

Received 13 April 2024, accepted 8 May 2024, date of publication 13 May 2024, date of current version 21 May 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3400448

## RESEARCH ARTICLE

# Automatic Modulation Classification for MIMO System Based on the Mutual Information Feature Extraction

N. USSIPOV<sup>1</sup>, S. AKHTANOV<sup>1</sup>, Z. ZHANABAEV<sup>1</sup>, D. TURLYKOZHAYEVA<sup>1</sup>,  
B. KARIBAYEV<sup>2</sup>, T. NAMAZBAYEV<sup>1</sup>, D. ALMEN<sup>1</sup>, A. AKHMETALI<sup>1</sup>,  
AND XIAO TANG<sup>3</sup>, (Member, IEEE)

<sup>1</sup>Faculty of Physics and Technology, Al-Farabi Kazakh National University, 050040 Almaty, Kazakhstan

<sup>2</sup>Department of Telecommunication Engineering, Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev, 050013 Almaty, Kazakhstan

<sup>3</sup>School of Electronics and Information, Northwestern Polytechnical University, Xi'an 710071, China

Corresponding author: D. Turlykozhayeva (dana.turlykozhayeva@kaznu.edu.kz)

This work was supported by the Committee of Science of the Ministry of Science and Higher Education of the Republic of Kazakhstan under Grant AP14872061.

**ABSTRACT** Automatic Modulation Classification (AMC) is an essential technology that is widely applied into various communications scenarios. In recent years, many Machine Learning and Deep-Learning methods have been introduced into AMC, and a lot of them apply different approaches to eliminate interference in complex Multiple-Input and Multiple-Output (MIMO) signals and improve classification performance. However, in practical communication systems, the perfect elimination of MIMO signal interference is impossible, and therefore classification performance suffers. In this paper, we propose a new AMC algorithm for MIMO system based on mutual information (MI) features extraction, which does not require a large amount of training data and the elimination of MIMO signal interference. In this approach, features based on mutual information are extracted using In-Phase and Quadrature (IQ) constellation diagrams of MIMO signals, which have not been explored previously. Our method can be effective since mutual information considers the interdependencies among variables and measures how much information about one variable reduces uncertainty about another, providing a valuable perspective for extracting higher-level and interesting features from the data. The effectiveness of our method is evaluated on several model and real-world datasets, and its applicability is proven.

**INDEX TERMS** Automatic modulation classification, classifier, feature extraction, mutual information, entropy, complex MIMO signals.

## I. INTRODUCTION

Automatic Modulation Classification (AMC) is an important technology that provides modulation information of the incoming radio signals and plays a significant role in practical civilian and military applications such as cognitive radio, interference recognition, spectrum management, data encryption, etc. AMC has gained considerable research interest in recent years since it aims to automatically identify the modulation types of wireless communications signals

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

without prior information [1], [2], [3], [4], [5]. Among the various modulation types, quadrature amplitude modulation (QAM) and phase-shift keying (PSK) find widespread application in communication systems. Additionally, when compared to other modulation types, classifying higher-order QAM poses a greater challenge [6]. Hence, this study specifically concentrates on high order 16QAM and 64QAM modulations.

Many modulation classification algorithms have been proposed for both single-input and single-output (SISO) systems and MIMO systems and these methods are usually classified into two major groups: likelihood-based (LB) and

feature-based (FB). LB methods are based on the likelihood function and are considered optimal but generally complex to implement [7], [8]. FB methods depend on feature extraction and classifier design and compared to LB methods are more applicable in practical implementations and more popular in AMC due to the ease of implementation [9], [10]. Various types of features, such as instantaneous time-domain features, transform-domain features, and statistical features have been investigated and used in AMC algorithms [11], [12]. After extracting the features, the classifier is used to determine the class of modulated signals. Until now, various types of classifiers have been used in the classification stage of AMC, including artificial neural networks (ANN), k-nearest neighbors (KNN), and support vector machines (SVM) [13].

Recently, numerous machine learning and deep learning-based AMC methods [2], [14], [15], [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27], [28], [29] have been developed for the classification of SISO and MIMO signals. These methods, as demonstrated in recent surveys [30], [31], [32], [33], [34], [35], show promise for achieving near-optimal performance with acceptable computational complexity. Some of these methods use various technologies to eliminate interference in complex MIMO signals and enhance classification performance. For instance, in [26] and [27] authors propose a convolutional neural network-based zero-forcing equalization AMC method for MIMO systems. However, it is worth noting that the perfect elimination of MIMO signal interference cannot be achieved in real communication systems, and therefore the classification performance is decreased.

In addition, authors in [36], [37], [38], and [39] have proposed AMC algorithms for the classification of communication signals based on information entropy. Entropy, which measures the uncertainty of signal distribution and represents the complexity degree of the signal, plays a key role in these methods. Although these methods are competitive among other approaches, feature extraction using mutual information (MI) can be preferable over Shannon entropy, since MI considers the dependencies between two or more variables. Unlike entropy, which primarily focuses on the uncertainty of individual variables, MI measures how much information about one variable reduces uncertainty about another [40], [41], [42]. This nuanced perspective is particularly valuable for extracting higher-level and interesting features from the data.

In this work, we propose a new AMC algorithm designed for MIMO system based on the mutual information feature extraction. Our algorithm stands out by specifically extracting features using IQ constellation diagrams of MIMO signals, a unique approach that has not been explored before. Furthermore, this method does not require the elimination of MIMO signal interference and a large amount of training data.

The rest of the paper is organized as follows. System model, data set generation and experimental setup are introduced in section II. In Section III, MI background, the proposed AMC method, and SVM classifier are proposed.

Section IV contains simulation results that demonstrate the correctness and validity of our method. Finally, conclusions are given in Section V.

## II. MIMO SYSTEM MODEL, DATA SET GENERATION AND EXPERIMENTAL SETUP

### A. MIMO SYSTEM MODEL

In this paper, we consider a MIMO system that consists of  $N_s$  transmitting antennas and  $N_r$  receiving antennas ( $N_r \geq N_s$ ). Assuming that the MIMO channel is an uncorrelated Rayleigh flat-fading and time-invariant channel [27], the received signal at the  $n$ -th sampling moment can be expressed as:

$$r(n) = Hs(n) + g(n), \quad (1)$$

where  $H$  denotes  $N_r \times N_s$  dimensional MIMO channel matrix, and it follows circular symmetric complex normal distribution;  $r(n) = [r_1(n), r_{N_r}(n)]^T$  is the received signal vector with a size  $N_r \times 1$ ;  $s(n) = [s_1(n), s_{N_t}(n)]^T$  is the transmitted signal vector with a size  $N_t \times 1$ ; and  $g(n) = [g_1(n), g_{N_r}(n)]^T$  is the additive white Gaussian noise vector with a size  $N_s \times 1$  [2], [3], [23].

### B. DATA SET GENERATION

Here we consider the system model of the complex multi-antenna, and the data set generating process is presented in Fig. 1. Particularly, random data are modulated using various modulation types, such as binary phase shift keying (BPSK), quadrature phase shift keying (QPSK), 8-phase shift keying (8PSK), 16-quadrature amplitude modulation (16QAM), 64-quadrature amplitude modulation (64QAM). The modulation signal vector, represented as  $S$ , with size  $1 \times N$  (where  $N$  is the number of symbols, and  $N = 256, 512, 1024$ ). To ensure a fair comparison,  $S$  is normalized with unit power, i. e.,  $\|S\|_2^2 = 1$ . After normalization,  $S$  is reshaped into a  $N_s \times N/N_s$  matrix, represented as  $[s_1, s_2, \dots, s_{N_s}]$ , here  $s_i = [s_i(1), s_i(2), \dots, s_i(N/N_s)]$ ,  $i \in [1, N_s]$ ; - is the transmitted signal vector at the  $i$ -th antenna.

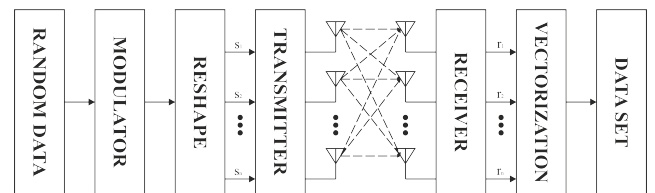


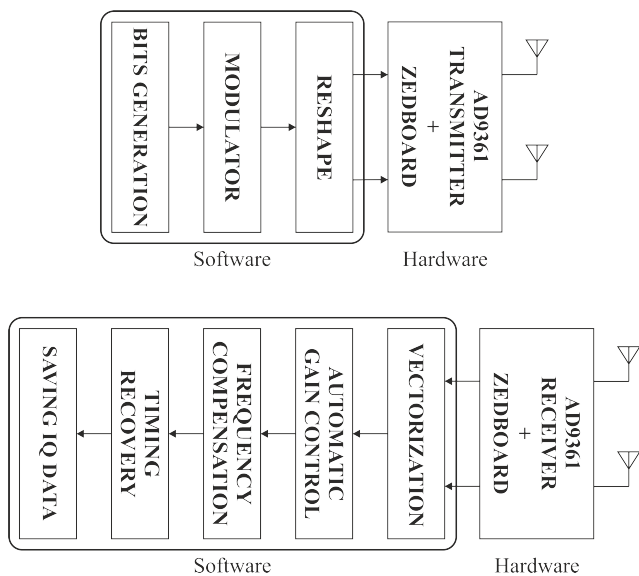
FIGURE 1. The process of dataset generation.

The received signal vector at the  $j$ -th receive antenna passing through the MIMO channel, is denoted as  $[r_1, r_2, \dots, r_{N_s}]$ , with size  $N_r \times N/N_s$ ,  $r_j = [r_j(1), r_j(2), \dots, r_j(N/N_s)]$ ,  $j \in [1, N_r]$ . Next, the received signal matrix can be vectorized into a  $1 \times N$  vector  $r$ . Test samples and training data are retrieved from  $r$ . Specifically, the real and imaginary parts of  $r$ :  $\mathcal{R}(r)$  and  $\mathcal{I}(r)$  are separated and combined into a  $2 \times N$  matrix  $[\mathcal{R}(r) \text{ and } \mathcal{I}(r)]$ , which is a sample for training and testing. It should be noted that 6,000 samples

are prepared for training and 4,000 samples for testing for each SNR value [24], [26].

**C. EXPERIMENTAL SETUP**

To evaluate the effectiveness and reliability of the proposed AMC method in real communication scenarios, we conducted a series of experiments using an integrated software and hardware complex for data collection of IQ constellation diagrams. The software component of this complex is implemented in the Matlab+Simulink program, while the hardware component comprises the Zedboard+AD9361 (Fig. 2). The Zedboard, connected to the PC via an Ethernet cable, provides communication between the software component and the RF module AD9361. The AD9361 RF module is compulsory as it provides wireless communication between receiver and transmitter.

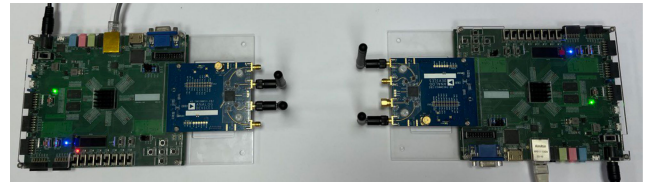


**FIGURE 2. Transmitter and receiver block diagram.**

The transmitting unit of software part consists of a random bit generator, a modulator, a reshape demultiplexer device, while the hardware part comprises AD9361 with 2 transmitting antennas.

The receiving unit of the software part accomplishes multiplexing for vectorization and signal processing for IQ constellation diagrams. Signal processing, in turn, includes tasks such as automatic gain control (AGC), frequency compensation, and timing recovery. The hardware part of this receiving unit consists of AD9361 with 2 antennas (Fig 2). IQ constellation diagram values were saved and used for testing the proposed AMC algorithm. Fig. 3 below illustrates the hardware part of the experimental setup designed for receiving and transmitting signals in MIMO channels.

The configuration of the AD9361 RF module was performed using Simulink, and the following parameters were selected: a frequency of 2.4 GHz, a baseband sample rate of 520.841 kHz. The transmitter and receiver gain parameters were adjusted to control the Signal-to-Noise Ratio (SNR)

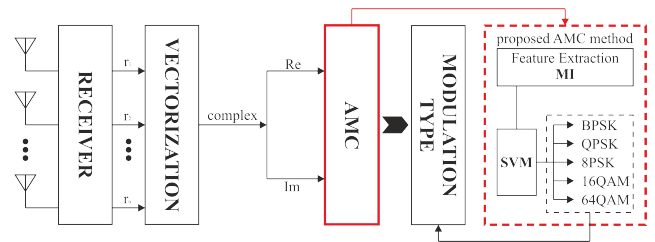


**FIGURE 3. The hardware part of the experimental setup.**

value. The experimental results confirm our theoretical data and serve as a qualitative addition to this research.

**III. THE PROPOSED AMC METHOD BASED ON MUTUAL INFORMATION FEATURE EXTRACTION**

In this section, we describe the structure of the proposed AMC method, which is shown in Fig 4. The method consists of two main parts: feature extraction based on mutual information  $I(X; Y)$  and SVM classifier. To make it easier to understand, we introduce the section from three aspects: MI background, detailed explanation of the proposed method and SVM classifier.



**FIGURE 4. The structure of the proposed AMC method.**

**A. MI BACKGROUND**

In probability theory and information theory, mutual information is a basic concept that measures the statistical dependence between two random variables. Specifically, it represents the amount of information learned about one random variable by observing another. Mutual information is closely related to the entropy because knowing information about one random variable can decrease the level of uncertainty of another variable. Therefore, a high value of  $I(X; Y)$  shows a substantial reduction of uncertainty, while a low value signifies a minor reduction. If the mutual information between two random variables is zero, it indicates that they are independent of each other [40], [41].

The mutual information between two discrete random variables  $X$  and  $Y$  is defined as a double sum:

$$I(X; Y) = \sum_{j \in Y} \sum_{i \in X} P_{(i,j)}(x, y) \log_2 \left( \frac{P_{(i,j)}(x, y)}{P_i(x)P_j(y)} \right), \quad (2)$$

where  $N$  is the number of possible values for  $X$  and  $M$  is the number of possible values for  $Y$ ,  $P_{(i,j)}(x, y)$  is the joint probability density function,  $P_i(x)$  and  $P_j(y)$  denote the marginal probability density function of  $X$  and  $Y$  [40], [41], [42], [43], [44].

As explained before, there is a strong relationship between mutual information and entropy. The following formula illustrates this relation:

$$I(X; Y) = H(X) - H(X/Y), \quad (3)$$

where  $H(X)$ ,  $H(X/Y)$  are the marginal and conditional entropies. To comprehend the meaning of  $I(X; Y)$ , it is necessary to explain both entropy and conditional entropy. Entropy ( $H$ ) quantifies the degree of expected uncertainty of a random variable, and it can be expressed as follows [43]:

$$H(X) = - \sum_i P_i(x) \log_2 P_i(x), \quad (4)$$

where  $P_i$  is the probability of event  $x_i$ . The conditional entropy quantifies the amount of uncertainty that a random variable  $X$  has when the value of  $Y$  is given, and it can be defined using the following formula [44]:

$$H(X/Y) = - \sum_{i=1}^N \sum_{j=1}^M P_{(i,j)}(x, y) \log_2(P_{(i,j)}(x/y)), \quad (5)$$

where  $N$  is the number of possible values for  $X$  and  $M$  is the number of possible values for  $Y$ ,  $P_{(i,j)}(x, y)$  denotes the joint probability, and  $P_{(i,j)}(x/y)$  denotes the conditional probability.

### B. THE PROPOSED METHOD

The proposed method involves the extraction of features based on mutual information (MI) using a histogram of IQ diagram. The process begins with constructing an IQ diagram that depicts the complex values of modulated signals. The in-phase component of the received signal is chosen as variable  $X$ , while the quadrature component is denoted as variable  $Y$ . Then, to calculate MI we built histogram (Fig. 5) of  $XY$  plane divided into bins of equal size ( $\Delta X \times \Delta Y$ ) with coordinates  $i, j$ . By counting the number  $k_{(i,j)}$  of samples in bin  $(i, j)$ , MI can be calculated using the following formula:

$$I(X; Y) = \sum_{j=1}^{k_y} \sum_{i=1}^{k_x} \left( \frac{k_{ij}}{N} \right) \log \left( \frac{k_{ij}N}{k_i k_j} \right) \quad (6)$$

To implement the proposed algorithm, the following steps must be performed:

1. Construct IQ diagram.
2. Build histogram of two-dimensional IQ diagram.
3. MI calculation according to equation (6).
4. In the final step, apply an SVM classifier to identify the type of modulation.

In Fig. 5, the histogram of the IQ diagrams shows significant differences corresponding to the type of modulation. The point density decreases as the number of symbols increases.

In the literature, several rules have been proposed to select the number of bins ( $b$ ), such as the choice of the square root (where  $b = \sqrt{N}$ ), the Sturges formula (where  $b = 1 + \log_2 N$ ), the Rice rule (where  $b = 2N^{1/3}$ ) [45]. The comparison of these rules is given in the results.

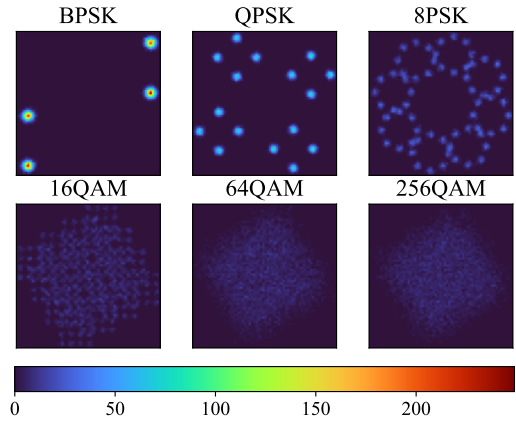


FIGURE 5. Histogram of IQ diagrams of four modulation schemes at SNR=25.

### C. SVM CLASSIFIER

Various types of classifiers have been used in the classification stage of AMC until those days. In this work, we evaluated the performance of selected classifiers based on their Probability Correct Classification (Table 1) and chose the SVM classifier.

Support Vector Machine (SVM) is a supervised machine learning algorithm developed by Vapnik and Chervonenkis based on statistical learning theory [46]. It is useful for solving small-sample learning problems and has been successfully applied in various fields, including pattern recognition, regression analysis, and density estimation. SVM using kernel functions maps the input data into a high-dimensional feature space, where a nonlinear classification problem can be treated as a linear one.

By applying the rule of structural risk minimization, SVM aims to not only reduce classification errors but also enhance the generalization ability of classifiers. Compared to traditional Neural Network algorithms, SVM is advantageous due to its simple structure, high success rates, better generalization, and small-sample-problem-solving ability. Detailed mathematical theories of SVM are available in references [47] and [48].

### IV. RESULTS AND DISCUSSION

This section presents simulation results to evaluate the performance of the suggested modulation classification scheme. We analyze a MIMO communication system with  $N_s = 2$  transmitting antennas and  $N_r = 2$  receiving antennas. A channel with relay fading and AWGN noise is assumed. Six modulation schemes are considered: BPSK, QPSK, 8PSK, 16QAM, 64QAM. The extraction of signal features using mutual information is calculated by averaging 100 simulations using the Monte Carlo method (Fig. 6).

According to Fig. 6, as the number of symbols on the constellation increases, the distance between the points increases, which leads to a decrease in the density of points on the constellation and, consequently, to a decrease in mutual information. To calculate the probability of correct



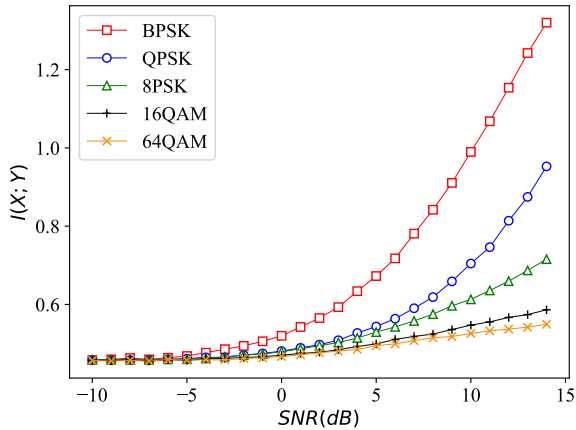


FIGURE 6. Feature extraction based on mutual information under different SNR.

classification, the following formula is used:

$$P_{cc} = N_{cc}/N_{rs}, \tag{7}$$

where  $N_{cc}$  is the number of correctly classified signals,  $N_{rs}$  is the total number of received signals. Fig. 7 below shows a comparison of the existing rules for selecting bins.

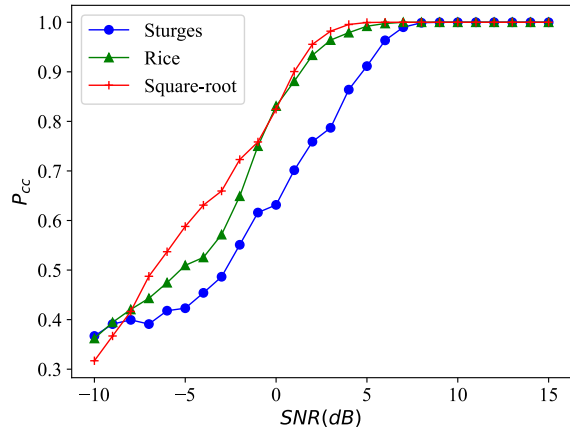


FIGURE 7. Comparison of empirical rules for selecting bins (N=1024).

According to Fig. 7, the results of the square root and Rice rule demonstrate better performance compared to the Sturges rule. This is because, from the perspective of the integral root-mean-square error, a larger number of bins enhances the accuracy of the density estimate.

Also, for the proposed AMC algorithm, we assessed the average probability of correct classification across different classifiers at SNR from  $-10$  to  $15$ , and we also examined the simulation time of these classifiers (Table 1). Simulations were conducted using a computer equipped with an 11th Gen Intel(R) Core (TM) i5-11400F processor operating at 2.60 GHz and 16.0 GB of RAM. The probability of correct classification of the classifier is generally considered the most crucial indicator for evaluating a feature extraction algorithm. We use five classifiers, including KNN, SVM, Adaboost, Random Forest (RF) and MLP to classify the original feature

set extracted by our proposed AMC method. These classifiers are evaluated using different methods for determining the number of bins in a histogram, including Sturges, Rice, and Square Root. The specific parameter for the KNN classifier was set as follows: the nearest neighbor number ( $k$ ) was set to 5. For the SVM classifier, the kernel function was set as RBF kernel function. The Adaboost classifier’s parameters were set with a depth of 10, a learning rate of 0.1, and 10 iterations. The Random Forest classifier had a depth of 9, a learning rate of 0.1, and 10 iterations. The MLP classifier was configured with a learning rate of 0.1, and 10 iterations. According to the data set, the simulation time and average probability of correct classification of different classifiers are calculated.

TABLE 1. Performance comparison of classifiers for the proposed AMC algorithm.

	Sturges		Rice		Square-root	
	$P_{cc}$	t (s)	$P_{cc}$	t (s)	$P_{cc}$	t (s)
SVM	79.21%	0.56	84.79%	0.48	85.24%	0.41
Adaboost	62.60%	0.25	66.29%	0.11	68.72%	0.25
MLP	78.92%	0.37	83.89%	0.34	84.37%	0.62
RF	78.84%	0.09	81.41%	0.19	82.31%	0.19
KNN	80.25%	0.10	83.85%	0.09	84.06%	0.09

Table 1 reveals that when applying the square root rule, SVM, MLP, and KNN classifiers achieve the highest average probability of correct classification compared to the Sturges and Rice methods. Specifically, SVM classifier leads with the highest average probability of correct classification, closely followed by MLP and KNN. On the other hand, Adaboost and RF classifiers exhibit lower average probabilities of correct classification across all methods. In terms of simulation times, KNN stands out as the most efficient among the three top-performing classifiers, demonstrating shorter times compared to SVM and MLP. Based on these results, in this paper, we favor the use of the SVM classifier because of its highest average probability of correct classification, prioritizing the sensitivity of feature extraction using MI over the simulation time of the classifiers.

Fig. 8 presents the simulation results of our proposed AMC method with a MIMO configuration of two transmit and two receive antennas ( $2 \times 2$ ). Fig. 8 shows the  $P_{cc}$  curves against different SNR values for various classifiers, including SVM and KNN classifiers. The proposed KNN and SVM classifiers based on mutual information is superior to the KNN and SVM classifier based on HOC, given in [49].

Our proposed future based classifier outperforms classifier based on HOC and achieves perfect classification at 5 dB SNR, whereas the compared classifier attains 100% classification at SNR of 13 dB. We conducted a comparison of the simulation times between our proposed method and the HOC-based approach. The algorithm execution time for our method was 0.1 seconds for KNN and 0.56 seconds for SVM, slightly longer than the HOC-based method, which recorded times of 0.53 seconds for SVM and 0.07 seconds for KNN.

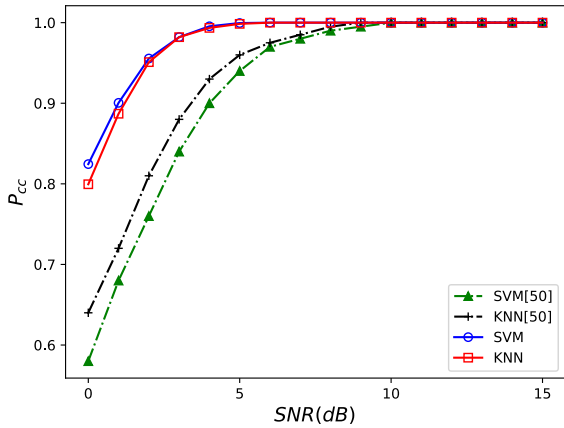


FIGURE 8. Performance comparison between MI-based and HOC-based classification methods.

Figure 9 shows the effect of N on the performance of our suggested method under different SNR values, where  $N = 256, 512, 1024$ . Figure 9 shows that the probability of correct classification increases with the growth of SNR, and then stabilizes. The five modulation types BPSK, QPSK, 8PSK, 16QAM, 64QAM are classified correctly with 100% probability when the SNR is about 5 dB. According to Figure 9, the probability of correct  $P_{cc}$  classification from the number of points varies with a small error, since mutual information is insensitive to the size of the data. Figure 10 below illustrates a comparison of the probability of correct classification for each modulation scheme relative to the other considered modulations.

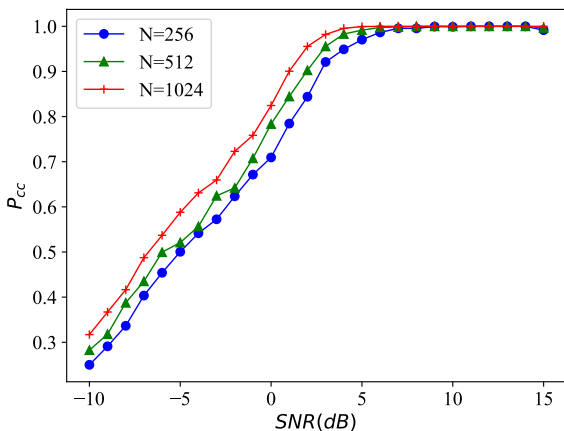


FIGURE 9. Average probability of correct classification for different N.

In general, the classification accuracy of QAM signals is determined better in comparison with PSK signals. Since with an increase in the modulation order of QAM signals, the points on the constellation are distributed more evenly. It is revealed that BPSK signals can be well identified using the classifier proposed by us even at low  $SNR = -5$  dB. The QPSK and 8PSK signals can be correctly recognized at SNR above 3 dB. Perfect 16QAM and 64 QAM detection is achieved at 5 dB SNR. Figure 11 shows the probabilities of

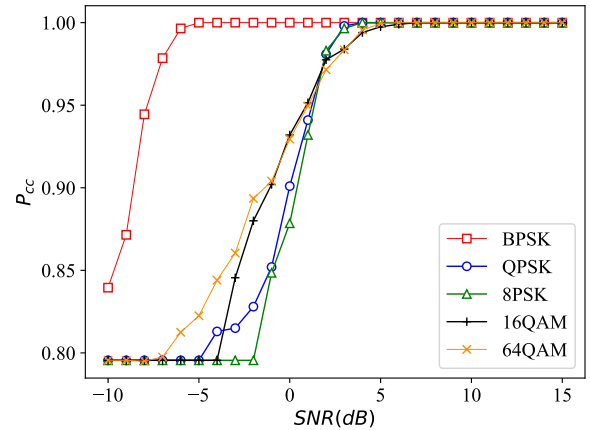


FIGURE 10. Probability of correct classification ( $P_{cc}$ ) among various modulation schemes ( $N=1024$ ).

correct classification of BPSK and QPSK signals obtained using the experimental setup.

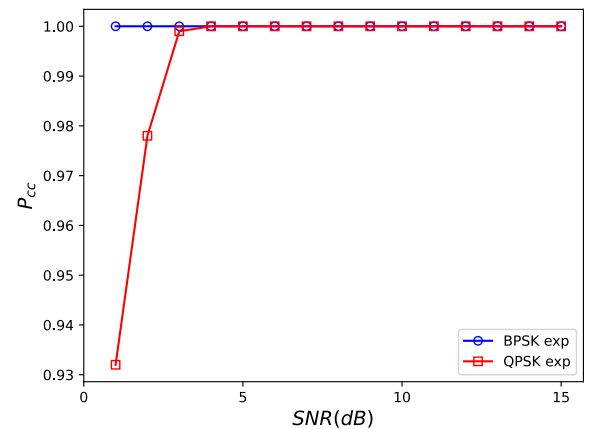


FIGURE 11. Average probability of correct classification of experimental data.

As shown in Figure 11, the QPSK signal can be correctly recognized at SNR above 3 dB, as in the theoretical results.

### V. CONCLUSION

In this article, we proposed a new classifier based on the extraction of mutual information features for recognizing the type of modulation schemes with low SNR for MIMO systems. This method does not require the elimination of MIMO signal interference and a large amount of training data, since our method is more resistant to noise uncertainty and spatial correlation in MIMO systems.

In addition, it is worth noting that feature extraction methods based on mutual information were not applied in AMC algorithms for MIMO systems. Our approach proves its effectiveness by using mutual information that considers the interdependencies between variables. The modeling and experimental results showed that our algorithm can work properly even at low SNR with a small number of observation samples, thereby representing a promising advancement

in the recognition of modulation schemes within MIMO systems.

## ACKNOWLEDGMENT

The authors thank the anonymous referee for constructive comments that helped to improve this article.

## REFERENCES

- [1] F. Meng, P. Chen, L. Wu, and X. Wang, "Automatic modulation classification: A deep learning enabled approach," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10760–10772, Nov. 2018.
- [2] 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.
- [3] 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. 21, no. 3, pp. 1–18, Dec. 2019.
- [4] P. Dileep, A. Singla, D. Das, and P. K. Bora, "Deep learning-based automatic modulation classification over MIMO keyhole channels," *IEEE Access*, vol. 10, pp. 119566–119574, 2022.
- [5] J. Lee, J. Kim, B. Kim, D. Yoon, and J. Choi, "Robust automatic modulation classification technique for fading channels via deep neural network," *Entropy*, vol. 19, no. 9, p. 454, Aug. 2017.
- [6] Z. Zhu, M. W. Aslam, and A. K. Nandi, "Support vector machine assisted genetic programming for MQAM classification," in *Proc. Int. Symp. Circuits Syst.*, Jun. 2011, pp. 1–6.
- [7] J. L. Xu, W. Su, and M. Zhou, "Likelihood-ratio approaches to automatic modulation classification," *IEEE Trans. Syst. Man, Cybern., Part C*, vol. 41, no. 4, pp. 455–469, Jul. 2011.
- [8] O. Ozdemir, R. Li, and P. K. Varshney, "Hybrid maximum likelihood modulation classification using multiple radios," *IEEE Commun. Lett.*, vol. 17, no. 10, pp. 1889–1892, Oct. 2013.
- [9] S. Majhi, R. Gupta, W. Xiang, and S. Glisic, "Hierarchical hypothesis and feature-based blind modulation classification for linearly modulated signals," *IEEE Trans. Veh. Technol.*, vol. 66, no. 12, pp. 11057–11069, Dec. 2017.
- [10] D. H. Al-Nuaimi, I. A. Hashim, I. S. Z. Abidin, L. B. Salman, and N. A. M. Isa, "Performance of feature-based techniques for automatic digital modulation recognition and classification—A review," *Electronics*, vol. 8, no. 12, pp. 2079–2092, 2019.
- [11] A. Hazza, M. Shoaib, S. A. Alshebeili, and A. Fahad, "An overview of feature-based methods for digital modulation classification," in *Proc. 1st Int. Conf. Commun., Signal Process., their Appl. (ICCSPA)*, Feb. 2013, pp. 1–6.
- [12] D. Hong, Z. Zhang, and X. Xu, "Automatic modulation classification using recurrent neural networks," in *Proc. 3rd IEEE Int. Conf. Comput. Commun. (ICCC)*, Dec. 2017, pp. 695–700.
- [13] M. A. Abdel-Moneim, W. El-Shafai, N. Abdel-Salam, E. M. El-Rabaie, and F. E. Abd El-Samie, "A survey of traditional and advanced automatic modulation classification techniques, challenges, and some novel trends," *Int. J. Commun. Syst.*, vol. 34, no. 10, pp. 1–15, Jul. 2021.
- [14] B. Kim, J. Kim, H. Chae, D. Yoon, and J. W. Choi, "Deep neural network-based automatic modulation classification technique," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2016, pp. 579–582.
- [15] K. Karra, S. Kuzdeba, and J. Petersen, "Modulation recognition using hierarchical deep neural networks," in *Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN)*, Mar. 2017, pp. 1–3.
- [16] W. Xie, S. Hu, C. Yu, P. Zhu, X. Peng, and J. Ouyang, "Deep learning in digital modulation recognition using high order cumulants," *IEEE Access*, vol. 7, pp. 63760–63766, 2019.
- [17] M. R. Bahloul, M. Z. Yusoff, A.-H. Abdel-Aty, M. N. M. Saad, and M. Al-Jemeli, "Modulation classification for MIMO systems: State of the art and research directions," *Chaos, Solitons Fractals*, vol. 89, pp. 497–505, Aug. 2016.
- [18] 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.
- [19] S. Hu, Y. Pei, P. P. Liang, and Y.-C. Liang, "Deep neural network for robust modulation classification under uncertain noise conditions," *IEEE Trans. Veh. Technol.*, vol. 69, no. 1, pp. 564–577, Jan. 2020.
- [20] T. Huynh-The, T.-V. Nguyen, Q.-V. Pham, D. B. da Costa, and D.-S. Kim, "MIMO-OFDM modulation classification using three-dimensional convolutional network," *IEEE Trans. Veh. Technol.*, vol. 71, no. 6, pp. 6738–6743, Jun. 2022.
- [21] Z. Zhao, Y. Zhou, F. Mei, and J. Li, "Automatic modulation classification by support vector machine," in *Proc. Int. Symp. Neural Netw.*, Aug. 2004, pp. 654–659.
- [22] J. Li, Q. Meng, G. Zhang, Y. Sun, L. Qiu, and W. Ma, "Automatic modulation classification using support vector machines and error correcting output codes," in *Proc. IEEE 2nd Inf. Technol., Netw., Electron. Autom. Control Conf. (ITNEC)*, Dec. 2017, pp. 60–63.
- [23] Y. Nie, X. Shen, S. Huang, Y. Zhang, and Z. Feng, "Automatic modulation classification based multiple cumulants and quasi-Newton method for MIMO system," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Mar. 2017, pp. 1–5.
- [24] S. Huang, L. Chai, Z. Li, D. Zhang, Y. Yao, Y. Zhang, and Z. Feng, "Automatic modulation classification using compressive convolutional neural network," *IEEE Access*, vol. 7, pp. 79636–79643, 2019.
- [25] J. Lee, B. Kim, J. Kim, D. Yoon, and J. W. Choi, "Deep neural network-based blind modulation classification for fading channels," in *Proc. Int. Conf. Inf. Commun. Technol. Converg. (ICTC)*, Oct. 2017, pp. 551–554.
- [26] Y. Wang, J. Gui, Y. Yin, J. Wang, J. Sun, G. Gui, H. Gacanian, H. Sari, and F. Adachi, "Automatic modulation classification for MIMO systems via deep learning and zero-forcing equalization," *IEEE Trans. Veh. Technol.*, vol. 69, no. 5, pp. 5688–5692, May 2020.
- [27] Y. Wang, G. Gui, H. Gacanian, T. Ohtsuki, H. Sari, and F. Adachi, "Transfer learning for semi-supervised automatic modulation classification in ZF-MIMO systems," *IEEE J. Emerg. Sel. Topics Circuits Syst.*, vol. 10, no. 2, pp. 231–239, Jun. 2020.
- [28] M. H. Shah and X.-Y. Dang, "An effective approach for low-complexity maximum likelihood based automatic modulation classification of STBC-MIMO systems," *Frontiers Inf. Technol. Electron. Eng.*, vol. 21, no. 3, pp. 465–475, Mar. 2020.
- [29] A. A. Salama, M. E. Morsy, S. H. Darwish, and E. I. Mohamed, "A novel SVM-based automatic modulation classifier," in *Proc. Int. Telecommun. Conf.*, Jul. 2022, pp. 1–3.
- [30] B. Jdid, K. Hassan, I. Dayoub, W. H. Lim, and M. Mokayef, "Machine learning based automatic modulation recognition for wireless communications: A comprehensive survey," *IEEE Access*, vol. 9, pp. 57851–57873, 2021.
- [31] T. Huynh-The, Q.-V. Pham, T.-V. Nguyen, T. T. Nguyen, R. Ruby, M. Zeng, and D.-S. Kim, "Automatic modulation classification: A deep architecture survey," *IEEE Access*, vol. 9, pp. 142950–142971, 2021.
- [32] S. Peng, S. Sun, and Y.-D. Yao, "A survey of modulation classification using deep learning: Signal representation and data preprocessing," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 12, pp. 7020–7038, Dec. 2022.
- [33] F. Zhang, C. Luo, J. Xu, Y. Luo, and F.-C. Zheng, "Deep learning based automatic modulation recognition: Models, datasets, and challenges," *Digit. Signal Process.*, vol. 129, Sep. 2022, Art. no. 103650.
- [34] R. Zhou, F. Liu, and C. W. Gravelle, "Deep learning for modulation recognition: A survey with a demonstration," *IEEE Access*, vol. 8, pp. 67366–67376, 2020.
- [35] Z. Zhu and A. K. Nandi, *Automatic Modulation Classification: Principles, Algorithms and Applications*. Hoboken, NJ, USA: Wiley, 2015.
- [36] Z. Zhang, Y. Li, S. Jin, Z. Zhang, H. Wang, L. Qi, and R. Zhou, "Modulation signal recognition based on information entropy and ensemble learning," *Entropy*, vol. 20, no. 3, p. 198, Mar. 2018.
- [37] Z. Zhang, Y. Li, X. Zhu, and Y. Lin, "A method for modulation recognition based on entropy features and random forest," in *Proc. IEEE Int. Conf. Softw. Qual., Rel. Secur. Companion (QRS-C)*, Jul. 2017, pp. 243–246.
- [38] H. Wang, L. Guo, Z. Dou, and Y. Lin, "A new method of cognitive signal recognition based on hybrid information entropy and D-S evidence theory," *Mobile Netw. Appl.*, vol. 23, no. 4, pp. 677–685, Aug. 2018.
- [39] 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.
- [40] T. M. Cover, *Elements of Information Theory*. Hoboken, NJ, USA: Wiley, 1999.
- [41] E. Archer, I. Park, and J. Pillow, "Bayesian and quasi-Bayesian estimators for mutual information from discrete data," *Entropy*, vol. 15, no. 12, pp. 1738–1755, May 2013.



- [42] S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Conditional entropy and mutual information," in *Numerical Recipes: The Art of Scientific Computing*, 2007.
- [43] L. Batina, B. Gierlichs, E. Prouff, M. Rivain, F.-X. Standaert, and N. Veyrat-Charvillon, "Mutual information analysis: A comprehensive study," *J. Cryptol.*, vol. 24, no. 2, pp. 269–291, Apr. 2011.
- [44] N. Veyrat-Charvillon and F. X. Standaert, "Mutual information analysis: How, when and why?" in *Proc. Int. Workshop Cryptograph. Hardw. Embedded Syst.*, 2009, pp. 429–443.
- [45] A. Balzanella and R. Verde, "Histogram-based clustering of multiple data streams," *Knowl. Inf. Syst.*, vol. 62, no. 1, pp. 203–238, Jan. 2020.
- [46] M. Awad and R. Khanna, "Support vector machines for classification," in *Efficient Learning Machines*. Amsterdam, The Netherlands: Elsevier, 2015, pp. 39–66.
- [47] H. Veisi, *Introduction to SVM*. Cham, Switzerland: Springer, 2023, pp. 3–18.
- [48] S. Ding, Z. Zhu, and X. Zhang, "An overview on semi-supervised support vector machine," *Neural Comput. Appl.*, vol. 28, no. 5, pp. 969–978, May 2017.
- [49] H. Tayakout, I. Dayoub, K. Ghanem, and H. Bousbia-Salah, "Automatic modulation classification for D-STBC cooperative relaying networks," *IEEE Wireless Commun. Lett.*, vol. 7, no. 5, pp. 780–783, Oct. 2018.



**N. USSIPOV** received the master's degree in engineering sciences from Taraz State University named after M. Kh. Dulaty, Taraz, Kazakhstan, in 2016. He is currently pursuing the Ph.D. degree with Al-Farabi Kazakh National University. He is a Senior Researcher and a Chief Programmer with the Department of Solid State Physics and Nonlinear Physics, Al-Farabi Kazakh National University. His current research interests include signal modulation classification, signal processing, network implementation, network information theory, and routing algorithms.



**S. AKHTANOV** received the Ph.D. degree in physics from Al-Farabi Kazakh National University, Almaty, Kazakhstan, in 2019. He is currently a Senior Researcher with the Department of Solid State Physics and Nonlinear Physics, Al-Farabi Kazakh National University. His current research interests include signal modulation classification, signal processing, network implementation, and optimization algorithms in wireless networks.



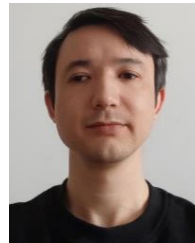
**Z. ZHANABAEV** received the Doctor of Physical and Mathematical Sciences degree from Al-Farabi Kazakh National University, Almaty, Kazakhstan, in 1996. In 1997, he received the title of professor from Al-Farabi Kazakh National University. Under his leadership, nine candidate dissertations, three doctoral dissertations, and five doctoral dissertations of international level (Ph.D.) were defended. He is the initiator of the development of a new scientific direction in Kazakhstan nonlinear physics and synergetics. He has written two monographs, 12 textbooks, and more than 300 scientific articles on modern areas of science.



**D. TURLYKOZHAYEVA** received the master's degree in physics from Tomsk Polytechnical University, Tomsk, Russia, in 2017. She is currently pursuing the Ph.D. degree with Al-Farabi Kazakh National University. She is a Senior Researcher and a Chief Editor with the Department of Solid State Physics and Nonlinear Physics, Al-Farabi Kazakh National University. Her current research interests include signal modulation classification, signal processing, network implementation, network theory, and routing of wireless networks.



**B. KARIBAYEV** received the Ph.D. degree in radio engineering, electronics and telecommunication from Al-Farabi Kazakh National University, Almaty, Kazakhstan, in 2019. He is currently a Senior Researcher with the Department of Telecommunication Engineering, Almaty University of Power Engineering and Telecommunications named after Gumarbek Daukeyev. His current research interests include signal modulation classification, signal processing, computer modeling, and MIMO antennas.



**T. NAMAZBAYEV** received the master's degree in radio engineering, electronics, and telecommunication from Al-Farabi Kazakh National University, Almaty, Kazakhstan, in 2019. He is currently a Senior Researcher with the Department of Solid State Physics and Nonlinear Physics, Al-Farabi Kazakh National University. His current research interests include signal modulation classification, signal processing, computer modeling, and MIMO antennas.



**D. ALMEN** received the master's degree in radio engineering, electronics and telecommunication from Al-Farabi Kazakh National University, Almaty, Kazakhstan, in 2021. She is currently a Researcher with the Department of Solid State Physics and Nonlinear Physics, Al-Farabi Kazakh National University. Her current research interests include signal modulation classification, signal processing, wireless networks theory, and wireless communication.



**A. AKHMETALI** is currently pursuing the degree with Al-Farabi Kazakh National University, Almaty, Kazakhstan. He is a Junior Researcher with the Department of Solid State Physics and Nonlinear Physics, Al-Farabi Kazakh National University. His current research interests include signal processing, network information theory, network implementation, and routing algorithms.



**XIAO TANG** (Member, IEEE) received the B.S. degree in information engineering (Elite Class named after Tsien Hsue-Shen) and the Ph.D. degree in information and communication engineering from Xi'an Jiaotong University, Xi'an, China, in 2011 and 2018, respectively. He is currently with the Department of Communication Engineering, Northwestern Polytechnical University, Xi'an. His research interests include wireless communications and networking, game theory, and physical layer security.

...