---

**Article:**
Sun, Y. orcid.org/0000-0001-6129-4290 and Ball, E. (2022) Automatic modulation classification using techniques from image classification. IET Communications, 16 (11). pp. 1303-1314. ISSN 1751-8628

https://doi.org/10.1049/cmu2.12335

---

Received: 23 July 2021 | Revised: 18 December 2021 | Accepted: 5 January 2022

IET Communications

DOI: 10.1049/cmu2.12335

The Institution of Engineering and Technology WILEY

ORIGINAL RESEARCH PAPER

# Automatic modulation classification using techniques from image classification

**Yilin Sun** | **Edward A Ball**

Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield, UK

**Correspondence**
Yilin Sun, Department of Electronic and Electrical Engineering, The University of Sheffield, Sheffield S1 4ET, UK
Email: ysun57@sheffield.ac.uk

**Abstract**

Automatic Modulation Classification (AMC) is a rapidly evolving technology, which can be employed in software defined radio structures, especially in 5G and 6G technology. Machine Learning (ML) can provide novel and efficient technology for modulation classification, especially for systems working in low signal to noise ratio (SNR). In this article, two dynamic systems not reliant on received signal phase lock and frequency lock are presented, with both employing ML to classify the modulation types for different received SNR. The first model is developed from the previous existing literatures, which utilises constellation images (CI) and image classification technology. Here, modulation types can be detected in a dynamic way without phase lock and frequency lock. In the second model, a new method named Graphic Representation of Features (GRF) is proposed, which represents the statistical features as a spider graph for ML. The concepts are tested and verified using simulations and RF data using a lab software defined radio (SDR). The results from the two models are compared. With the GRF techniques an overall classification accuracy of 59% is observed for 0 dB SNR and 86% at 10 dB SNR, compared to a random guess accuracy of 25%.

## 1 | INTRODUCTION

Digital communication systems continue to develop, with improvements to radio spectrum usage efficiency becoming vital. To respond to this requirement, dynamic spectrum access (DSA) is an important tool requiring spectrum sensing and signal classification. In this case, the modulation classification performs a significant role and can be widely employed in a variety of applications, such as software defined radio system and radar communication in the military [1, 2]. There is a high demand for radio frequency (RF) bands. In a crowded spectrum situation, the automatic modulation classification (AMC) [3] technology can respond to the requirements optimising signal detection and subsequent demodulation when multiple complex/unknown signals are to be handled, or for cognitive radio primary-user detection. Radio spectrum is a valuable resource, hence AMC is the way of identifying user activity in a spectrum and also identifying unused spectrum [4]. It can also be used by spectrum managers to ensure correct usage of spectrum. AMC can also be used by cognitive radio systems to identify the presence of primary users.

This work concentrates on modulation classification provided by different methods in the deep learning area. Signals can appear like noises at very low signal to noise ratio (SNR). The purpose of this work is to develop the accuracy of classification in the low SNR area with more efficient methods. With the improvement of Artificial Intelligence (AI) technology, many areas have achieved new progress by this novel area [5]. For the traditional statistical methods of machine learning, classification of the modulation types by the statistical features of the signals is common [6]. From this, deep learning (DL), working as a subproject of machine learning, reached a new step by applying the essence of biological information processing system [7]. DL also works as an efficient technique in image classification [8]. In this work, we provide two models using image classification from DL technology. The main difference between them is the training dataset. The first model utilises constellation images (CI) as the dataset and the second model utilises the graphic

representation of features (GRF). Based on the previous research, the CI model can classify the modulation types in a dynamic system without phase lock and frequency lock at low SNR levels. Transfer learning (TL) [9] is also introduced as an efficient AI technology in our model. We use the CI model as a development of the traditional method [10], to be compared to the GRF model.

In the second model, the GRF method shows advantages from both statistical features and image classification. Presenting the modulated signals in GRF, they can be fed to the pretrained network for the existing advanced image classifiers to test. To make a fair comparison, the same TL networks are also applied in this model. For the GRF, spider graphs of the statistical features provide a new way to classify the signals. Both models are utilised to detect the modulated signals from −10 to 20 dB SNR.

In previous published research, Maximum-likelihood decision theory is used as a critical method [11]. PSK and QAM signals are distinguished with accuracy of 90% above 9 dB SNR [12]. For the pattern recognition algorithms, the features of signals are involved in the models, where high order cumulants play a critical role in the AMC algorithm. After extracting the efficient features, a support vector machine (SVM) [13] is applied in the recognition process. The accuracy can reach 96 % at 10 dB SNR for 200 samples [14]. However, the probability of correct classification is between 50% and 70% at circa 0 dB SNR, suggesting a need for more improvement. Compared to the previous classifiers, the binary hierarchical polynomial classifiers are also proposed with the probability of correct classification of 56% at 0 dB SNR with 1000 symbols [15]. Hence, there is still the challenge to optimise the AMC system, especially at low SNR.

Our work introduces two AMC systems with DL technology. The first system is tested directly using the received CI. The second proposed system utilises the GRF, which benefits from the statistical features and image classification with DL. In both systems, convolutional neural networks, GoogleNet [16], SqueezeNet [17] and Inception-v3 [18] are trialled and compared. The main difference between the two systems is the training data and use of GRF. We develop the recognition images by extracting the statistical features for the GRF, whilst the constellation graphs are captured from the received signals directly. By using pre-existing neural networks aimed at generic image classification, we leverage this existing technology but repurpose it for modulation classification and thereby also assess its suitability.

There are four key contributions of this work. For our first contribution, we make the assessment of TL using the pretrained neural network for constellation recognition. The good usage of the TL reduces the complexity of calculation and improves the accuracy of modulation classification. The CI model uses the TL with constellation images without phase lock or frequency lock. For our second contribution, we develop and assess the novel method named graphic representation of features (GRF) to indicate the statistical features of modulation types, represented graphically and used as image classification data. This method takes the advantages of both statistical characterisation and image classification. As our third contribution, we compare the simulated data and lab collected data in our system. Our final contribution is the overall assessment of generic DL image classifiers as applied to communications AMC usage.

In the following parts, we will introduce the two system models, classification methods, and finally discuss results and performance.

## 2 | SYSTEM MODEL AND PROBLEM DESCRIPTION

In this section, the model of the received signal, the constellations and the statistical features are introduced. The constellations of the modulated signals are utilised for the DL model. The statistical features are extracted in graphical form for the GRF method.

### 2.1 | Signal model

The received signal in baseband is defined by $r(t)$ and it is given by (1):

$$(t) = s(t) + n(t) \tag{1}$$

Here, $(t)$ is the original signal which transmits through the additive white Gaussian noise (AWGN) channel, and $n(t)$ represents the noise applied to the signals.

For the demands of features calculation and analysis, the raw data should be represented by convention using in-phase and quadrature components ($I/Q$), (2).

$$a\ [i] = a_I\ [i] + j{*}a_Q\,[i] \tag{2}$$

Hence, the signals are composed of a real part and imaginary part, representing the characters of their constellation diagrams. All the statistical features and constellations are calculated and captured by this model. Furthermore, the data captured by lab test hardware is also read as $I/Q$ data [19].

The four kinds of modulated signals investigated and applied in this work are: BPSK, QPSK, 8PSK and QAM16. There are 100 symbols utilised, sampled at the rate of 50 samples per symbol.

SNR is also an important metric to describe the noise and represent the signals in the real world, here defined by (3):

$$SNR = \frac{\text{power of signal}}{\text{variance of noise}} = 20log_{10}\left\{\frac{\sqrt{\frac{1}{N}\sum |a\,[i]|^2}}{std\ (|\text{noise}|)}\right\} \tag{3}$$

In this work, signals are considered from −10 to 20 dB SNR. Note that the focus of previous works was on the classification above 10 dB SNR [20]. SNR values less than 0 dB are

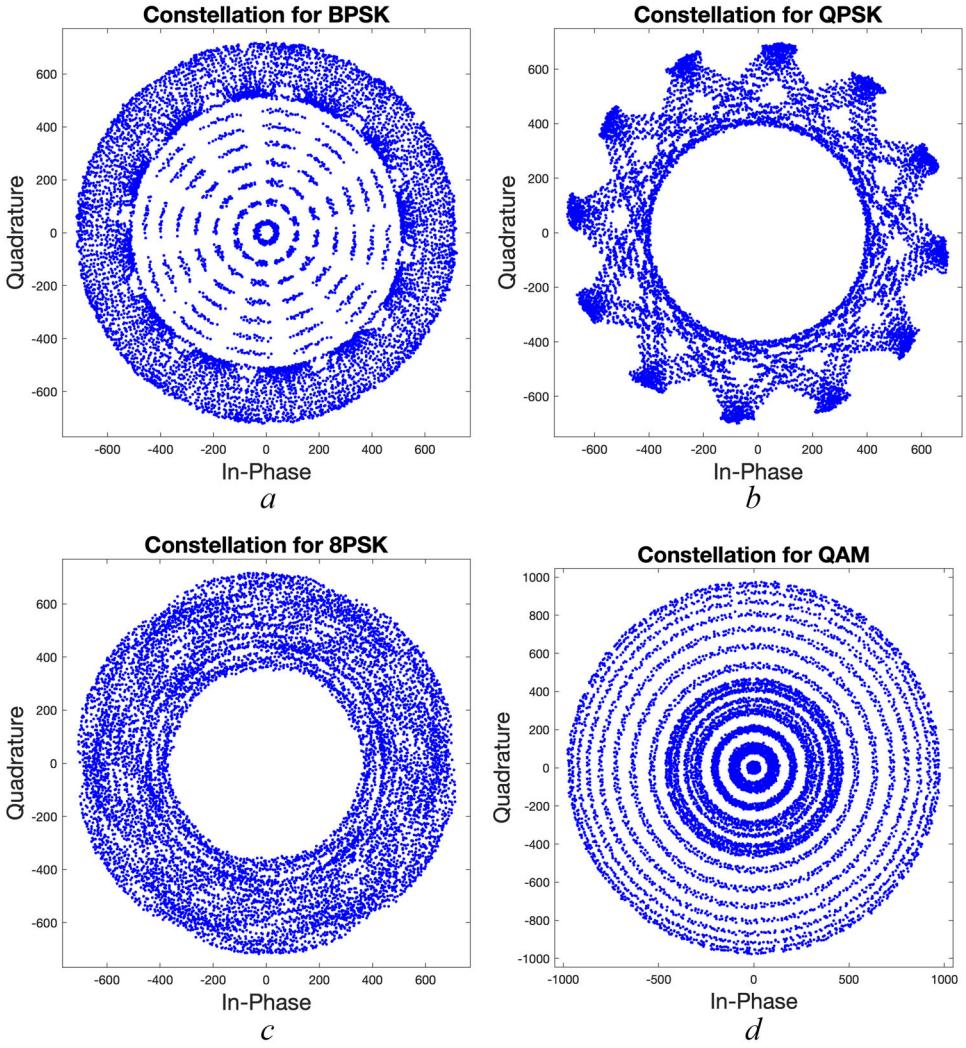


**FIGURE 1** Lab signals collected from SDR without phase and frequency lock (a) BPSK, (b) QPSK, (c) 8PSK, (d) QAM16

very important and should also be tested, to improve the performance in distinguishing between modulation types. We have developed our models from −10 to 20 dB SNR to provide a comprehensive dataset.

## 2.2 | Modulation constellations

For the DL model, the main principle applied is that of image classification. In applying DL, the modulated signals are transferred to the CI representations. We feed the network significant datasets, which are captured by increasing number of samples in each constellation. Figure 1 shows the constellations collected from the lab as examples of the four modulation types (BPSK, QPSK, 8PSK, QAM16) at 10 dB SNR.

Initially, for the CI model which we developed from the traditional method using constellations, our main contribution is we remove the requirement for phase and frequency lock. We generate the constellation data using a lab signal generator (Agilent E4437B) and collect it using an SDR (Ettus E310). The SNR effects and constellations can be observed in Figure 1, which also shows the received data rotation. In a dynamic system, this kind of data is not easy to train in the neural network especially at low SNR level.

To detect these kinds of constellations, we divided the data into small groups which can be observed in Figure 2. To improve the accuracy of classification, we also gradually increase and plot the samples from 200 to 500 samples in each small group which can be observed in Figure 3.

In neural networks with the dataset in the form of constellations, the constellations are captured in JPG format and stored with the labels of the modulation names. To obtain efficient CI for the DL, the sampled data are represented in small groups. In this way, the sampled dots and hence modulation-defining trajectories can be represented clearly, whilst avoiding the need of phase lock and symbol lock. This also allows the progression of the symbols to be captured, representing characteristic changes in the vector due to modulation symbol change separate to any phase lock issues. In our implementation, 2000 images were captured to feed the DL models. All the images are used as a dataset, with 70% of the images utilised for training, 20% of images for the validation process and the remaining 10% used for testing:

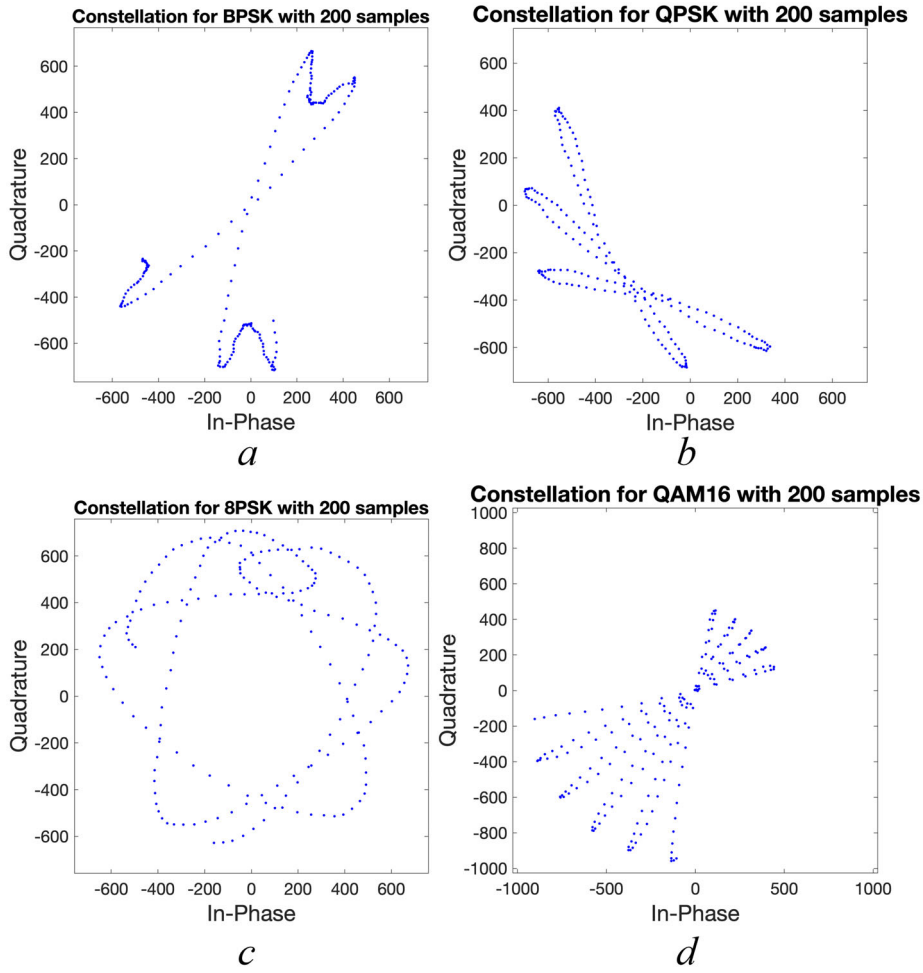


**FIGURE 2** Constellations of signal samples in small groups (a) BPSK, (b) QPSK, (c) 8PSK, (d) QAM16

Proportions as used by others [21]. Each DL model uses ten times iterative experiments. During these experiments, the models are trained to adjust the new dataset by getting feedback from the experiment results.

## 2.3 | Statistical features used for the spider graph representation

The common statistical methods of machine learning are proposed in the previous research, for classifying digital modulations [22]. The features proposed in the literature for artificial intelligence technologies were thus considered. The statistical models can obtain the results by capturing the statistical features from the four modulation types, from −10 to 20 dB SNR. The useful features are discussed below.

The signal power ratio of the in-phase and quadrature part, $\beta$ is extracted by (4):

$$\beta = \frac{\sum_n a_Q^2[i]}{\sum_n a_I^2[i]} \tag{4}$$

The standard deviation of the direct instantaneous phase, $\sigma_{dp}$, is extracted by (5):

$$\sigma_{dp} = \sqrt{\frac{1}{N}\left(\sum_{a_n[i]} \varphi_{NL}^2[i]\right) - \left(\frac{1}{N}\sum_{a_n[i]} \varphi_{NL}[i]\right)^2} \tag{5}$$

Here, $N$ is the number of the samples and $\varphi_{NL}[i]$ is the instantaneous phase which is defined by $\varphi_{NL}[i] = \tan^{-1}\frac{a_Q[i]}{a_I[i]}$.

The standard deviation of the signal instantaneous normalised amplitude, $\sigma_{aa}$, is extracted by (6) and (7):

$$\sigma_{aa} = \sqrt{\frac{1}{N}\left(\sum_{i=1}^{N} a_{cn}^2[i]\right) - \left(\frac{1}{N}\sum_{i=1}^{N} |a_{cn}[i]|\right)^2} \tag{6}$$

$$a_{cn}[i] = \frac{a[i]}{E(a[i])} - 1 \tag{7}$$

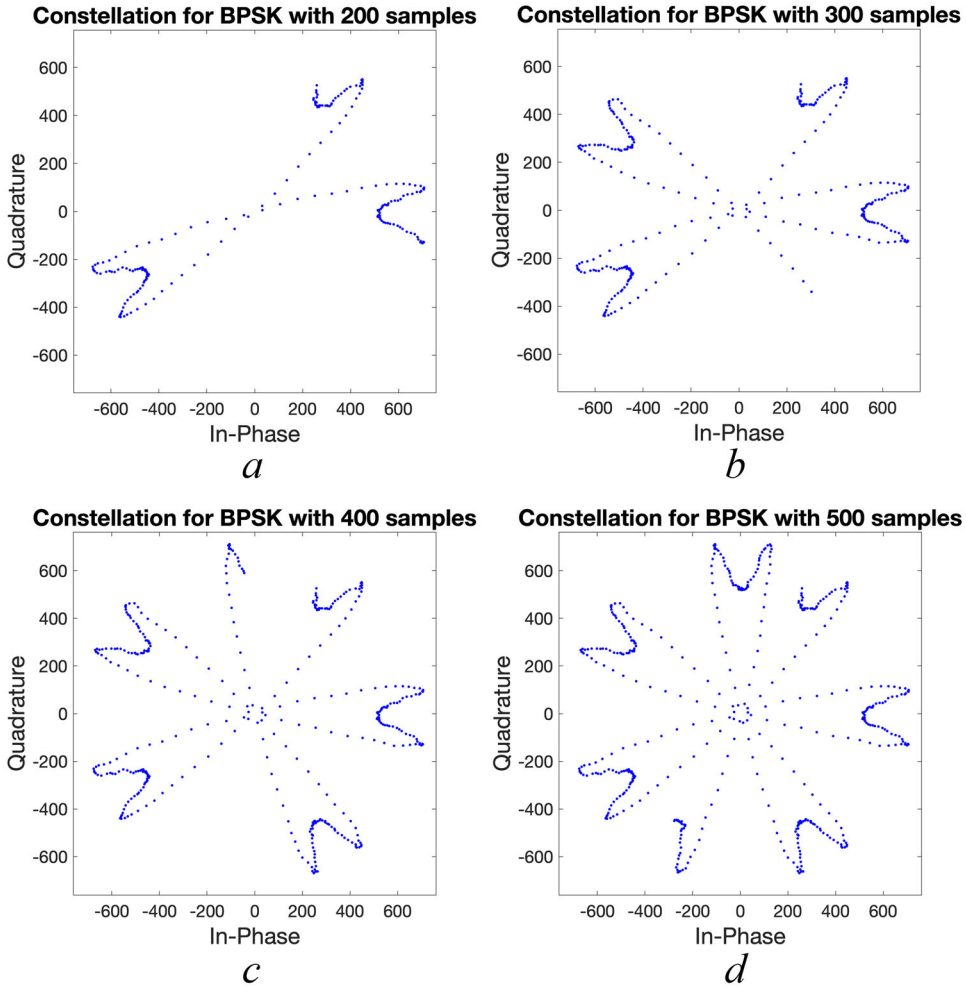


**FIGURE 3** Constellations of BPSK with increasing samples (a) BPSK with 200 samples, (b) BPSK with 300 samples, (c) BPSK with 400 samples, (d) BPSK with 500 samples

The standard deviation of the signal normalised amplitude of signal, $\sigma_v$, is extracted by (8) and (9):

$$\sigma_v = \sqrt{\frac{1}{N}\left(\sum_{i=1}^{N} a_v^2[i]\right) - \left(\frac{1}{N}\sum_{i=1}^{N}|a_v[i]|\right)^2} \tag{8}$$

$$a_v[i] = \sqrt{\frac{a[i]}{var(a[i])}} - 1 \tag{9}$$

The standard deviation can describe the statistical dispersion, with the above three types of standard deviation results utilised for describing the signals in the main characteristics of phase and amplitude [23].

The mixed order moment of signals, $v_{20}$, is extracted by (10):

$$v_{20} = \frac{M_{42}(y)}{M_{21}^2(y)} = \frac{E\left(|a[i]|^4\right)}{E\left(|a[i]|^2\right)} \tag{10}$$

The Mixed order moment is defined by the fourth order moment $M_{42}$ and second order moment $M_{21}$ of the signal. This feature employs the Joint Power Estimation and Modulation Classification (JPEMC) algorithm [20], which is associated with the power of signal and noise. This feature can represent the received power.

The mean value of the signal samples, $X$, is extracted by (11):

$$X = \frac{1}{N}\sum_{n=1}^{N}|a[i]| \tag{11}$$

The normalised square root of signal, $X_2$, is extracted by (12):

$$X_2 = \frac{\sqrt{\sum_{i=1}^{N}|a[i]|}}{N} \tag{12}$$

The normalised square root of signal can describe the amplitude scale of the signals.

The maximum value of power spectral density (PSD) is defined as $\gamma_{max}$, (13):

$$\gamma_{max} = \frac{1}{N} max|DFT\ (a_{cn}\ [i])|^2 \tag{13}$$

Parameter $\gamma_{max}$ relates to the power spectrum of signals in the frequency domain, using the discrete Fourier transform (DFT).

The Cumulants of signals are extracted by (14)–(20):

$$C_{20} = E\left[a^2\ [n]\right] \tag{14}$$

$$C_{21} = E\left[|a\ [n]|^2\right] \tag{15}$$

$$C_{40} = M_{41} - 3M_{20}^2 \tag{16}$$

$$C_{41} = M_{41} - 3M_{20}M_{21} \tag{17}$$

$$C_{42} = M_{42} - |M_{20}|^2 - 2M_{21}^2 \tag{18}$$

$$C_{63} = M_{63} - 6M_{20}M_{40} - 9M_{42}M_{21} + 18M_{20}^2M_{21} + 12M_{21}^3 \tag{19}$$

$$C_{80} = M_{80} - 35M_{40}^2 - 28M_{60}M_{20} + 420M_{40} - 630M_{20}^4 \tag{20}$$

$M_{p+q,p}$ can be defined by $E[a[n]^p a[n]^{*q}]$. Cumulants and moments are widely used in a variety of classification techniques, especially the high-order cumulants, which can reduce the influence of additive white Gaussian noise (AWGN) [24]. Cumulants are the coefficients of Maclaurin series of the generating function. They are proposed to represent the moments and the moments can measure functions

quantitatively [25]. When the higher order cumulants of the received signals are calculated, the effect of the Gaussian noise is removed and the signals are independent [26].

The Kurtosis of signals, *K*, is extracted by (21):

$$K = \left| \frac{E\left[(a - E\ [a])^4\right]}{E\left[(a - E\ [a])^2\right]^2} \right| \tag{21}$$

Kurtosis is a feature which can describe the steepness or flatness of the distribution of signals [27].

The Skewness of signals, *S*, is extracted by (22):

$$S = \left| \frac{E\left[(a - E\ [a])^3\right]}{E\left[(a - E\ [a])^2\right]^{3/2}} \right| \tag{22}$$

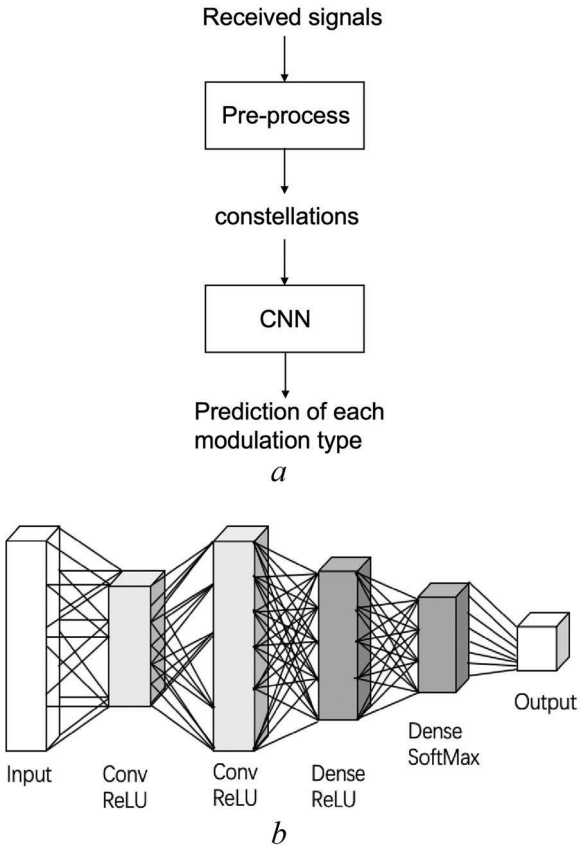


**FIGURE 4** DL system with CI (a) system structure, (b) structure of neural network

Skewness can describe the position of the tapering side of the distribution, which is also the third central moment. Both the Kurtosis and Skewness are used to assess the shape of the signal's distribution.

The ratio of peak-to-RMS, *PR*, is extracted by (23):

$$PR = \frac{max|a|^2}{\frac{1}{N}\sum_{i=1}^{N}(a\ [i])^2} \tag{23}$$

The ratio of peak-to-average of the signal, *PA*, is extracted by (24):

$$PA = \frac{max\ |a|}{\frac{1}{N}\sum_{i=1}^{N} a\ [i]} \tag{24}$$

The two features, *PA* and *PR*, further aim to describe aspects of the shape of different signals in detail.

In the next section, we test the above statistical features using the collected data, to find the appropriate and necessary features to build the GRF system.

## 3 | CLASSIFICATION METHOD

In this section, the classification methods are introduced. The first DL model achieves the classification based on the CI. To represent and support detection of the dynamic signals, the

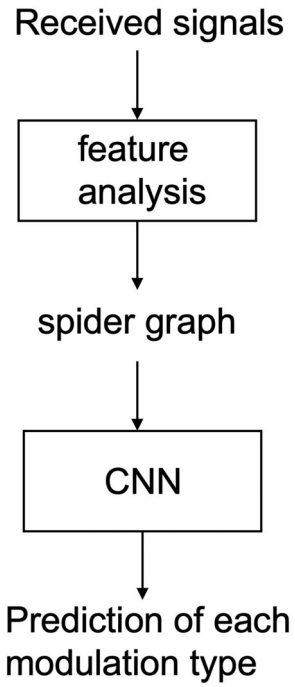


**FIGURE 5** DL system with GRF

training dataset includes many constellations in different situations. The second DL model is based on the GRF images from the modulated signals. The spider graphs are developed from the statistical features of the collected data, which can describe the characteristics in a stable way. In both models, we trained with four different structured CNNs to find the one with best accuracy at low SNR levels.

## 3.1 | Deep learning with CI

In this section, four CNN models are introduced, a simple CNN developed from the Iris case [28] with multiple layers, the SqueezeNet model [17], the GoogleNet model and the Inception-v3 model [29]. The last three models are pretrained networks, which are used for TL. The DL system with CI is shown in Figure 4a. We pre-process the signals and obtain the constellations, which are sent to the CNN.

As Figure 4(b) shows, an example of a simple CNN structure can consist of convolutional layers and dense fully connected layers. The characteristics of the signals can be captured by the convolutional layers and the feature map can be produced by these layers. Subsequently, the features can be learned by the sequential layers. Finally, the results of the classification are outputted.

### 3.1.1 | CNN

This model is created as a development from the Iris recognition case. There are fifteen layers in total. Three convolutional layers are followed by the batch normalisation layers and ReLu layers. Two max-pooling layers are set between the other three blocks, which can down-sample the input data and reduce the risk of overfitting. The convolutional layers can capture the physical features of the images by the filters, such as the profile and grayscale of the images. In this way, the convolutional layer is significant in being able to influence the classification of the images. The batch normalisation layers can normalise the input channel. The ReLu layers calculate the threshold for the elements [30].

### 3.1.2 | Transfer learning models

The SqueezeNet has 68 layers, and its coefficients of the network are 227-by-227 pixels [17]. The GoogleNet has 144 layers and is pretrained to classify the images into 1000 categories and 365 places. The input image size is 224-by-224 pixels [16]. The Inception-v3 has 315 layers and was trained by over a million images to classify into 1000 categories [18]. The input size is 299-by-299 pixels. Images need to be modified in

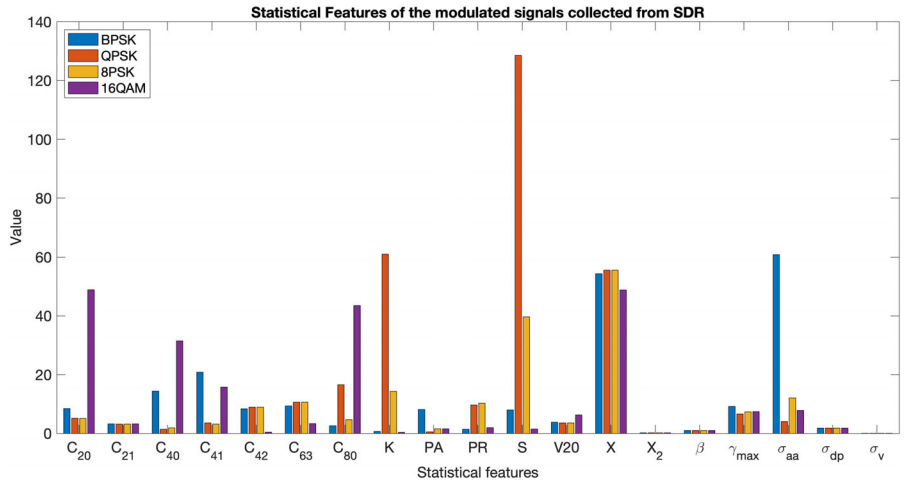


**FIGURE 6** Features of the tested modulated signals at 10 dB SNR

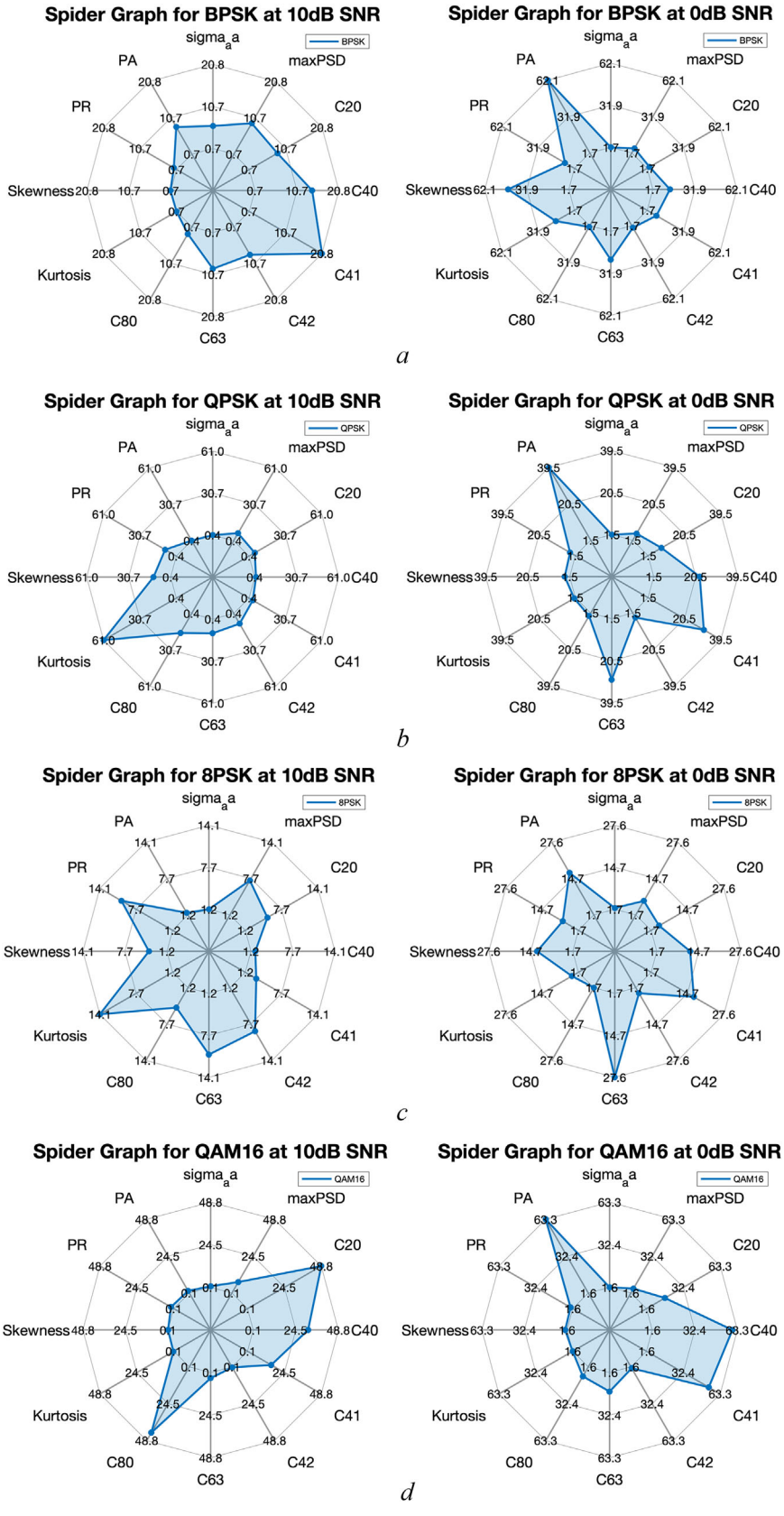


**FIGURE 7** Spider Graphs of modulated signals at 10 and 0 dB SNR (a) BPSK, (b) QPSK, (c) 8PSK, (d) QAM16

size before feeding to models, to fit the input size criteria. The final learning layers are modified, to provide only four outputs, corresponding to each of the modulation types. In our experiments, 70% of images are used as training and the rest are used as validation. We rotate the images in −90° to 90° steps (representing constellation phase errors) and rescale images from 1 to 1.5 randomly (representing amplitude variations), which can help improve the amount of training data and avoid overfitting [10].

### 3.2 | Graphical representation of features (GRF)

For the DL with GRF, the statistical features need to be analysed and selected to work in the model. GRF is a new method to plot the statistical characteristics graphically. In this method, we calculated the data from the dynamic receiving system without phase and frequency lock, to indicate the signal characteristics in a stable way. By using the images with GRF, we can use the pretrained network with existing advanced image classifiers. The classification system can be observed in Figure 5. The received signals are analysed and extracted as the spider graphs, which efficiently represent the graphical features.

Based on the formulas of features from Section 2.3, the features from the four kinds of modulation types were calculated. Figure 6 shows the statistical features of the four modulated signals at 10 dB SNR. To show the characteristics of all the features in one graph, some of the features are displayed in logarithmic form, such as $\sigma_{aa}$, $X_2$, $\gamma_{max}$, which is a clear way to visualise the magnitude data differences. According to Figure 6, it is obvious that $\beta$, $\sigma_{dp}$, $\sigma_v$ and $X_2$. cannot help to classify the modulations (BPSK, QPSK, 8PSK and QAM16) - these four features remain constant around fixed values.

Figure 7 indicates how the selected statistical features are applied in the spider graph. Based on the previous analysis, we choose twelve features of the modulated signals. We use the features as the axes of the graph, each modulation type can then display the values on the graph. The spider graphs make the statistical features into a single visual representation. For the different magnitudes of the feature values, a log function is employed where appropriate. As seen in Figure 7, at 10 and 0 dB SNR levels, each modulated signal shows a different form of representation. After collecting the graphs, all the image forms are used for our neural network based on image classification.

## 4 | EVALUATED PERFORMANCE

This section presents the results from the different classification methods. All the detecting systems are based on the DL technology. The previously described systems use MATLAB simulation data and lab data collected from an SDR. In both systems, different DL structures are compared. MATLAB simulation data and lab measured data are tested in this section.

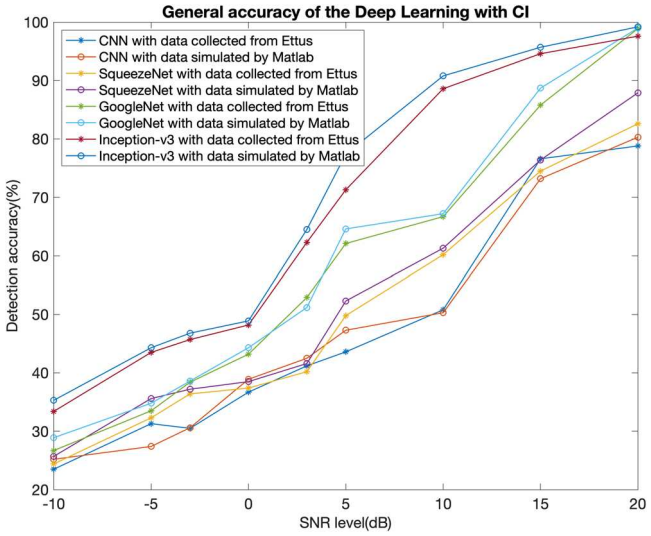


**FIGURE 8** General accuracy of DL with CI over different SNR levels (all modulation types)

The lab equipment used was an Agilent E4437B signal generator and Ettus Research E310 SDR. We baseband-modulated the RF signals using a PN15 sequence. The symbol rate used was 100 ksps and the sample rate was 2 MHz.

### 4.1 | Deep learning with CI

In Figure 8, the results are shown from the four CNN models. The CNN developed using the Iris case performs worse than other traditional DL methods. This is likely due to the structures and coefficients of this CNN variant being potentially very sensitive and thus significantly influencing the classification system.

The three other models are also forms of CNN, but they have more complicated structures, and are pretrained by millions of generic images. Although the images applied for pretraining are not the constellations of modulated signals, they can also help the models to classify the images by modifying the coefficient iteratively and repeatedly.

Figure 8 shows the general accuracy of each DL model at different SNR levels. The data are also collected in two ways: Simulation in MATLAB and lab signals from the Ettus SDR. When in the low SNR area, both sources of the data have significant noise. That is why the accuracy of both are quite similar and low. From Figure 8 we can find the Inception-v3 compared to the CNN, SqueezeNet and GoogleNet shows best performance over the whole detection range.

Figure 9 presents examples for the classification results from DL with CI at 10, 0, −10 dB SNR. Figure 9 also shows the probability of modulation recognition. Based on the results provided in Figure 8, we tested the data in the Inception-v3 model. At 10 dB SNR, the classification is valid according to Figure 9a. In general, the random guess detection rate would be 25% (since there are four possible modulation types). This model also provides accuracy at −10 dB SNR slightly higher than that from a

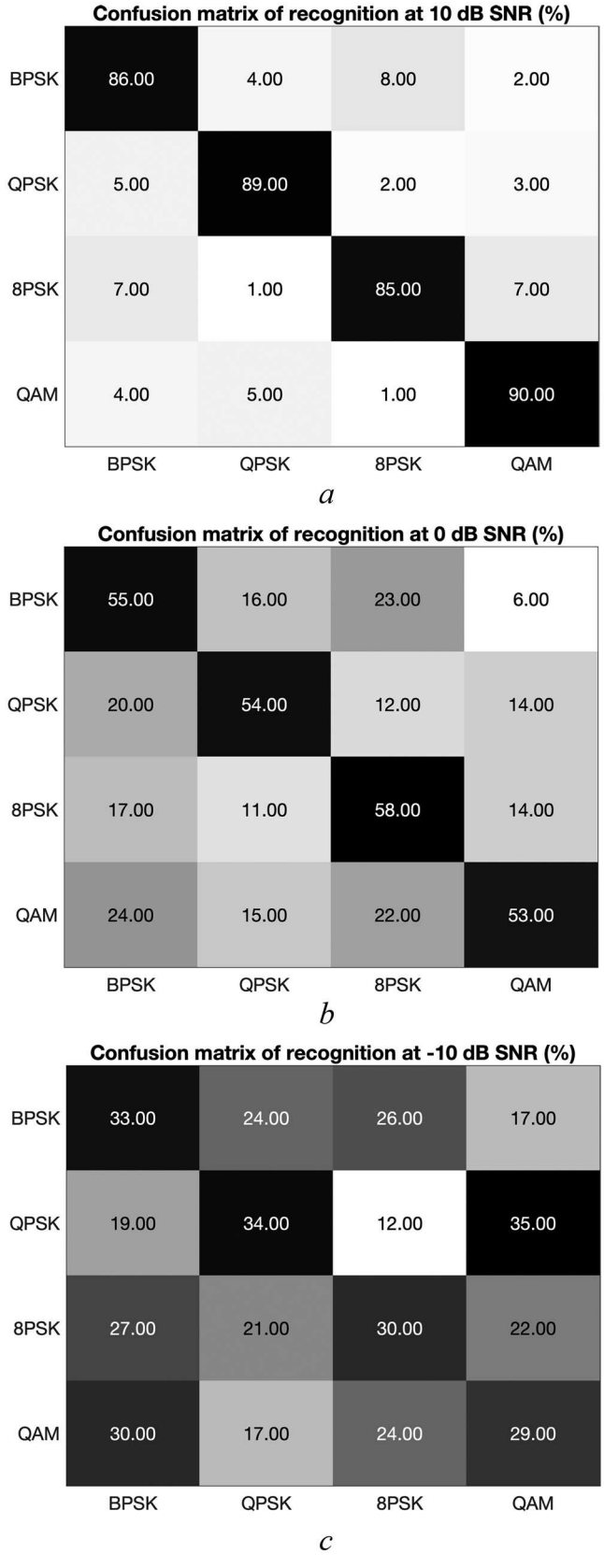


**FIGURE 9** Accuracy of DL with CI (Lab data). (a) Confusion matrix of recognition at 10 dB SNR, (b) confusion matrix of recognition at 0 dB SNR, (c) confusion matrix of recognition at −10 dB SNR

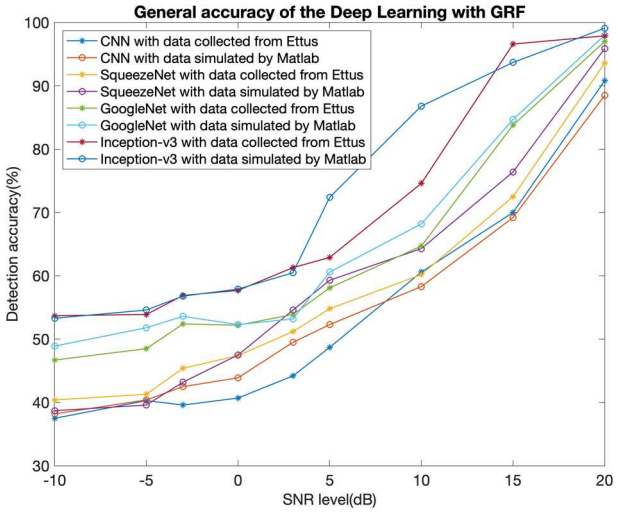


**FIGURE 10** General accuracy of DL with GRF over different SNR levels (all modulation types)

random guess. We use the simulated data as a reference in the previous experiment to choose the best neural network structure, so Figure 9 utilises the data collected from Ettus SDR, representing real-world signals. From Figure 9a–c, we can also see the details of each prediction.

## 4.2 | Deep learning with GRF

The high order cumulants were employed for this system, as the second classification process to be trialled. The Kurtosis, Skewness, PR, and PA can describe the shape of the signals. According to Figure 6 the standard deviation of the signal instantaneous normalised amplitude, $\sigma_{aa}$ and the maximum value of power spectral density, $\gamma_{max}$ can also help to classify the modulation types. In the training data, the signals were varied from −10 to 20 dB SNR, which caused the features to also change dramatically. These changes of the appropriate features result in the spider graph having different shapes for different modulation types, as expected.

Figure 10 shows the general results of GRF with different neural network structures. Two kinds of training data are collected and employed in the same network and the results are compared: (1) from MATLAB simulation and (2) the SDR platform in the lab. At low SNR levels, both measurement and simulation produce similar accuracy. Compared to the probability of a random guess (25%), the detection accuracy is significantly better for the GRF DL approach. The Inception-v3 model also shows the best classification ability.

The accuracy was improved (especially at low SNR which reaches over 50 % with Inception-v3 model) by training with the whole SNR range of data. The prediction results of the four kinds of modulation types at 10, 0 and −10 dB SNR are shown in Figure 11. We utilise the data collected from Ettus SDR, representing real-world signals in Figure 11.

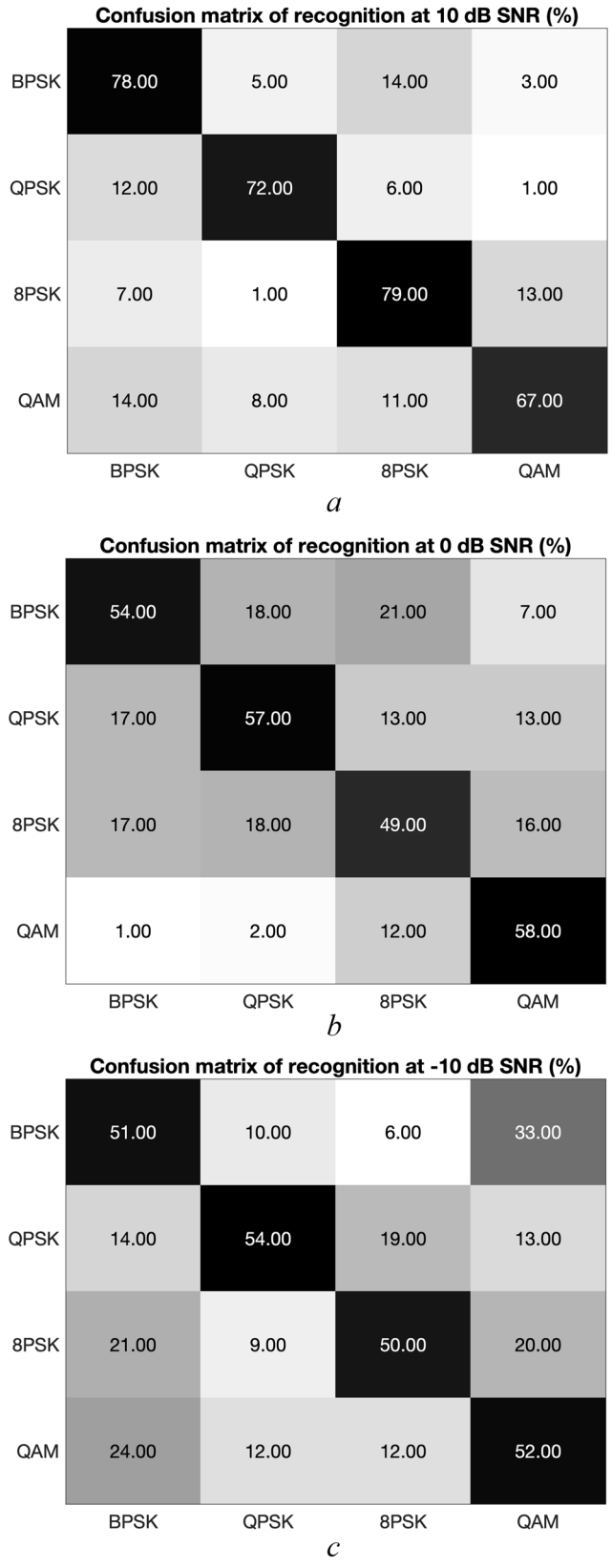


**FIGURE 11** Accuracy of DL with GRF (Lab data). (a) Confusion matrix of recognition at 10 dB SNR, (b) confusion matrix of recognition at 0 dB SNR, (c) confusion matrix of recognition at −10 dB SNR

**TABLE 1** Detection accuracy of DL with GRF

| Model name | Detection accuracy |
|---|---|
| DL with CI | 65.1 % |
| DL with GRF | 69.3 % |

**TABLE 2** Comparison of detection accuracy at 0 dB SNR

| Model name | Detection accuracy |
|---|---|
| SVM [14] | 50% |
| Binary Hierarchical polynomial classifier [15] | 56% |
| This work: DL with GRF | 59% |

There is still a limitation for the proposed techniques here. The features need to use a log function, or other pre-processing, to make the value fit the graph for imaging, which will influence the classification. However, we propose this to be a minor issue.

## 5 | DISCUSSION

From Figure 8, when the SNR is more than 10 dB, the detection accuracy of constellation-based classification achieves close to 100%. However, from Figure 10, the detection accuracy of DL with GRF is not as good as CI model at high SNR levels. After comparing Figures 8 and 10, we also find key results in low SNR area. The detection accuracy of CI classification increases from around 35% at −10 dB SNR to 50% at 0 dB SNR. However, the detection of accuracy of DL with GRF increases from 52% at −10 dB SNR to 59% at 0 dB SNR. DL with GRF has shown its advantages by our results here, for low SNR use cases. Hence, to improve the classification accuracy in low SNR, we can use the GRF classification as a supplement to CI recognition.

Table 1 shows the detection accuracy of different systems over all the tested SNR range. From this, we can find the DL system with GRF has the highest accuracy 69.3%. But from Figure 8, the Inception-v3 model with constellations can achieve near 100% for SNRs greater than 10 dB SNR and so this model provides the best results in the high SNR area. The DL method with the GRF performs well in our general results, especially for low SNR. Table 2 compares the detection results at 0 dB SNR between the GRF model and previous published works [14, 15]. There is an improvement of accuracy in our work. Note that in our systems we also explore below 0 dB SNR, which others do not.

These methods are all based on image classification, but with different datasets: CI and GRF. The GRF system has an efficient use of statistical features, requiring only twelve features yet achieving 69.3% detection accuracy overall (and significantly better at lower SNR, as discussed above) with the random guess accuracy of the system being 25%. The TL systems detect the constellation dataset directly which reduces the mathematical complexity.

## 6 | CONCLUSIONS

In this work, we discussed different classification models based on DL, applied to modulation recognition in the dynamic receiving system without phase and frequency lock. Firstly, we developed the traditional method: Classification with CI. After that, we presented the GRF method. Both methods show their advantages. In high SNR scenarios, the Inception-v3 with constellation model reaches the highest accuracy. But for low SNR levels, the GRF provides the superior accuracy. Therefore, for our future work, we are exploring a hybrid approach which could combine the advantages of both techniques to achieve better overall performance over a wide SNR range.

### FUNDING INFORMATION
None

### CONFLICT OF INTEREST
None

### DATA AVAILABILITY STATEMENT
Author elects to not share data

### ORCID
*Yilin Sun* https://orcid.org/0000-0001-6129-4290
*Edward A Ball* https://orcid.org/0000-0002-6283-5949

### REFERENCES

1. Kulin, M., Kazaz, T., Moerman, I., De Poorter, E.: End-to-End learning from spectrum data: A deep learning approach for wireless signal identification in spectrum monitoring applications. IEEE Access 6, 18484–18501 (2018)
2. Hamid, M., Slimane, S.B., Moer, W.V., Björsell, N.: Spectrum sensing challenges: Blind sensing and sensing optimization. IEEE Instrum. Meas. Mag. 19(2), 44–52, (2016)
3. Zhechen, Z., Asoke, K.N.: Automatic Modulation Classification : Principles, Algorithms and Applications. John Wiley & Sons, New York (2015)
4. Hindia, M.N., Qamar, F., Ojukwu, H., Dimyati, K., Al-Samman, A.M., Amiri, I.S.: On platform to enable the cognitive radio over 5G Networks. Wirel. Pers. Commun. 113(2), 1241–1262 (2020)
5. West, N.E., O'Shea, T.: Deep architectures for modulation recognition. In: Proceedings of the 2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN), Baltimore, USA, pp. 1–6 (2017)
6. Kim, J., Lee, B., Lee, H., Kim, Y., Lee, J.: Deep learning-assisted multidimensional modulation and resource mapping for advanced OFDM systems. In: Proceedings of the 2018 IEEE Globecom Workshops (GC Wkshps), Abu Dhabi, United Arab Emirates, pp. 1–6 (2019)
7. Schmidhuber, J.: Deep Learning in neural networks: An overview. Neural Networks 61, 85–117, (2015)
8. Chen, H., Wang, Z., Zhang, L.: Collaborative spectrum sensing for illegal drone detection: A deep learning-based image classification perspective. China Commun. 17(2), 1–92 (2020)
9. Gao, Y., Mosalam, K.M.: Deep transfer learning for image-based structural damage recognition. Comput. Civ. Infrastruct. Eng. 33(9), 748–768 (2018)
10. Peng, S., Jiang, H., Wang, H., Alwageed, H., Yao, Y.D.: Modulation classification using convolutional neural network based deep learning model. In: Proceedings of the 26th Wireless and Optical Communication Conference (WOCC), Newark, NJ, pp. 1–5 (2017)
11. Sills, J.A.: Maximum-likelihood modulation classification. In: Proceedings of the MILCOM 1999. IEEE Military Communications, Atlantic City, USA, pp. 217–220 (1999)
12. Whelchel, J.E., McNeill, D.L., Hughes, R.D., Loos, M.M.: Signal understanding: An artificial intelligence approach to modulation classification. In: Proceedings of the IEEE International Workshop on Tools for Artificial Intelligence, Fairfax, USA, pp. 231–236 (1989)
13. Corinna, C.; Vladimir, V.: Support-Vector Networks. Mach. Learn. 20, 273–297 (1995)
14. Gang, H., Jiandong, L., Donghua, L.: Study of modulation recognition based on HOCs and SVM. In: Proceedings of the IEEE 59th Vehicular Technology Conference, Milan, Italy, pp. 898–902 (2004)
15. Abdelmutalab, A., Assaleh, K., El-Tarhuni, M.: Automatic modulation classification based on high order cumulants and hierarchical polynomial classifiers. Phys. Commun. 21, 10–18 (2016)
16. Szegedy, C., Liu, W., Jia, Y., et al.: Going deeper with convolutions. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, pp. 1–9 (2015)
17. Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J. & Keutzer, K.: SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size. 2017 The International Conference on Learning Representations, (2017).
18. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, pp. 2818–2826 (2016)
19. O'Shea, T.J., Roy, T., Clancy, T.C.: Over-the-air deep learning based radio signal classification. IEEE J. Sel. Top. Signal Process. 12(1), 168–179 (2018)
20. Chen, K.Z., Hu, A.Q.: MPSK demodulation algorithm based on pattern recognition. In: Proceedings of the International Conference on Neural Networks and Signal Processing, Nanjing, China, pp. 182–186 (2008)
21. O'Shea, T., Hoydis, J.: An introduction to deep learning for the physical layer. IEEE Trans. Cogn. Commun. Netw. 3(4), 563–575 (2017)
22. Zhang, M., Diao, M., Guo, L.: Convolutional neural networks for automatic cognitive radio waveform recognition. IEEE Access 5, 11074–11082 (2017)
23. Rajendran, S., Meert, W., Giustiniano, D., Lenders, V., Pollin, S.: Deep learning models for wireless signal classification with distributed low-cost spectrum sensors. IEEE Trans. Cogn. Commun. Netw. 4(3), 433–445 (2018)
24. An, N., Li, B., Huang, M.: Modulation classification of higher order MQAM signals using mixed-order moments and fisher criterion. In: Proceedings of the 2nd International Conference on Computer and Automation Engineering (ICCAE), Singapore, 150–153 (2010)
25. Spooner, C.M.: On the utility of sixth-order cyclic cumulants for RF signal classification. In: Proceedings of the Conference Record of Thirty-Fifth Asilomar Conference on Signals, Systems and Computers, Boston, MA, USA, pp. 890–897 (2001)
26. Dobre, O.A., Bar-Ness, Y., Su, W.: Higher-order cyclic cumulants for high order modulation classification. In: Proceedings of the IEEE Military Communications Conference, Boston, MA, pp. 112–117 (2003)
27. Lee, J., Kim, B., Kim, J., Yoon, D., Choi, J.W.: Deep neural network-based blind modulation classification for fading channels. In: Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Republic of Korea, pp. 551–554 (2017)
28. Fisher R. A.: The use of multiple measurements in taxonomic problems. Annals of Eugenics. 7(2), 179–188 (1936). https://doi.org/10.1111/j.1469-1809.1936.tb02137.x
29. Grm, K., Struc, V., Artiges, A., Caron, M., Ekenel, H.K.: Strengths and weaknesses of deep learning models for face recognition against image degradations. IET Biometrics 7(1), 81–89 (2018)
30. Alsina-Pages, R.M., Hervas, M., Vilasis-Cardona, X., Vinyoles-Serra, M.: QPSK demodulation using cellular neural networks. In: Proceedings of the 14th International Workshop on Cellular Nanoscale Networks and their Applications (CNNA), Notre Dame, pp. 1–2 (2014)

**How to cite this article:** Sun Y., Ball E.A.: Automatic modulation classification using techniques from image classification. IET Commun. 1–12 (2022). https://doi.org/10.1049/cmu2.12335