CENTRAL ASIAN JOURNAL OF MATHEMATICAL THEORY AND COMPUTER SCIENCES

https://cajmtcs.centralasianstudies.org/index.php/CAJMTCS

Volume: 06 Issue: 03 | July 2025 ISSN: 2660-5309

Article

A Hybrid Deep Learning and Bagging Method for Automatic Modulation Recognition Utilizing Time-Frequency Data

Intisar K. Saleh

1. Technical Engineering College-Kirkuk, Northern Technical University

* Correspondence: intisarks@ntu.edu.iq

Abstract: In satellite communications systems, submarine communications, and military communications, determining out the type of modulation is a crucial issue. A newly developed digital modulation classification model is introduced in this study to identify many types of modulated signals. At the first step, the density of specter for the frequencies accompanied with the modulation signals at the scalogram image is visually represented using continuous wavelet transform (CWT). Then, a deep convolutional neural network (CNN) is utilized to extract features from the scalogram pictures. The MRMR method is then used to get the best features. By decreasing the size of the features, the MRMR method improves classification speed and model interpretation. Using the group learning technique, the modulations are categorized in the fourth stage. Modulated signals with various levels of noise and SNRs ranging from 0 to 25 dB are taken into consideration in the simulations. The simulations' result shows that the suggested model outperforms other earlier research and functions effectively related to various noise levels.

Keywords: *Neural network, Deep learning, Modulation, Bagging method, Noise level.*

Citation: Saleh, I. K. A Hybrid Deep Learning and Bagging Method for Automatic Modulation Recognition Utilizing Time Frequency Data. Central Asian Journal of Mathematical Theory and Computer Sciences 2025, 6(3), 567-580.

Received: 03th Feb 2025
Revised: 11th Mar 2025
Accepted: 24th Apr 2025
Published: 21th May 2025

1. Introduction

Different modulation schemes are used in wireless communication, which complicates the communication environment. The technology of wireless communication is constantly changing. As a result, the ability to automatically and swiftly interpret and clarify communication signals is becoming more and more crucial. In blind signal processing, automated modulation classification (AMC), which includes software-defined radios (SDRs) [1] and cognitive radios (CR) [2], is essential. There are two kinds of AMC : The probability function is computed under various assumptions to identify the potential classes of arrived signals, and the results are then compared to a predefined threshold [3]. What is known as "Bayesian LB techniques" frequently have a high sensitivity to unidentified channel properties and necessitate costly calculations or a significant quantity of previous knowledge [4], [5]. The less computationally complex FB algorithms are able to identify less-than-ideal strategies. By looking at the properties of the received signals, one can identify the modulation type. High-quality features can result in dependable performance at a fair cost. Prior research has examined several of these traits, including WT, cyclostationary characteristics, and Higher-Order Cumulants (HOC) [6]. For AMC systems, prior studies frequently employed support vector machine (SVM), and D-Tree, in addition to Random Forests (RF) [7], and KNNs. These techniques are time-consuming, though, because handcrafted features need technical know-how and domain knowledge. Due to its remarkable achievements in different fields, such as computer view, examining emotion, in addition to speech detection, deep learning (DL)

has been increasingly popular in recent years. In contrast to traditional approaches for data analysis and processing, DL must be able to handle complex data with many dimensions in automatic way with no need for manually created characteristics [8], [9]. This has led to its expansion into other fields, such as communications [10]. Numerous deep learning approaches have been presented up to this point, including RNN, deep convolutional networks, LSTM networks, and others.

In order to get the complex relationships found in time-related signals over channels of Rayleigh fading via different added noise conditions, an automated modulation categorization (AMC) method utilizing recurrent neural networks (RNNs) is proposed in [11]. A long short-term memory (LSTM) network is a sophisticated RNN structure specifically developed to investigate the long-term reliance of signals with variable-length modulation in the temporal domain. For multi-class classification tasks, convolutional neural networks (CNNs) have outperformed long short-term memory (LSTM) with repeated neural network (RNN) regarding their ability to extract important discriminative characteristics from feature representations with multiple scales. The rate of classifying on the 24-modulation DeepSig data set is significantly improved by using a compact CNN architecture that consists of multiple residual convolutional stacks. In terms of recognition performance, deep learning (DL) surpasses conventional machine learning (ML) techniques, removing the need for feature engineering expertise [12]. In, the authors used the model of AlexNet CNN and arrangement diagrams to perform network training and classification tasks. ACGAN-based data augmentation was employed by the authors in [13], [14]. The Caffe technique was applied using the modulation categorization procedure. Peng et al. produced grid-like images for AMC (AlexNet and GoogLeNet) using two pre-trained models. Performance can be improved by feature fusion using various time-frequency adjustments with the addition white Gaussian noise signal technique [15]. Although these techniques convert the AMC into several good-researched image identification problems, they all necessitate a sophisticated image processing step. The research claims that the AMC problem can be resolved utilizing machine learning models and several techniques for feature extraction and selection. Getting rid of requiring any prior understanding of feature engineering and DL surpasses ML in recognizing [16], [17], [18]. The majority of current automatic modulation categorization (AMC) techniques perform worse when widely used convolutional layer topologies in extremely complicated networks. This problem is exacerbated by the fact that the extracted feature maps are ineffective at capturing the intended representations and that the process's trainable parameters are computationally costly. Furthermore, these techniques require a large amount of training time. This study presents an efficient AMC technique based on the display of time-frequency information utilizing deep convolutional neural networks CNN and CWT. The density of specter of the frequencies for modulated signals is visually represented in its first stage using the continuous wavelet transform (CWT). The second step involves extracting features from scalogram images using a deep convolutional neural network technique. The network's internal convolution layer discovers and expands the frequency-time sequences in the images of scalogram for high power by reducing the high dimensions of the images without sacrificing information [19].

Minimum Redundancy Maximum Relevance (MRMR) method is then utilized to choose the best features by applying the notions of minimum redundancy and maximum relevance. By reducing the feature dimensions, the MRMR method not only speeds up execution but also improves and makes it easier to grasp the classification model. The ensemble learning technique is used in the fourth phase to classify modulations. The bagging technique uses a group of weak learners to generate a strong learning model. This method's main advantage is its capacity to successfully handle the over-fitting problem. Contributions made to this work include the following:

a. A potent four-stage noise-resistant approach is used to detect digital modulation kinds with excellent accuracy.
b. The CWT transformation offers a visual depiction of the frequency spectrum densities for the modulated signals.

c. The scalogram transformation is used to present time-frequency data, which enhances the deep convolutional network's capacity to identify patterns connected to any digital modulation.
d. The comprehension of the model of classification is improved by reducing the feature space dimensionality and removing redundancy of feature by the application of the MRMR method. As a result, the diagnosis's accuracy and the system's overall efficacy have both increased [20].
e. A group of weak learners has been used to generate a strong learning model in the classification stage, which reduces the variance in noisy data and addresses the issue of overfitting.

2. Materials and Methods

In this work, a new classification system covering modulation signals having noise powers ranging from 0 dB to 25 dB is presented. The suggested method generates signals after modeling the telecommunications environment. CWT first processes these modulated signals to expand the time-frequency information and create scalogram images. In order to extract features, the CNN deep neural networks are then fed the scalogram images. The MRMR technique is then used to minimize the dimensionality of the characteristics that the deep convolutional neural network has extracted. Lastly, the bagging method is used to identify the modulation kinds by classifying the chosen features. The following subsections provide an explanation of each stage in the suggested strategy. The suggested method's flowchart is displayed in Figure 1.

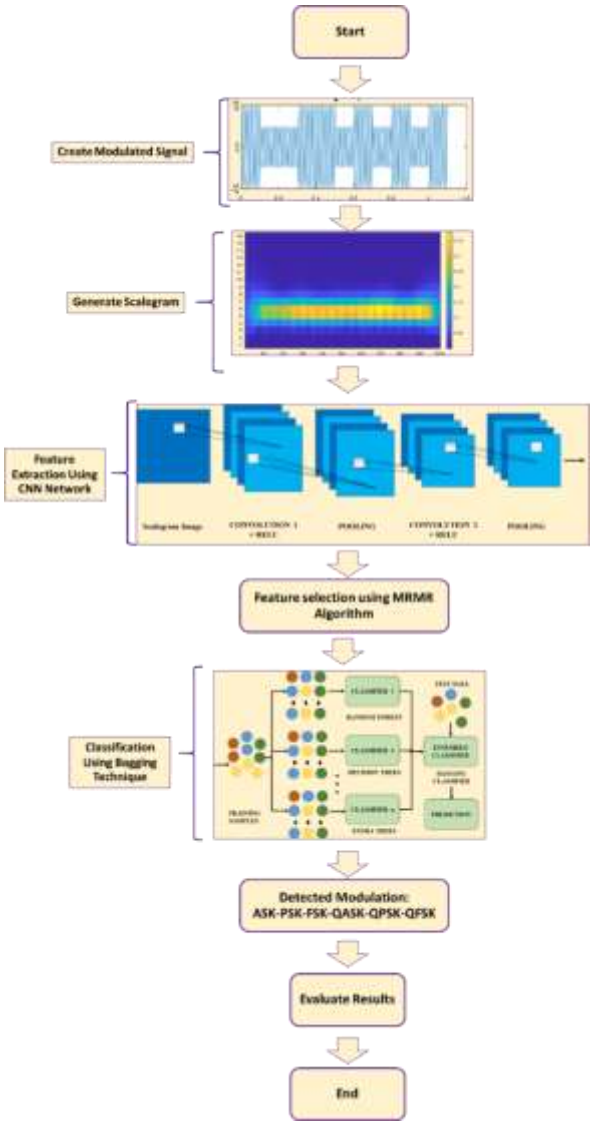

Figure 1. Process of the suggested algorithm.

2.1. Simulations for environment and communication signals

The digital base-band signal is changed to a high-frequency pass-band signal, according to get constraints, by using digital passband modulation. With the use of this digital transmission technique, various modulations that take place in the transition band can be categorized into three different types: phase shift keying (PSK), frequency shift keying (FSK), and amplitude shift keying (ASK).

a. Amplitude shifts keying: The binary representation of the baseband signal is transmitted by A1 and A2 using the same frequency carriers.
b. Frequency shifts keying: The baseband signal's binary values are represented using the frequency shift keying (FSK) method. Two frequencies having the same amplitude are utilized in FSK.
c. Phase- shift keying: A modulation which uses phase changes of the same amplitude and frequency to alter bits 0 and 1 of baseband signals.

The basic modulation formulae for ASK, PSK, and FSK are shown below:

$$X_{m_{ASK}}(t) = A_m \cos \omega_c t \quad (1)$$

$$X_{m_{PSK}}(t) = A \cos(\omega_c t + \theta_m) \quad (2)$$

$$X_{m_{FSK}}(t) = A \cos(\omega_{c_m} t) \quad (3)$$

According on the modulation type used for multi-level transmission, the use of the M-piece can have various amplitudes (MASK), frequencies (MFSK), in addition to phase values (MPSK). This is done by varying the carrier's phase, amplitude, or frequency among different quantities [37]. Using numerous carriers instead of one can boost the throughput of data transmission. The binary value of the related bit group via the signal's information is carried and sent via the carrier linked to the M-bit symbol. For multiple-level modulations, QASK, OFSK, and OPSK modulations are frequently employed. Those modulations quadruple the transmission speed and send two bits at a time. There is a significant increase in the data transfer rate when 8, 16, and 32 carriers are added. Furthermore, as carriers proliferate, the demodulator circuits must also grow more intricate. Six different modulation types have been simulated and transmitted across a communication channel (AWGN) in this study: three (BASK, BFSK, and BPSK) binary modulations and three (QASK, QFSK, and QPSK) quadratic modulations. Lastly, the bagging technique classifies the chosen features to automatically identify the modulation type.

2.2. Visualizing time-frequency data with the continuous wavelet transform (CWT)

Whenever the CWT is applied to signals having frequency changes over time, a diagram of time-frequency is produced. The time-frequency domain transformation technique is essential to pattern recognition methods. The wavelet transform is a good fit for this transformation. Non-stationary signals like ECG, EEG, and EMG can be successfully transformed using wavelet transforms like Daubechies, Morlet, Symlets, and Gaussian [38, 39]. The signal in a wavelet function can be altered using these wavelet functions.

The following is the Continuous Wavelet Transform (CWT) definition having a signal:

$$Z(a.b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} s(t) \psi^* \left(\frac{t-b}{a} \right) dt \quad (4)$$

The variables (a and b) control wavelet's scaling in addition to translation, whereas (t) stands for a finite signal of energy and $\psi*$ for the primary wavelet's complex conjugate. High frequencies of the signal are disclosed when the wavelet contracts with the scale values go smaller, while the low frequency information of the signal is revealed when the wavelet expands with time and scale values are larger [40]. In order to calculate the CWT, the a and b parameters are continually varied through the signal's lengths and scales, respectively [21]. Similar to a spectrogram produced by the fast Fourier transform (FFT), a scalogram means a graphical representation for a signal's Continuous Wavelet

Transform (CWT) or representation of Short-Time Fourier Transform (STFT). In terms of frequency and temporal resolution, the Continuous Wavelet Transform (CWT) outperforms the Short-Time Fourier Transform (STFT). By allowing utilizing of analysis windows having different sizes at various frequencies, it does this. The CWT-developed scalograms display the frequencies in the signals at various times. An illustration of the modulated signal with its matching scalogram can be found in Figure 2.

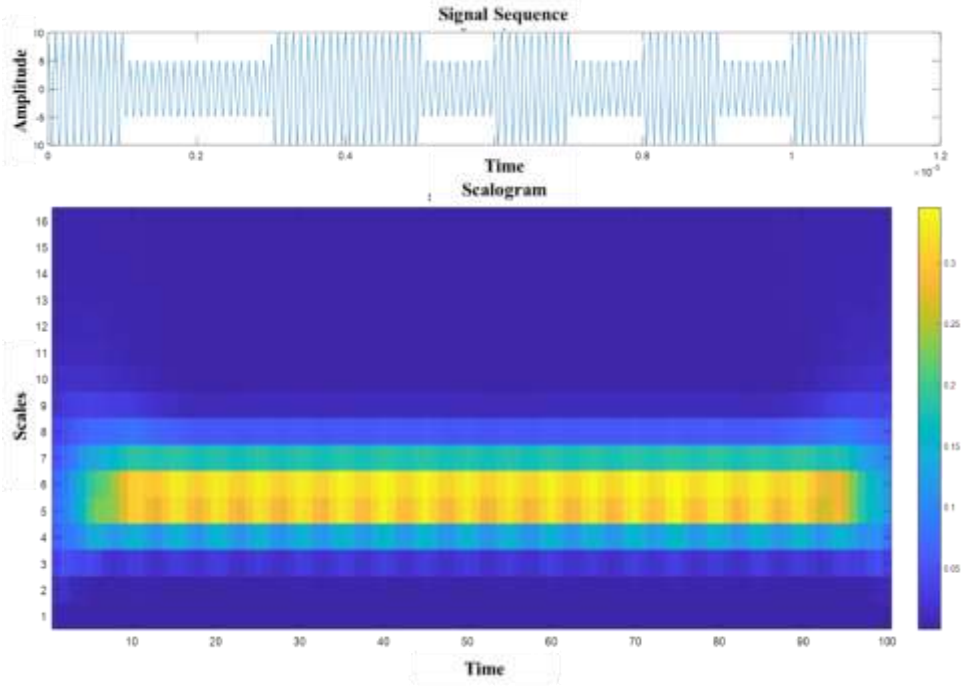

Figure 2. Illustration of the modulated signal with the associated scalogram.

2.3. Utilizing a convolutional neural network model for feature extraction

Other well-liked neural architecture type is CNN. The primary two distinctions between it and the ANN are the architecture in addition to the input data. CNN employs pictures, while ANN uses numerical values. A collection of pixels via the dimensions w, h, and d clears an image I. These images are scaled by height, width, and depth in order of precedence. Depending on the color model used, the depth of the image changes [22]. The depth (d) equals to three in the color space model RGB, that makes use of three-color channels. Convolutionally, fully connected, and layers of pooling make up the CNN model's framework, accordingly. To process and extract unique features via the input image, convolutional layers use filters. Every pixel in the $I_{(x,y)}$ image and a k (p × p matrix) filter go through a analytical process known as a star operation* during the convolution operation. Coordinates (x, y) are subjected to this process in a separate and an independent manner.

$$k * I_{x,y} = \sum_{i=1}^{p} \sum_{j=1}^{p} k_{i,j} . I_{x+i-1,y+j-1} + b_1 \quad (5)$$

Where the bias is b_1. The pooling layer is utilized to reduce the size of the image's file. Here, every pixel and its surrounding pixels are examined using the function ω(.), which applies procedures like minimum, maximum, and average computations. The smaller image ω(.) is supplemented with the function-performing pixel. To make it more understandable, we could write it like this:

$$\omega(I_{x,y}) = \max_{i,j\in\{-1,0,1\}} I_{x-i,y-j} \quad (6)$$

The picture resizing calculation can be written like (w-k)/(s+1) × (h-k)/(s+1), where s is the shift for kernel. Before arriving at the completely connected layer, that represents the final layer in the model design. It is possible to repeat these two resizing layers [23]. There are several hidden layers and one output layer in this model. The CNN-bagging structure is displayed in Figure 3. Convolutional layers, a totally linked layer, with two pooling layers make up this structure. As can be observed, grouping is done utilizing the output of the layer for the second pooling. The suggested CNN produces six probabilities for each of the six modulation classes after processing a scalogram image.

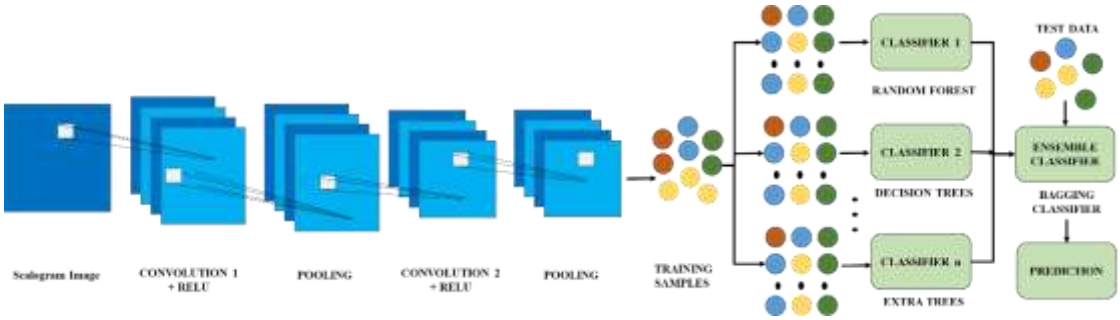

Figure 3. Structure of the model of CNN-bagging in the current investigation.

2.4. Minimum redundancy maximum relevance (MRMR) technique

The method of meticulously choosing a smaller group of characteristics from a bigger collection while removing redundant and unnecessary ones is known as feature selection. In addition to decreasing the feature's dimensionality set and, thus, the quantity for data needed to learning, this strategy also lessens the effect of high-dimensional dataset on the total performance of algorithms, improving capabilities for the generalization. In addition, it improves the capacity to interpret the models and the speed at which they are executed [24], [25], [26]. The MRMR technique aims to maximize the relation between characteristics and related labels (class). With the simultaneous goal of lowering the correlation between each feature, this technique uses mutual information to calculate how similar two variables are. If X and Y are two variables, so the reciprocal information may be found using the following equation (7):

$$I(X;Y) = \sum_{x\epsilon X}\sum_{y\epsilon Y} p(x,y)\log\frac{p(x,y)}{p(x)p(y)} \quad (7)$$

Here: $p\ (x, y)$: The probability density of variables X and Y. The concepts of lowest redundancy (R) and maximum relevance (D) are combined in the MRMR approach. Therefore, equations (8) and (9) can be used to maximize D and R at the same time.

2.4.1 The criterion for maximum relevance

The following formula is used to determine this criterion, which stipulates that the relationship between each class's attributes and its label should be maximized:

$$Max\ D(S,c), D = \frac{1}{|S|}\sum_{xi\epsilon S} I \quad (8)$$

In the above relation:
S, set of features
$|S|$, the feature size of set for the S-space.
xi: the characteristics of the individual,
c: are classes.

$I\ (xi, c)$: A common comprehension of the characteristics of each class with the class label.

2.4.2 The criterion for minimum redundancy

The following definition applies to this criterion, which specifies that the link between features should be minimized:

$$Min\ R(S), R = \frac{1}{|S|^2} \sum_{xi,xj\ \in S} I\ (xi, xj) \tag{9}$$

Incremental search techniques are employed in applications to find nearly optimal attributes [27], [28]. The following formula should be used to calculate the Sm-1 feature subset in order to determine the ideal feature subset (with m-1 feature):

$$max_{xj \epsilon X - S_{m-1}} \left[I(x_i; c) - \frac{1}{m-1} \sum_{xi\ \epsilon\ S_{m-1}} I(xj; xj) \right] \tag{10}$$

2.5 Types of modulation categorization employing bagging methods

A fraction of the basic data is sent to every classifier in this study's automatic modulation classification using the bagging approach. This indicates that every classifier looks at a subset of the set of data while uses that subset to build its model. This subset is chosen using replacement, which implies that every sample may be selected more than once [29], [30]. According to study, the categorization approach can enhance learning and recognition skills with greater precision for all kinds of data. Figure 4 shows the general operation of this approach.

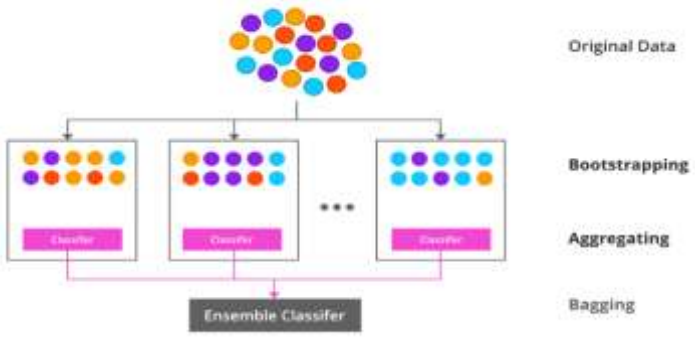

Figure 4. Bagging algorithm classification model.

3. Result

This part compares the offered model performance with other state-of-the-art approaches and measures the effectiveness of the suggested approach using common assessment criteria. MATLAB (2022b) software is used to train the model, which needs an NVIDIA card of graphics with 8 GB for onboard RAM. Signal-to-Noise Ratio (SNR) rates between 0 and 25 decibels (dB) are used to assess the model's performance. In the system modeling phase, 1000 scalogram images are produced for each modulation class, which translates to 6000 modulation signals. Several preprocessing methods were applied to the scalogram images. Standard Continuous Wavelet Transform (CWT) images were automatically cropped to remove all unwanted white spaces, so the corresponding resolutions were then decreased to 227 by 227 by 3 pixels from 657 by 535 by 3. Certain areas of interest in scalogram images might be emphasized. 70% of the scalogram pictures are used to train the CNN network [31], [32]. Stated otherwise, the CNN model receives

4200 scalogram images as train input, whereas 1800 scalogram images are taken into consideration during the test phase. A neural network extracts 1000 features from each scalogram image. The MRMR method then reduces each characteristic vector's dimensionality to 500 dimensions. Lastly, the bagging technique is used to classify these feature vectors. Additionally, a validation technique that divides the dataset to 10 folds with iteratively trains and tests the model utilizing each fold like a validation set is used to calculate the classification accuracy [33]. Notably, the data is divided to training with examining using a random permutation method, and the outcomes mentioned are the average for the program's 50 runs.

3.1. Criteria used for evaluation

Precision, accuracy, recall, and score of F1 are performance metrics employed in this study and are computed using the following formulas [41, 42]:

$$Accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \tag{11}$$

The ratio of pertinent examples to all examples discovered is referred to as precision (P). P is computed using the formula:

$$Precision = \frac{TP}{TP\ +\ FP} \tag{12}$$

The percentage of pertinent cases that are found among all the pertinent examples is known as recall (R). R can be calculated with the formula below:

$$Recall = \frac{TP}{TP\ +\ FN} \tag{13}$$

Due to the fact the precision (P) besides recall (R) assessments can occasionally yield contradictory results, the score of F1 is a commonly utilized evaluation statistic. The following formula illustrates how the score of F1 is determined by taking the mean of harmonic for precision and recall:

$$F1 - score\ = \frac{2PR}{(P + R)} \tag{14}$$

In 11-14 equations: The term "True Positive" (TP) refers to detections that are accurately positive. False Positive (FP): Denotes detections that are incorrectly positive. False Negative (FN): Indicates incorrectly negative results. True Negative (TN): Denotes detections that are correctly negative.

3.2. Training process evaluation

Curves of learning for accuracy and loss in the training set over 150 periods are displayed in Figures 5 and 6, respectively. The current model improves as accuracy rises because it captures the model's performance. As Figure 5 illustrates, the accuracy curve is increasing with time, suggesting that the model improves with experience (learning). Furthermore, loss (or cost) shows "how much poorly this model is performing" or the mistake in the model. Reducing the amount of loss is the model's goal [34], [35]. This is accomplished through the use of techniques like gradient descent. Consequently, and temporarily, the current model will perform better the loss gets. A loss or cost function is calculated to quantify the loss. The expected result of the learning process is shown in Figure 6. It is evident that the training set's loss curve is getting smaller with time. This indicates that the current model is gaining knowledge. Learning curves fluctuate slightly, but with time, the accuracy rises and the loss decreases, indicating that the model is learning.

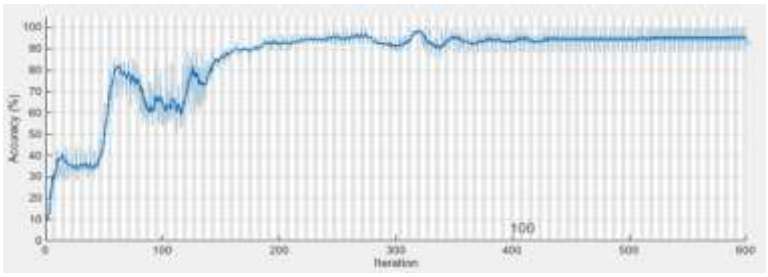

Figure 5. A graph shows how accuracy of the model increased through training.

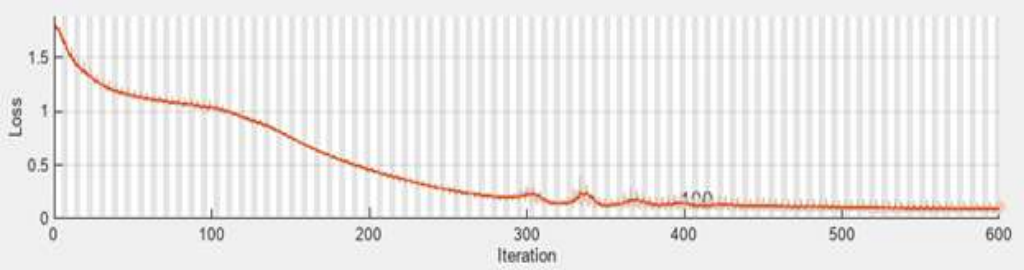

Figure 6. A graph shows how the model's loss decreased through training.

3.3. An evaluation for the proposed model with comparative study

A confusion matrix is developed to evaluate performance of the model on classification of the digital modulation challenge. The calculated confusion matrix of the suggested approach with 0 dB SNR is displayed in Table 1. which shows the results of classification for every category in the matrix of confusion. Stated otherwise, it implies that the classification performance of the suggested model is suitable.

Table 1. A confusion matrix displaying the outcomes for classification with 0 dB SNR at every class.

		Actual						
		FSK	QFSK	PSK	QPSK	ASK	QASK	
Predicted	FSK	299	0	0	0	1	0	300
	QFSK	1	299	0	1	0	0	301
	PSK	0	0	299	0	0	0	299
	QPSK	0	1	0	299	0	0	300
	ASK	0	0	0	0	299	0	299
	QASK	0	0	1	0	0	300	301
		300	300	300	300	300	300	

The performance of a binary classifier for different cutoff points is displayed in machine learning using a ROC curve, which depends on two basic metrics: True Positive Rate (TPR) in addition to False Positive Rate (FPR). Ratio for true positives to all positive cases is known as TPR, sensitivity, or recall. The ratio of negative cases that are mistakenly categorized as positive is known as the FPR, and it has the following definition:

$$FPR = \frac{FP}{TN + FP} \tag{15}$$

Drawing the TPR versus the FPR for each potential threshold allows one to observe the trade-off classifies between those two metrics. A good classifier will have a low FPR and a high TPR, placing it in the upper-left corner of the ROC curve. In contrast, a bad classifier will have a reduces TPR and an increased FPR, placing it in the lower-right quadrant of the curves of ROC. A random classifier will eventually have an equal TPR and FPR, placing it on the diagonal line for ROC curve. Figure 7 shows the plot of the current model's ROC curve. As can be observed, this curve is shown toward the top-left corner and has a low FPR and a high TPR. Therefore, it can be said that the current model has a high degree of accuracy in classifying modulations [36], [37], [38]. The accuracy of the suggested approach for various SNRs is displayed in Figure 8. The results show that 0 dB SNR has the lowest accuracy of detection, while 20 dB with 25 dB SNRs have the highest. Figure 7 clears how the accuracy for detection rose in tandem with SNR. This

experiment has achieved the expected increase in detection rate since the signal strength increases relative to the noise as SNR increases.

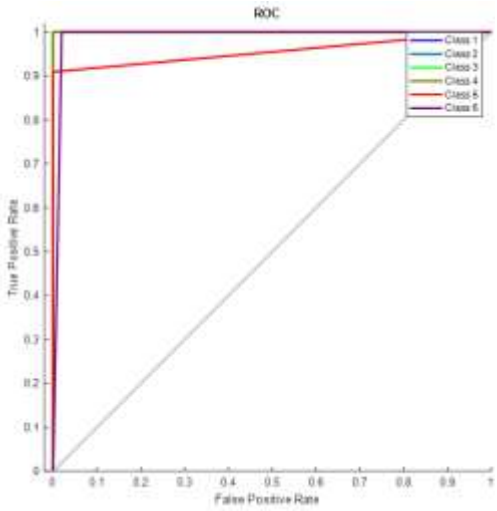

Figure 7. ROC curve for modulation categorization classes that the model produced.

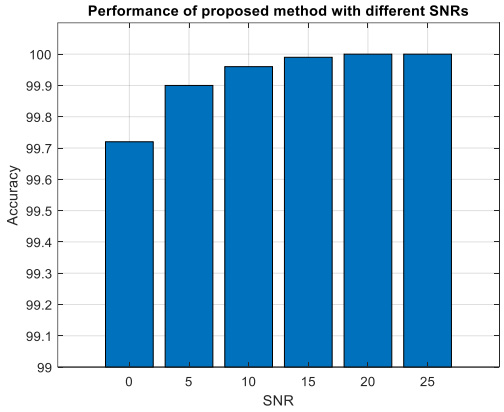

Figure 8. Classification results' accuracy at varying SNR rates.

A bar chart comparing the modulation detection outcomes of the suggested approach to a number of pre-trained designs is shown in Figure 9. Outcomes of the investigations for SNRs range between 0 and 25 db are averaged to produce all of the criteria utilized in this graphic [39], [40]. The outcomes are assessed for the AlexNet, VGG-19, VGG-16, and GoogLeNet techniques, as illustrated in Fig. 9. According to the findings, the accuracies of the AlexNet, GoogLeNet, VGG-16, VGG-19, and techniques are 99.4, 99.55, 99.90, and 99.80, respectively [41], [42], [43], [44], [45], [46]. The accuracy rate of the suggested approach was 99.92%. To classify digital modulation, the suggested method is compared with a number of convolutional neural network-based techniques. According to Table 2, the proposed CNN model produced a high classification accuracy of more than 99.9%.

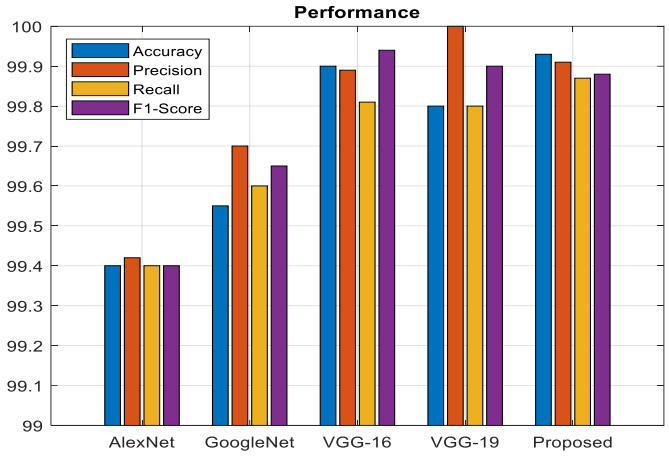

Figure 9. The efficiency of the suggested strategy and trained models.

Table 2. Results for the proposed method with several previous CNNs methods.

No.	Title	Year	Ref.	Accuracy
1	Using the Dyadic Aggregated Autoregressive Model (DASAR) to Classify Modulations Automatically	2020		96.7
2	R. Lin, W. Ren, X. Sun, Z. Yang, K. Fu, "A Hybrid Neural Network for Fast Automatic Modulation Classification", IEEE Access, Vol. 8, 2020, pp. 130314 – 130322.	2020		87.4
3	Recognizing modulation automatically with CNN deep learning models	2023		94.39
4	A multiscale network-based automatic modulation recognition system that incorporates statistical aspects	2023		97
5	Investigation into modulation recognition using sequential feature fusion	2023		93
6	Proposed Method	-	-	99.92

4. Discussion

The results obtained from this study demonstrate the superior performance of the proposed hybrid method combining deep convolutional neural networks (CNN), the Minimum Redundancy Maximum Relevance (MRMR) feature selection method, and the bagging ensemble technique for automatic modulation classification (AMC). The approach effectively utilizes the continuous wavelet transform (CWT) to visualize time-frequency representations, providing rich input features for the CNN to extract relevant patterns. By integrating MRMR, the model successfully reduces feature dimensionality, enhancing classification speed and interpretability without sacrificing performance. The use of bagging addresses common challenges such as overfitting, particularly in low Signal-to-Noise Ratio (SNR) environments. Comparative analysis against established architectures like AlexNet, VGG-16, VGG-19, and GoogLeNet reveals that the proposed method achieves the highest accuracy (99.92%) across various SNR levels, outperforming prior studies. Furthermore, the ROC curve analysis confirms the robustness and reliability of the classifier, even at 0 dB SNR. However, while the model exhibits excellent performance, its dependency on a substantial amount of computational resources for training deep networks could limit its deployment in resource-constrained environments. Future work should explore model optimization techniques, such as lightweight network architectures or knowledge distillation, to enhance the applicability of the proposed framework in real-world communication systems with limited hardware capabilities.

5. Conclusion

A four-step technique for automatically detecting digital modulation based on visualization was presented in this research. The basic idea behind the suggested approach is to use continuous wavelet transformation to extract time-frequency information in addition to display it as a scalogram image. Modulation is then found utilizing this information, convolutional deep neural networks, with the bagging technique. Additionally, the MRMR algorithm was used to select the best features in addition to minimize the size of the space of feature. Experiments were carried out at various SNR levels between 0 and 25 dB to examine the robustness of the suggested approach. The simulations' outcomes demonstrated that the suggested method's accuracy in every SNR was greater than 99%, demonstrating its resistance to variations in SNR. A comparative analysis to evaluate the method's efficiency against other widely utilized approaches is also performed. The outcomes unequivocally showed that the suggested procedure outperformed other cutting-edge techniques.

REFERENCES

[1] D. Zhang, Y. Lu, Y. Li, *et al*., "Frequency learning attention networks based on deep learning for automatic modulation classification in wireless communication," *Pattern Recognit.*, vol. 137, p. 109345, 2023.

[2] S. Wei, Z. Sun, Z. Wang, *et al*., "An efficient data augmentation method for automatic modulation recognition from low-data imbalanced-class regime," *Appl. Sci.*, vol. 13, p. 3177, 2023, doi: 10.3390/app13053177.

[3] B. Jdid, K. Hassan, I. Dayoub, *et al*., "Machine learning based automatic modulation recognition for wireless communications: A comprehensive survey," *IEEE Access*, vol. 9, pp. 57851–57873, 2021, doi: 10.1109/ACCESS.2021.3071801.

[4] W. Chen, Z. Xie, L. Ma, *et al*., "A faster maximum-likelihood modulation classification in flat fading non-Gaussian channels," *IEEE Commun. Lett.*, vol. 23, pp. 454–457, 2019.

[5] A. Krayani, A. S. Alam, M. Calipari, *et al*., "Automatic modulation classification in cognitive-IoT radios using generalized dynamic Bayesian networks," in *Proc. 7th IEEE World Forum Internet Things (WF-IoT)*, pp. 235–240, 2021, doi: 10.1109/WF-IoT51360.2021.9594936.

[6] P. Urriza, E. Rebeiz, P. Pawelczak, and D. Cabric, "Computationally efficient modulation level classification based on probability distribution distance functions," *IEEE Commun. Lett.*, vol. 15, pp. 476–478, 2011, doi: 10.1109/LCOMM.2011.032811.110316.

[7] M. R. Ismael, H. J. Abd, and M. T. Gatte, "Recognition of APSK digital modulation signal based on wavelet scattering transform," in *Lecture Notes in Networks and Systems*, Springer, pp. 469–478, 2022.

[8] X. Wu, L. Lu, and M. Jiang, "Deep learning aided cyclostationary feature analysis for blind modulation recognition in massive MIMO systems," *Digit. Signal Process.*, p. 103890, 2023.

[9] X. Tan, Z. Xie, X. Yuan, *et al*., "Small sample signal modulation recognition based on higher-order cumulants and CatBoost," in *Proc. 7th Int. Conf. Commun., Image Signal Process. (CCISP)*, IEEE, pp. 324–329, 2022.

[10] . Liu and Y. Liu, "Modulation recognition with pre-denoising convolutional neural network," *Electron. Lett.*, vol. 56, pp. 255–257, 2020, doi: 10.1049/el.2019.3586.

[11] P. Ghasemzadeh, "A novel graph neural network-based framework for automatic modulation classification in mobile environments," 2023.

[12] 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. (ITC-Egypt)*, 2022, doi: 10.1109/ITC-Egypt55520.2022.9855683.

[13] X. Sun, S. Su, Z. Zuo, *et al*., "Modulation classification using compressed sensing and decision tree–support vector machine in cognitive radio system," *Sensors (Switzerland)*, vol. 20, 2020, doi: 10.3390/s20051438.

[14] B. A. Alhadi, T. M. Hasan, and H. A. Hamed, "Digitally modulated signal recognition based on feature extraction optimization and random forest classifier," in *New Trends Inf. Commun. Technol. Appl.: Proc. 4th Int. Conf. NTICT*, Springer, pp. 75–84, 2020.

[15] Y. Li, G. Hong, and C. Feng, "Application and effectiveness of weighted KNN in pattern recognition of communication modulated signals," in *Proc. IEEE 4th Int. Conf. Civil Aviation Safety Inf. Technol. (ICCASIT)*, IEEE, pp. 744–748, 2022.

[16] A. Kamble, P. H. Ghare, and V. Kumar, "Deep-learning-based BCI for automatic imagined speech recognition using SPWVD," *IEEE Trans. Instrum. Meas.*, vol. 72, pp. 1–10, 2022.

[17] S. Peng, L. Cao, Y. Zhou, *et al*., "A survey on deep learning for textual emotion analysis in social networks," *Digit. Commun. Netw.*, vol. 8, pp. 745–762, 2022.

[18] Z. Lingxin, S. Junkai, and Z. Baijie, "A review of the research and application of deep learning-based computer vision in structural damage detection," *Earthq. Eng. Eng. Vib.*, vol. 21, pp. 1–21, 2022.

[19] H. Tayakout, E. Boutellaa, and F. Z. Bouchibane, "On the robustness of digital modulation recognition for cooperative relaying networks under imperfect CSI," in *Proc. 7th Int. Conf. Image Signal Process. Appl. (ISPA)*, IEEE, pp. 1–6, 2022.

[20] C. T. Nguyen, N. Van Huynh, N. H. Chu, *et al.*, "Transfer learning for wireless networks: A comprehensive survey," *Proc. IEEE*, 2022.

[21] Z. Wang, P. Wang, and P. Lan, "Automatic modulation classification based on CNN, LSTM and attention mechanism," in *Proc. IEEE 8th Int. Conf. Comput. Commun. (ICCC)*, IEEE, pp. 105–110, 2022.

[22] Y. Wang, S. Fang, Y. Fan, and Z. Wang, "A CLSTM network algorithm for automatic modulation recognition," in *Proc. Int. Conf. Signal Process., Comput. Netw., Commun. (SPCNC)*, SPIE, pp. 456–462, 2022

[23] N. Chakravarty, M. Dua, and S. Dua, "Automatic modulation classification using amalgam CNN-LSTM," in *Proc. IEEE Radio Antenna Days Indian Ocean (RADIO)*, IEEE, pp. 1–2, 2023.

[24] S. Hanna, C. Dick, and D. Cabric, "Signal processing-based deep learning for blind symbol decoding and modulation classification," *IEEE J. Sel. Areas Commun.*, vol. 40, pp. 82–96, 2022, doi: 10.1109/JSAC.2021.3126088.

[25] P. Ghasemzadeh, M. Hempel, and H. Sharif, "A robust graph convolutional neural network-based classifier for automatic modulation recognition," in *Proc. Int. Wireless Commun. Mobile Comput. (IWCMC)*, IEEE, pp. 907–912, 2022.

[26] C. Hou, G. Liu, Q. Tian, *et al.*, "Multi-signal modulation classification using sliding window detection and complex convolutional network in frequency domain," *IEEE Internet Things J.*, vol. 9, pp. 19438–19449, 2022, doi: 10.1109/JIOT.2022.3167107.

[27] K. Liao, Y. Zhao, J. Gu, *et al.*, "Sequential convolutional recurrent neural networks for fast automatic modulation classification," *IEEE Access*, vol. 9, pp. 27182–27188, 2021, doi: 10.1109/ACCESS.2021.3053427.

[28] S. Huang, R. Dai, J. Huang, *et al.*, "Automatic modulation classification using gated recurrent residual network," *IEEE Internet Things J.*, vol. 7, pp. 7795–7807, 2020.

[29] R. Utrilla, E. Fonseca, A. Araujo, and L. A. Dasilva, "Gated recurrent unit neural networks for automatic modulation classification with resource-constrained end-devices," *IEEE Access*, vol. 8, pp. 112783–112794, 2020, doi: 10.1109/ACCESS.2020.3002770.

[30] Q. Zhou, X. Jing, Y. He, *et al.*, "LSTM-based automatic modulation classification," in *Proc. IEEE Int. Symp. Broadband Multimedia Syst. Broadcast. (BMSB)*, IEEE, pp. 1–4, 2020.

[31] Y. Li, G. Shao, and B. Wang, "Automatic modulation classification based on bispectrum and CNN," in *Proc. IEEE 8th Joint Int. Inf. Technol. Artif. Intell. Conf. (ITAIC)*, IEEE, pp. 311–316, 2019.

[32] M. Patel, X. Wang, and S. Mao, "Data augmentation with conditional GAN for automatic modulation classification," in *Proc. 2nd ACM Workshop Wirel. Secur. Mach. Learn. (WiseML)*, pp. 31–36, 2020, doi: 10.1145/3395352.3402622.

[33] A. Jagannath and J. Jagannath, "Multi-task learning approach for automatic modulation and wireless signal classification," in *Proc. IEEE Int. Conf. Commun. (ICC)*, 2021, doi: 10.1109/ICC42927.2021.9500447.

[34] C. Hou, Y. Li, X. Chen, and J. Zhang, "Automatic modulation classification using KELM with joint features of CNN and LBP," *Phys. Commun.*, vol. 45, p. 101259, 2021, doi: 10.1016/j.phycom.2020.101259.

[35] A. H. Shah, A. H. Miry, and T. M. Salman, "Automatic modulation classification based deep learning with mixed feature," *Int. J. Electr. Comput. Eng.*, vol. 13, pp. 1647–1653, 2023, doi: 10.11591/ijece.v13i2.pp1647-1653.

[36] X. Fu, G. Gui, Y. Wang, *et al.*, "Automatic modulation classification based on decentralized learning and ensemble learning," *IEEE Trans. Veh. Technol.*, vol. 71, pp. 7942–7946, 2022, doi: 10.1109/TVT.2022.3164935.

[37] D. Anandkumar and R. G. Sangeetha, "A survey on performance enhancement in free space optical communication system through channel models and modulation techniques," Springer US, 2021.

[38] A. Nambisan, V. Gajjar, and K. Kosbar, "Scalogram aided automatic modulation classification," *Int. Found. Telemetering*, 2022.

[39] O. Almanza-Conejo, D. L. Almanza-Ojeda, J. L. Contreras-Hernandez, and M. A. Ibarra-Manzano, "Emotion recognition in EEG signals using the continuous wavelet transform and CNNs," *Neural Comput. Appl.*, vol. 35, pp. 1409–1422, 2023, doi: 10.1007/s00521-022-07843-9.

[40] H. S. Ghanem, M. R. Shoaib, S. El-Gazar, *et al.*, "Automatic modulation classification with 2D transforms and convolutional neural network," *Trans. Emerg. Telecommun. Technol.*, vol. 33, p. e4623, 2022.

[41] B. Mirzaei, B. Nikpour, and H. Nezamabadi-pour, "CDBH: A clustering and density-based hybrid approach for imbalanced data classification," *Expert Syst. Appl.*, vol. 164, p. 114035, 2021, doi: 10.1016/j.eswa.2020.114035.

[42] B. Mirzaei, F. Rahmati, and H. Nezamabadi-pour, "A score-based preprocessing technique for class imbalance problems," *Pattern Anal. Appl.*, vol. 25, pp. 913–931, 2022.

[43] R. Lin, W. Ren, X. Sun, *et al*., "A hybrid neural network for fast automatic modulation classification," *IEEE Access*, vol. 8, pp. 130314–130322, 2020, doi: 10.1109/ACCESS.2020.3009471.

[44] S. Mohsen, A. M. Ali, and A. Emam, "Automatic modulation recognition using CNN deep learning models," *Multimed. Tools Appl.*, 2023, doi: 10.1007/s11042-023-15814-y.

[45] K. Liu and F. Li, "Automatic modulation recognition based on a multiscale network with statistical features," *Phys. Commun.*, vol. 58, p. 102052, 2023, doi: 10.1016/j.phycom.2023.102052.

[46] L. Qian, H. Wu, T. Zhang, and X. Yang, "Research and implementation of modulation recognition based on cascaded feature fusion," *IET Commun.*, vol. 17, pp. 1037–1047, 2023, doi: 10.1049/cmu2.12604.