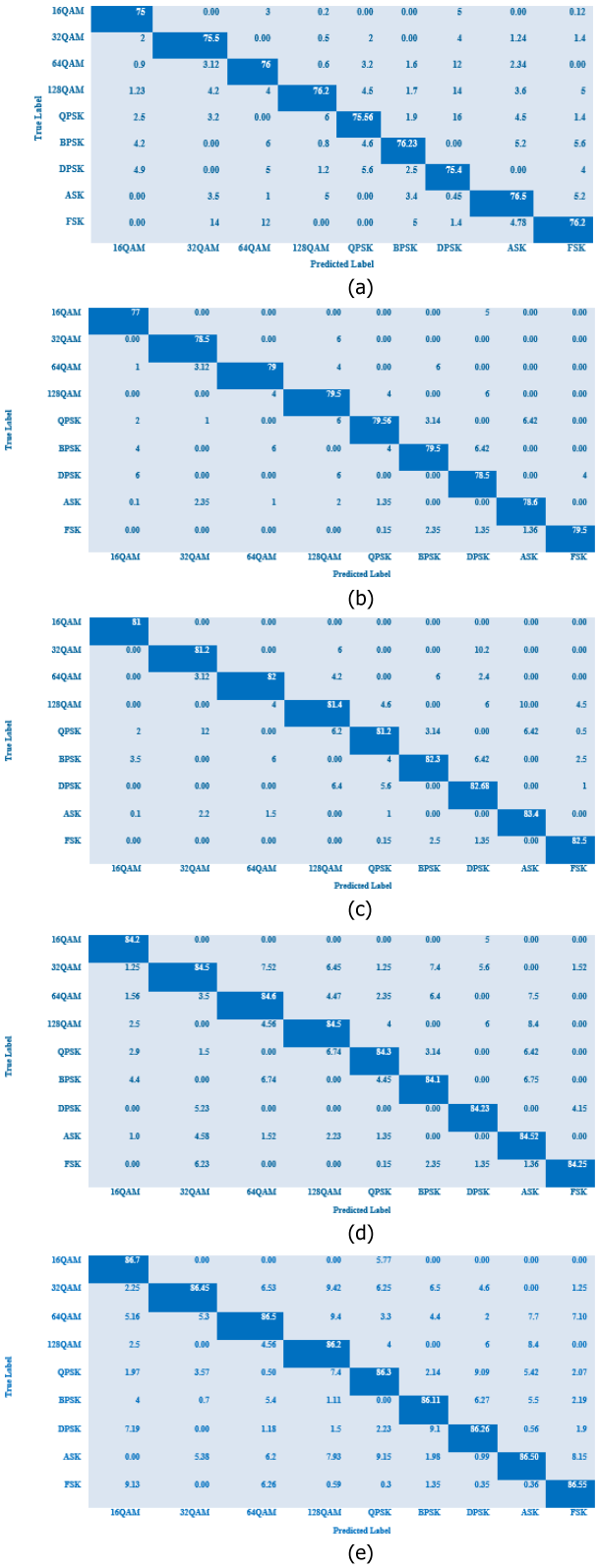

FIGURE 11. (a) Confusion matrix for OFDM. (b) Confusion matrix for NCA. (c) Confusion matrix for PSO-DNN. (d) Confusion matrix for RBF-DNN. (e) Confusion matrix for AMC2N. (f) confusion matrix for AMC^2-pyramid.

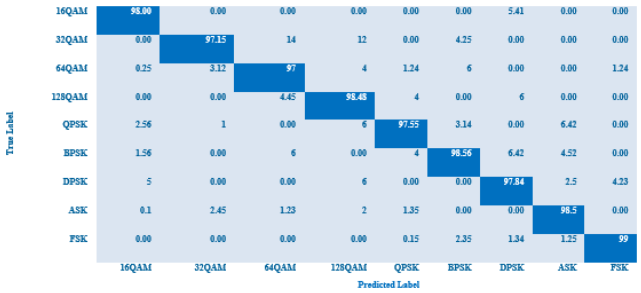

FIGURE 11. *(continued.)* (a) Confusion matrix for OFDM. (b) Confusion matrix for NCA. (c) Confusion matrix for PSO-DNN. (d) Confusion matrix for RBF-DNN. (e) Confusion matrix for AMC2N. (f) confusion matrix for AMC^2-pyramid.

features and algorithms for each processing, which provides better results than previous works. When the sample size is getting increased the accuracy of our proposed work is 98.8% which gives better results than AMC2N (92%), RBF-DNN (84.3%), NCA (80.9%), PSO-DNN (78.7%) and OFDM (75.65%). Thus, the proposed model obtained better accuracy although the number of samples increases. Form the comparison it is said that our proposed work achieve (6.8% - 23.15%) high accuracy in terms of sample size than existing works

In Figure. 10 (b), the overall classification accuracy is given for SNR variations, and the analysis defines that the proposed model obtains higher classification accuracy than previous works. The major reason behind this higher accuracy is feature extraction, clustering and distance computation. All these processes are significant in a signal modulation classification. For example, spectral features not only improve the classification but also help in reducing the demodulation error. The previous works are better when the SNR is greater than 0 dB but is worse when SNR is less than 0 dB. The reason is the lack of feature extraction for a modulated signal. Different modulation modes are defined with different probability rates, which are different based on the signal sampling point and a data frame. Under the low SNR (e.g. -10 dB), the proposed work obtains 86% of accuracy, whereas the previous works have obtained 72%, 69%, 48%, 45% and 40% for AMC2N, RBF-DNN PSO-DNN, NCA and OFDM respectively. Form the comparison it is said that our proposed work achieves (14% - 46%) high accuracy in terms of SNR at -10 dB level than existing works.

A result of the confusion matrix for accuracy is computed for SNR at 10 dB. A confusion matrix is a $N \times N$ matrix that aims to evaluate the performance of a classifier, in which N represents the number of classes to be classified. This matrix compares the true value of the class in the training to the predicted value in testing. For our case, that is, multi-modulation type classification problem, the size of the confusion matrix is 9×9. The TP, TN, FP and FN for each modulation are computed by adding the cell values.

The value of the confusion matrix is evaluated to the proposed AMC^2-pyramid method and previous works as OFDM, NCA, PSO-DN, RBF-DNN and AMC2N which are shown in Figure. 11(a)–(f). For all modulation types, namely, 16QAM, 32QAM, 64QAM, 128QAM, QPSK, BPS, DPSK, ASK and FSK, the AMC^2-pyramid method exhibits better results than previous works. In detail, the AMC^2-pyramid method has obtained $\cong$ 99% of accuracy under 10 dB SNR value.

2) PRECISION

Precision is defined as the measure of percentage for true positive count (predicted accurate modulation types) to the total number of obtained types of modulations and is computed as follows:

$$Precision = \frac{TN}{TN + FP} \tag{52}$$

Figure 12(a) shows the result of precision in terms of sample size. The use of spectral-domain information helps in improving the performance of precision than the HOS and HOC features. Previous works used a raw set of features for classification, which is not effective even under high SNR values. Furthermore, previous works established the classification model with the addition of white Gaussian noise. As a result of these effects, the precision is greatly reduced in previous works. The precision with respect to the overall sample size ($N = 200$) for proposed is 97.52% and for existing works are 92.56%, 82.25%, 79.6%, 72.22% and 68% for AMC2N, RBF-DNN, PSO-DNN, NCA, OFDM respectively. Form the above results it is shown that the proposed work precision rate in terms of sample size is (4.96% - 29.5%) higher than existing works.

By considering the demodulation error in IQL, the reward is accurately computed, and action is taken based on the current reward. In addition, the precision of the AMC^2-pyramid model increases rapidly with SNR from low to high. The proposed AMC^2-pyramid model is highly capable of classifying the modulation type with good precision. Under low SNR ranges, previous algorithms are constrained by AWGN noise and also poor CFO compensation. Thus, the performance of precision has not reached the extent. When the SNR is 10 dB, the proposed method can reach above 99% when compared to existing works which has precision in terms of SNR at -10 dB for proposed is 88.5% and for existing works are AMC2N (72%), RBF-DNN (69.5%), PSO-DNN (42.5%), NCA (41.5%) and OFDM (40%). Figure 12(b) depicts the simulation results with respect to SNR variation. From the above results it is shown that the proposed work precision rate in terms of SNR at -10 dB is (16.5% - 48.5 dB) higher than existing works.

3) RECALL

To better explain the efficiency of the proposed work, recall is evaluated with respect to the number of samples and SNR

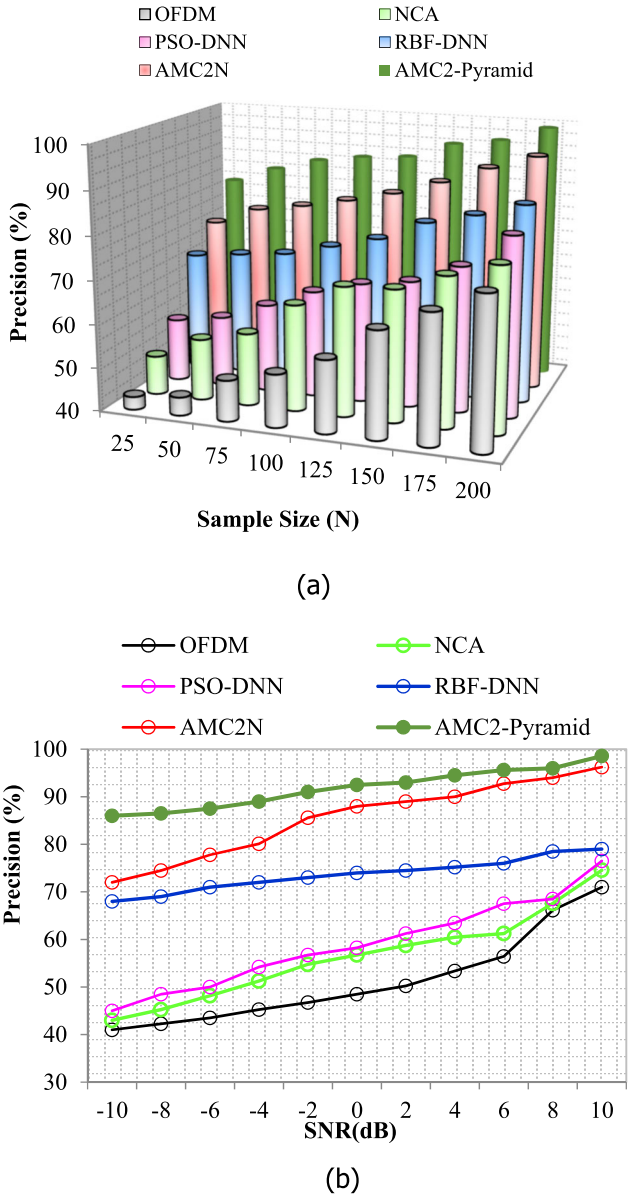

FIGURE 12. (a) Sample size vs. precision. (b) SNR vs. precision.

variations. Recall is described as the computation of the sum of accurate modulation type identified samples from the total number of positive predictions. This metric describes the comprehensiveness of any classification system as it indicates the missed positive predicted result, which is computed as follows:

$$Recall = \frac{TP}{TP + FN} \tag{53}$$

Figures 13(a) and (b) shows the recall of the proposed work and previous works with respect to sample size and different SNR levels. From the analysis, the performance of the proposed work in terms of recall outperforms the existing

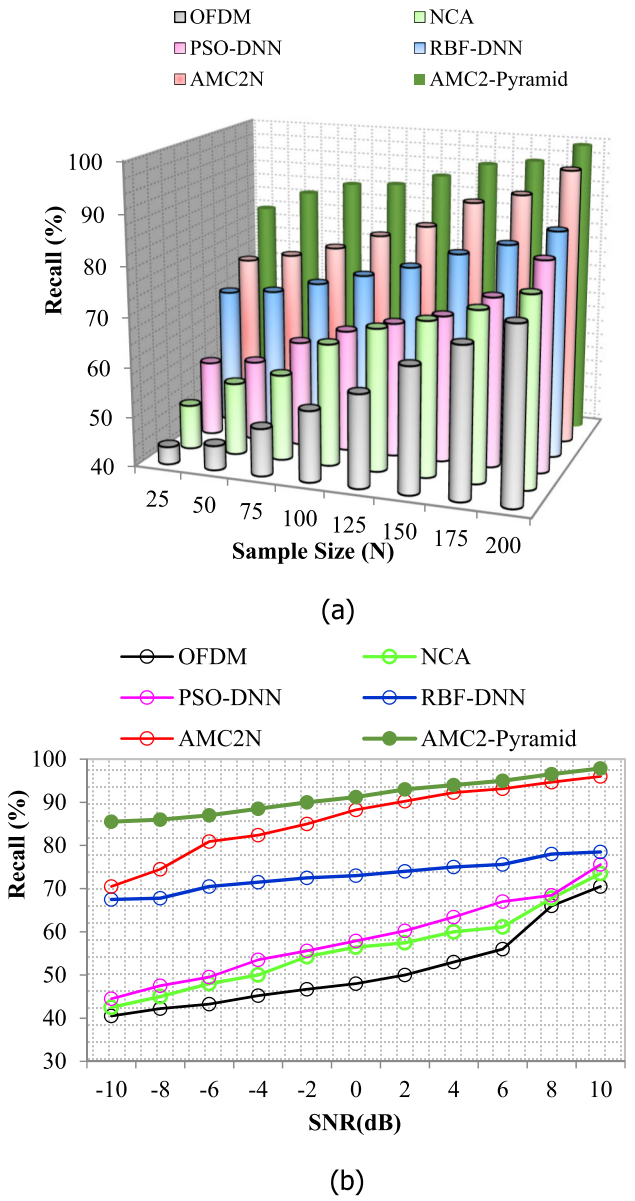

FIGURE 13. (a) Sample size vs. recall. (b) SNR vs. recall.

works. However, signal features, such as spectral domain, statistical, transform and constellation features, are used in the classification step. For any high-order modulation scheme, the proposed work has obtained a higher recall than that of previous works. Owing to the inefficiency in pre-processing steps, namely, denoising, equalising, quantisation and CFO compensation, and limited set of features, the recall performance is lower in terms of sample size and SNR variations.

As shown in Figures 13(a) and (b), the proposed work performance in terms of sample size vs recall 99% which shows the better performance than the existing works AMC2N (97%), RBF-DNN (87.8%), PSO-DNN (80%), NCA (73%) and OFDM (69.23%). The proposed work performance in terms of SNR at −10 dB vs recall is 85% which shows better performance than existing works AMC2N (71%), RBF-DNN (68.5%), PSO-DNN (46.5%), NCA (42%) and OFDM (40%). From the above results it is shown that the recall rate in terms of sample size is (2 – 29.76) high and recall rate in terms of SNR at −10 dB is (14% - 45%) high than existing works.

4) F-SCORE

F-score is a harmonic mean of the precision and recall values, and f-score is computed through the precision and recall of the weighted value. When F-score reaches a value of 1, the classification model performance is high, which is expressed as follows:

$$F - Score = \frac{2 * Precision * Recall}{Precision + Recall} \tag{54}$$

The simulation result of the f-score is computed for two cases: sample size and SNR variations. Fig. 14(a) and (b) shows the representation of graphical plots for f-score. According to the results depicted in Fig. 14(a) and (b), the proposed work in f-score obtains higher values than the previous methods. In 14 (a) the comparison of F-score with sample size which shows the proposed method f-score is 98.5% weather the existing methods such as AMC2N (96.5%), RBF-DNN (85%), PSO-DNN (82%), NCA (78%) and OFDM (68.5%). In 14 (b) the comparison of F-score with SNR at −10 dB which shows the proposed method F-score is 88% weather existing methods such as AMC2N (72.5%), RBF-DNN (68%), PSO-DNN (46%), NCA (42%) and OFDM (41.5%). From the above results it is shown that F-score value in terms of sample size is (2%- 30%) high and F-score value in terms of SNR at −10 dB is (15.5%-46.5%) high than existing works. When the modulation type class is similar to the trained result, the F-score has a better value when imbalanced modulation type predictions exist and then the f-score is lower. In most real-world communication systems, such as MIMO and OFDM, F-score is a better metric to analyse the performance of modulated signals and addressed imbalanced type classification problems. For our proposed AMC^2-pyramid model, no sample size or SNR values are worse than those of the existing methods.

Furthermore, we obtain a higher precision and recall, and the combination of the f-score obtains a better outcome where we predict positive classes for all modulated signals received at the receiver by ≅98.56%. Hence, we say that the proposed work is better for any sample size and any SNR variations.

5) ERROR RATE

On any classification system, the error rate must be lower to prove better performance, which is defined as the proportion

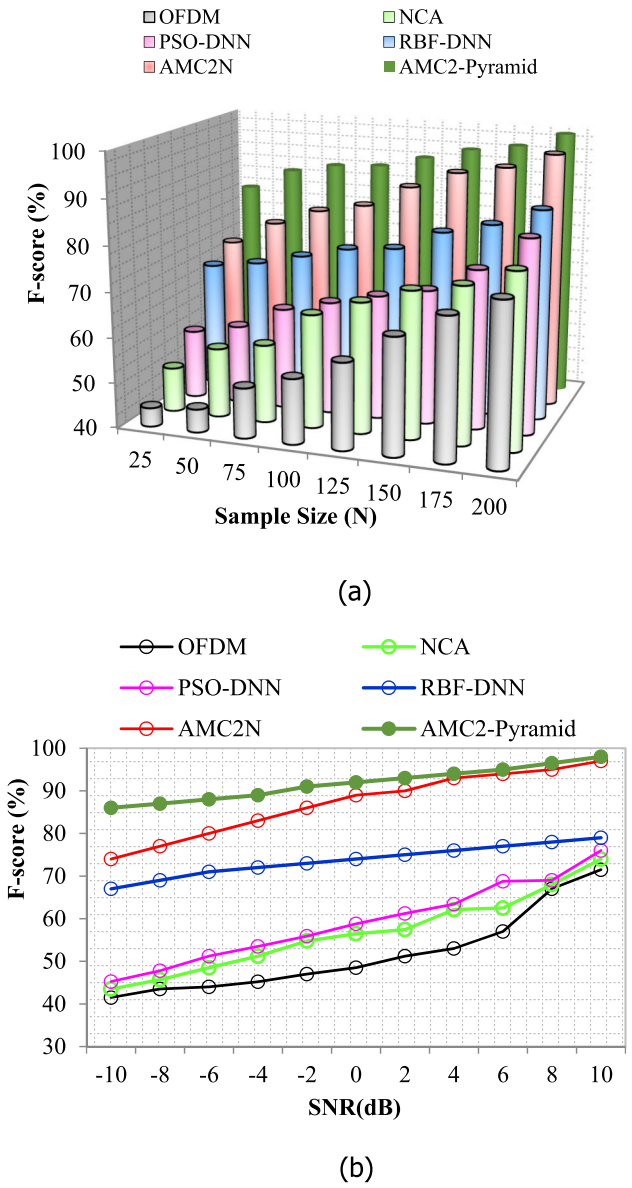

FIGURE 14. (a) Sample size vs. F-score. (b) SNR vs. F-score.

of samples misclassified over the whole set of samples.

$$Error\ Rate = \frac{\#\ of\ Errors}{\#\ of\ Samples} \tag{55}$$

$$= \frac{FN + FP}{TP + TN + FP + FN} \tag{56}$$

When the trained set is much optimistic or imbalanced, the classification error rate is high. However, estimating the total error or loss as once for an average computation is not effective in error rate computation.

Figure 15(a) and (b) shows the performance of the error rate for sample size and SNR, respectively. Given the uncertainty condition prediction in a gated feature pyramid network, the expected loss is estimated. From that analysis, the error rate is reduced much in the proposed work compared

ate computation.

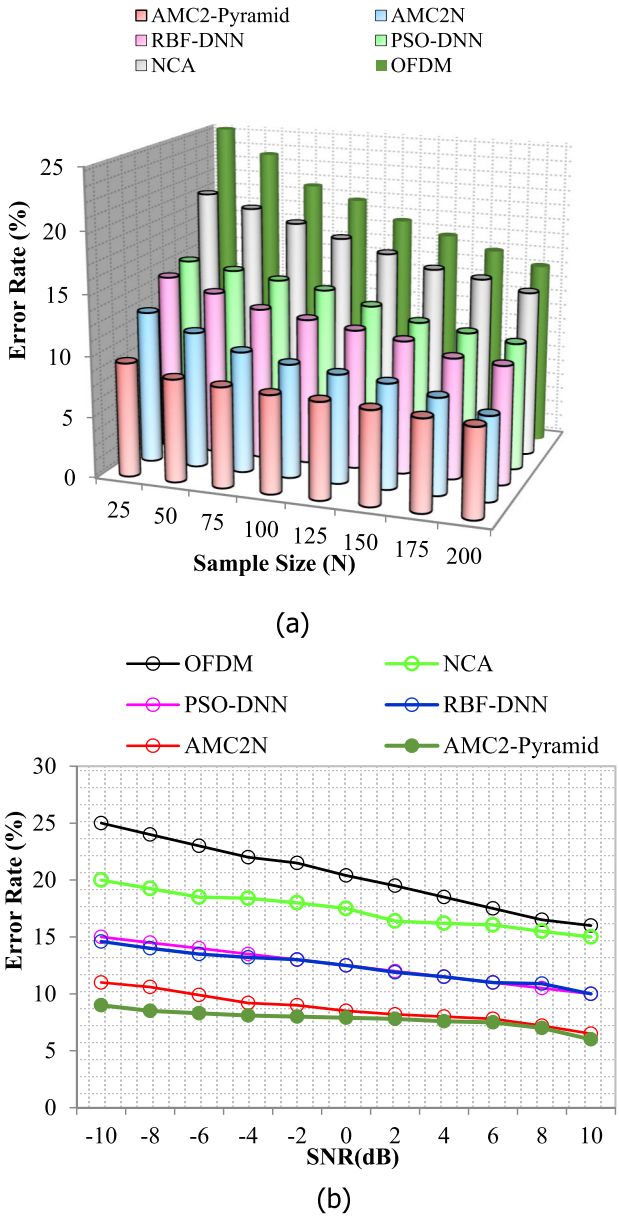

FIGURE 15. (a) Sample size vs. error rate. (b) SNR vs. error rate.

with the previous methods. Furthermore, the proposed work provides an optimistic estimate of the true error probability. Notably, all the previous methods require a number of operations that increase with n order, where a large number of samples utilised in the classification system requires additional excessive computational efforts. To address this issue, a loss of information is computed in training and also clusters similar feature vectors to the sample class, which results in a less error rate. As the number of samples increases, the classification error rate is lesser in the proposed work and is obtained by minimum value in the proposed work when the SNR range is reached from minimum

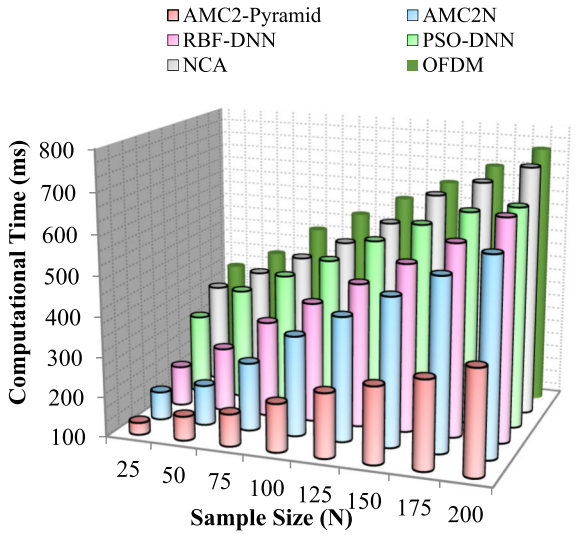

FIGURE 16. Computational time vs. sample size.

to maximum. In 15 (a) the error rate is compared with the sample size which shows the proposed work error rate at samples (N=200) is 8% when compared to the existing works such as AMC2N (8.7%), RBF-DNN (13.8%), PSO-DNN (14.9%), NCA (17.5%) and OFDM (19.5%). In 15 (b) the error rate is compared with SNR at −10 dB is 25% when compared to existing works such as AMC2N (21.5%), RBF-DNN (15.8%), PSO-DNN (14.9%), NCA (12%) and OFDM (8%). From the above results it is shown the error rate in terms of sample size is (0.7%- 11.5%) low and in terms of SNR at −10 dB is (4.5%- 17%) high than existing works.

6) COMPUTATIONAL TIME

Computational time is a time duration, which is measured by increasing the sample size considered in the classification step. The computational time of the proposed work is lesser as compared with the previous methods.

As given in Fig. 16, the proposed work minimises the computational time compared with the previous methods. The proposed work uses fast and lightweight algorithms that require a minimum number of iterations, and also, this work requires minimum training and testing time. In particular, a feature clustering operation reduces the time in classification, and also, preprocessing tasks reduce the computational overhead for modulation type classification. The previous methods consume a considerable amount of computational time to implement preprocessing, feature extraction and classification. In figure.16 proposed work computation time at samples (N=200) is 380 ms as the number of sample size increases which is lesser than the existing works that has computation time of AMC2N (550ms), RBF-DNN (650ms), PSO-DNN (680ms), NCA (720ms) and OFDM (780ms). From the results it is shown that the computational time in terms of sample size is (170ms-400ms) lower than the existing works. The existing method requires more than 100 ms for 10 signals and is 50% greater than the proposed work because the proposed work obtains 52 ms for 10 signals.

C. RESULTS AND DISCUSSION

The proposed AMC-pyramid model uses an intelligent gated feature network for feature engineering, and thus, the proposed work obtains the best performance. Furthermore, this model sacrifices the complexity, and multiple features are trained over time, that is, the spectral domain, transformation domain, constellation and statistical features. Therefore, the PCC, that is, the overall classification accuracy is reached to the peak and also outperforms for different cases as low to high sample size and low to high SNR range. Then, based on the result of computation time in the simulation, the pyramid-based system model has a fast velocity because determining the optimum results in a minimum number of iterations.

VII. CONCLUSION AND FUTURE WORK

Considering the increasing rate of applications in wireless communication systems, AMC has been developed dramatically. On the receiver side, an accurate and fast modulation type detection method is in high demand to demodulate the transmitter signal. Obtaining a high modulation classification accuracy in a multi-carrier system with low to high SNR variations is a more hindering task. In this study, we aim to develop the AMC^2-pyramid system, which aims to classify modulated signals as any of the following classes: 16QAM, 32QAM, 64QAM, 128QAM, QPSK, BPSK, DPSK, ASK and FSK. The proposed AMC^2-pyramid model consists of four operations to obtain the above performance, namely, pre-processing, feature extraction, feature clustering and classification. Firstly, pre-processing is executed first to improve the quality of the signal. If the received signal consists of good quality, then the feature extraction step is performed. Quality testing for all kinds of signals introduces complexity, and thus, the signal quality enrichment step is presented after the quality assessment. In the signal quality augmentation process, noise is eliminated, and quantisation, equalisation and CFO compensation are performed. Subsequently, multiple features are extracted using GFPN, which extract spectrum, statistical, transform and constellation features. Then, the twin function-based human mental search algorithm is used to form clusters by similar feature vectors. For modulation type classification, the MDNC^2 algorithm is used, and the final classification is implemented by the IQL, in which the demodulation error is updated in reward computation. The obtained results of all performance metrics are compared with the several existing methods, such as OFDM, RBF-DNN, NCA and PSO-DNN. Finally, the proposed work outperforms the existing methods in all performance metrics in terms of high precision, recall, low error rate and low computation time. The main advantage of this work is efficient

classification accuracy of multi carrier modulation with low error rate. Only limited number of modulation scheme is classified. In future more signal modulation types will be considered and will be classified using enhanced methods.

REFERENCES

[1] D. H. Al-Nuaimi, M. F. Akbar, L. B. Salman, I. S. Z. Abidin, and N. A. M. Isa, "AMC2N: Automatic modulation classification using feature clustering-based two-lane capsule networks," *Electronics*, vol. 10, no. 1, p. 76, Jan. 2021.

[2] A. P. Hermawan, R. R. Ginanjar, D.-S. Kim, and J.-M. Lee, "CNN-based automatic modulation classification for beyond 5G communications," *IEEE Commun. Lett.*, vol. 24, no. 5, pp. 1038–1041, May 2020.

[3] Y. Kumar, M. Sheoran, G. Jajoo, and S. K. Yadav, "Automatic modulation classification based on constellation density using deep learning," *IEEE Commun. Lett.*, vol. 24, no. 6, pp. 1275–1278, Jun. 2020.

[4] M. A. Pinto-Orellana and H. L. Hammer, "Dyadic aggregated autoregressive model (DASAR) for automatic modulation classification," *IEEE Access*, vol. 8, pp. 156096–156103, 2020.

[5] M. Zhang, Z. Yu, H. Wang, H. Qin, W. Zhao, and Y. Liu, "Automatic digital modulation classification based on curriculum learning," *Appl. Sci.*, vol. 9, no. 10, p. 2171, May 2019.

[6] N. Daldal, Z. Cömert, and K. Polat, "Automatic determination of digital modulation types with different noises using convolutional neural network based on time–frequency information," *Appl. Soft Comput.*, vol. 86, Jan. 2020, Art. no. 105834.

[7] P. Wu, B. Sun, S. Su, J. Wei, J. Zhao, and X. Wen, "Automatic modulation classification based on deep learning for software-defined radio," *Math. Problems Eng.*, vol. 2020, Nov. 2020, Art. no. 2678310.

[8] R. Zhang, Z. Yin, Z. Wu, and S. Zhou, "A novel automatic modulation classification method using attention mechanism and hybrid parallel neural network," *Appl. Sci.*, vol. 11, no. 3, p. 1327, Feb. 2021.

[9] Y. Zhang, D. Liu, J. Liu, Y. Xian, and X. Wang, "Improved deep neural network for OFDM signal recognition using hybrid grey wolf optimization," *IEEE Access*, vol. 8, pp. 133622–133632, 2020.

[10] J. C. Clement, N. Indira, P. Vijayakumar, and R. Nandakumar, "Deep learning based modulation classification for 5G and beyond wireless systems," *Peer-Peer Netw. Appl.*, vol. 14, no. 1, pp. 319–332, Jan. 2021.

[11] A. Güner, Ö. F. Alçin, and A. Şengür, "Automatic digital modulation classification using extreme learning machine with local binary pattern histogram features," *Measurement*, vol. 145, pp. 214–225, Oct. 2019.

[12] J. P. Mouton, M. Ferreira, and A. S. J. Helberg, "A comparison of clustering algorithms for automatic modulation classification," *Expert Syst. Appl.*, vol. 151, Aug. 2020, Art. no. 113317.

[13] Y. Wei, S. Fang, and X. Wang, "Automatic modulation classification of digital communication signals using SVM based on hybrid features, cyclostationary, and information entropy," *Entropy*, vol. 21, no. 8, p. 745, Jul. 2019.

[14] S. H. Lee, K.-Y. Kim, and Y. Shin, "Effective feature selection method for deep learning-based automatic modulation classification scheme using higher-order statistics," *Appl. Sci.*, vol. 10, no. 2, p. 588, Jan. 2020.

[15] V. G. Sannikov and V. P. Volchkov, "Multi-carrier modulations digital modem with the narrow-band optimal signals and high spectral-energy efficiency," in *Proc. 22th Int. Conf. Digit. Signal Process. Appl. (DSPA)*, Mar. 2020, pp. 1–5.

[16] S. Zhou, Z. Yin, Z. Wu, Y. Chen, N. Zhao, and Z. Yang, "A robust modulation classification method using convolutional neural networks," *EURASIP J. Adv. Signal Process.*, vol. 2019, no. 1, pp. 1–15, Dec. 2019.

[17] X. Shang, H. Hu, X. Li, T. Xu, and T. Zhou, "Dive into deep learning based automatic modulation classification: A disentangled approach," *IEEE Access*, vol. 8, pp. 113271–113284, 2020.

[18] H. Ma, G. Xu, H. Meng, M. Wang, S. Yang, R. Wu, and W. Wang, "Cross model deep learning scheme for automatic modulation classification," *IEEE Access*, vol. 8, pp. 78923–78931, 2020.

[19] X. Zhang, J. Sun, and X. Zhang, "Automatic modulation classification based on novel feature extraction algorithms," *IEEE Access*, vol. 8, pp. 16362–16371, 2020.

[20] A. O. A. Salam, R. E. Sheriff, Y.-F. Hu, S. R. Al-Araji, and K. Mezher, "Automatic modulation classification using interacting multiple model Kalman filter for channel estimation," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 8928–8939, Sep. 2019.

[21] W. Li, Z. Dou, L. Qi, and C. Shi, "Wavelet transform based modulation classification for 5G and UAV communication in multipath fading channel," *Phys. Commun.*, vol. 34, pp. 272–282, Jun. 2019.

[22] A. Hussain, F. Sohail, S. Alam, S. A. Ghauri, and I. M. Qureshi, "Classification of M-QAM and M-PSK signals using genetic programming (GP)," *Neural Comput. Appl.*, vol. 31, pp. 6141–6149, Mar. 2018.

[23] A. K. Ali and E. Erçelebi, "An M-QAM signal modulation recognition algorithm in AWGN channel," *Sci. Program.*, vol. 2019, May 2019, Art. no. 6752694.

[24] Y. Liang, X. Xiang, Y. Sun, X. Da, C. Li, and L. Yin, "Novel modulation recognition for WFRFT-based system using 4th-order cumulants," *IEEE Access*, vol. 7, pp. 86018–86025, 2019.

[25] Z. Xing and Y. Gao, "A modulation classification algorithm for multipath signals based on cepstrum," *IEEE Trans. Instrum. Meas.*, vol. 69, no. 7, pp. 4742–4752, Jul. 2020.

[26] S. S. Mathad and C. Vijaya, "Revised architecture for automatic modulation recognition," *Int. J. Inf. Technol.*, vol. 12, no. 2, pp. 605–610, Jun. 2020.

[27] L. Pu, H.-C. Wu, K. Yan, Z. Gao, X. Wang, and W. Xiang, "Novel three-hierarchy multiple-tag-recognition technique for next generation RFID systems," *IEEE Trans. Wireless Commun.*, vol. 19, no. 2, pp. 1237–1249, Feb. 2020.

[28] Z. Zhang, C. Wang, C. Gan, S. Sun, and M. Wang, "Automatic modulation classification using convolutional neural network with features fusion of SPWVD and BJD," *IEEE Trans. Signal Inf. Process. Netw.*, vol. 5, no. 3, pp. 469–478, Sep. 2019.

[29] M. Kulin, T. Kazaz, I. Moerman, and E. De Poorter, "End-to-end learning from spectrum data: A deep learning approach for wireless signal identification in spectrum monitoring applications," *IEEE Access*, vol. 6, pp. 18484–18501, 2018.

[30] K. Jiang, J. Zhang, H. Wu, A. Wang, and Y. Iwahori, "A novel digital modulation recognition algorithm based on deep convolutional neural network," *Appl. Sci.*, vol. 10, no. 3, p. 1166, Feb. 2020.

[31] Y. Wang, J. Wang, W. Zhang, J. Yang, and G. Gui, "Deep learning-based cooperative automatic modulation classification method for MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 69, no. 4, pp. 4575–4579, Apr. 2020.

[32] Y. Wang, J. Yang, M. Liu, and G. Gui, "LightAMC: Lightweight automatic modulation classification using deep learning and compressive sensing," *IEEE Trans. Veh. Technol.*, vol. 49, no. 3, pp. 3491–3495, Mar. 2020.

[33] H. Zhang, Y. Wang, L. Xu, T. Aaron Gulliver, and C. Cao, "Automatic modulation classification using a deep multi-stream neural network," *IEEE Access*, vol. 8, pp. 43888–43897, 2020.

[34] S. Chen, Y. Zhang, Z. He, J. Nie, and W. Zhang, "A novel attention cooperative framework for automatic modulation recognition," *IEEE Access*, vol. 8, pp. 15673–15686, 2020.

[35] N. Daldal, K. Polat, and Y. Guo, "Classification of multi-carrier digital modulation signals using NCM clustering based feature-weighting method," *Comput. Ind.*, vol. 109, pp. 45–58, Aug. 2019.

[36] W. Shi, D. Liu, X. Cheng, Y. Li, and Y. Zhao, "Particle swarm optimization-based deep neural network for digital modulation recognition," *IEEE Access*, vol. 7, pp. 104591–104600, 2019.

[37] B. Dehri, M. Besseghier, A. B. Djebbar, and I. Dayoub, "Blind digital modulation classification for STBC-OFDM system in presence of CFO and channels estimation errors," *IET Commun.*, vol. 13, no. 17, pp. 2827–2833, Oct. 2019.

[38] S. Duan, K. Chen, X. Yu, and M. Qian, "Automatic multicarrier waveform classification via PCA and convolutional neural networks," *IEEE Access*, vol. 6, pp. 51365–51373, 2018.

[39] M. H. Shah and X. Dang, "Low-complexity deep learning and RBFN architectures for modulation classification of space-time block-code (STBC)-MIMO systems," *Digit. Signal Process.*, vol. 99, Apr. 2020, Art. no. 102656.

[40] Y. Zeng, M. Zhang, F. Han, Y. Gong, and J. Zhang, "Spectrum analysis and convolutional neural network for automatic modulation recognition," *IEEE Wireless Commun. Lett.*, vol. 8, no. 3, pp. 929–932, Jun. 2019.

[41] P. Ghasemzadeh, S. Banerjee, M. Hempel, and H. Sharif, "Accuracy analysis of feature-based automatic modulation classification with blind modulation detection," in *Proc. Int. Conf. Comput., Netw. Commun. (ICNC)*, Feb. 2019, pp. 1000–1004.

[42] C. Bai, R. Zhang, Z. Xu, R. Cheng, B. Jin, and J. Chen, "L1-norm-based kernel entropy components," *Pattern Recognit.*, vol. 96, Dec. 2019, Art. no. 106990.

[43] G. Wang, J. Zhu, and Z. Xu, "Asymptotically optimal one-bit quantizer design for weak-signal detection in generalized Gaussian noise and lossy binary communication channel," 2018, *arXiv:1805.03024*. [Online]. Available: http://arxiv.org/abs/1805.03024
[44] Y. Liu, X. Zhu, E. G. Lim, Y. Jiang, and Y. Huang, "Iterative semi-blind CFO estimation, Si cancelation and signal detection for full-duplex systems," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2018, pp. 1–7.
[45] Z. Shen, H. Shi, J. Yu, H. Phan, R. Feris, L. Cao, D. Liu, X. Wang, T. Huang, and M. Savvides, "Improving object detection from scratch via gated feature reuse," in *Proc. BMVC*, 2019, pp. 1–12.
[46] X. Xie, Q. Liao, L. Ma, and X. Jin, "Gated feature pyramid network for object detection," in *Proc. PRCV*, 2018, pp. 199–208.
[47] D. H. Al-Nuaimi, I. A. Hashim, I. S. Z. Abidin, L. B. Salman, and N. A. Mat Isa, "Performance of feature-based techniques for automatic digital modulation recognition and classification—A review," *Electronics*, vol. 8, no. 12, p. 1407, Nov. 2019.
[48] S. J. Mousavirad, H. Ebrahimpour-Komleh, and G. Schaefer, "Effective image clustering based on human mental search," *Appl. Soft Comput.*, vol. 78, pp. 209–220, May 2019.

DHAMYAA HUSAM AL-NUAIMI received the bachelor's and M.Sc. degrees in electrical engineering from the University of Technology, Baghdad, Iraq, in 2003 and 2006, respectively, and the Ph.D. degree in digital signal processing from the School of Electrical and Electronic Engineering, Universiti Sains Malaysia (USM), Malaysia. She is currently a Lecturer with the Communication Engineering Department, Al-Mansour University College, Baghdad.

NOR ASHIDI MAT ISA received the B.Eng. degree (Hons.) in electrical and electronic engineering from Universiti Sains Malaysia (USM), in 1999, and the Ph.D. degree in electronic engineering (majoring in image processing and artificial neural network), in 2003. He is currently a Professor and the Deputy Dean (Academic, Career, and International) at the School of Electrical and Electronic Engineering, USM. His research interests include intelligent systems, image processing, neural networks, biomedical engineering, and intelligent diagnostic systems and algorithms.

MUHAMMAD FIRDAUS AKBAR (Member, IEEE) received the B.Sc. degree in communication engineering from the International Islamic University Malaysia (IIUM), Malaysia, in 2010, and the M.Sc. and Ph.D. degrees from The University of Manchester, Manchester, U.K., in 2012 and 2018, respectively. From 2010 to 2011, he was with Motorola Solutions, Penang, Malaysia, as a Research and Development Engineer. From 2012 to 2014, he was an Electrical Engineer with Usains Infotech Sdn Bhd, Penang. He is currently a Senior Lecturer at Universiti Sains Malaysia (USM). His current research interests include electromagnetics, microwave nondestructive testing, and microwave sensor and imaging.

INTAN SORFINA ZAINAL ABIDIN received the M.Eng. degree in electronics engineering from the University of Surrey, Guildford, U.K., in 2008, and the Ph.D. degree from the 5G Innovation Centre (5GIC), Institute of Communication System (ICS), University of Surrey, in 2017. Then, she worked with Motorola Solutions (M), Penang, Malaysia, as an Electrical Engineer, for two years, before moving to Celestica (M), Kulim, Malaysia, to work as a Product Engineer. In 2013, she returned to the University of Surrey for her Ph.D. degree under 5GIC, ICS. She proceeds to work as a Senior Lecturer at the School of Electrical and Electronic Engineering, Universiti Sains Malaysia (USM), specializing in electronics communication system, antenna, channel propagation, RF, microwave engineering, and MIMO system.