**Proceedings Paper:**

Sun, Y. orcid.org/0000-0001-6129-4290 and Ball, E. (2022) Automatic modulation classification based on machine learning. In: Laribi, M.A., Carbone, G. and Jiang, Z., (eds.) Advances in Automation, Mechanical and Design Engineering: SAMDE 2021. 2021 International Symposium on Automation, Mechanical and Design Engineering: SAMDE 2021, 03-05 Dec 2021, Beijing, China. Mechanisms and Machine Science . Springer Cham , pp. 53-67. ISBN 9783031099083

https://doi.org/10.1007/978-3-031-09909-0_5

# Automatic Modulation Classification Based on Machine Learning

**Yilin Sun, Eddie Ball**
The University of Sheffield

Ysun57@sheffield.ac.uk

**Abstract**. Automatic Modulation Classification (AMC) is a rapidly evolving technology, which can be employed in software defined radio structures, such as for military communication. Machine Learning can provide novel and efficient technology for modulation classification, especially for systems working in low Signal to Noise Ratio (SNR). For this work, a dynamic modulation classification system without phase lock is trialed. The signals are captured with different SNR and duration. Traditional Machine Learning based on the mathematical features is compared with Deep Learning based on the constellations. Based on these two methods, a hybrid model is provided. This model involved the novel Deep Learning at first and the feature classification as a supplement, which achieves good performance at low SNR area.

## 1. Introduction

As digital communication systems develop, improving radio spectrum usage efficiency becomes vital. To respond to this requirement, Dynamic Spectrum Access (DSA) is a good starting point, with 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. There is a high demand of radio frequency (RF) bands. In the crowded spectrum situation, the AMC technology can respond to the requirements, optimizing signal demodulation, information extraction and interference detection without limitation from various complex emitters [1][2].

This work concentrates on modulation classification provided by different methods in the machine learning area. With the improvement of the Artificial Intelligence (AI) technology, many areas have achieved new progress by this novel area. For the traditional statistical methods of machine learning, classification of the modulation types by the statistical features of the signals is common. The most suitable model will be selected after the experiments of the calculation and comparison among SVM, K-Nearest-Neighbors (KNN), Decision Tree, and so on. After that, the Deep Learning working as a subproject of machine learning reached a new step, which extracts the essence from the biological information processing system. The advantages are also obviously shown in audio recognition, image classification and semantic segmentation. In this paper, the theory of the image classification will be utilized, Convolutional

Neural Network (CNN) of deep learning is introduced and tested. Based on this novel method, the feature classification based on machine learning works as a supplement to improve accuracy. This new hybrid model detects the modulated signals from -10dB SNR to 20dB SNR.

In the previous research, Maximum-likelihood decision theory is used as a critical method. PSK and QAM signals are distinguished with accuracy of 90% above 9dB SNR [3]. For the pattern 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, the Support Vector Machines (SVM) is applied in the recognition process. The accuracy can reach to 96% around 10 dB SNR for 200 samples [4]. But the probability of correct classification is between 50% to 70% at around 0dB SNR, which still needs to be improved. Compared to the previous classifiers, the binary hierarchical polynomial classifiers are also proposed with the probability of correct classification of 65% [5]. There is still the challenge to optimize the AMC system, especially at low SNR.

This work will introduce the machine learning and deep learning methods. Images are captured by the dynamic system and used as training data. In this paper, the CNN model with the best accuracy and the high order cumulants will be involved in the hybrid system. The results at low SNR will also be analyzed in the following works.

## 2. Problem Presentation

In this section, the model of the signal, the statistical features and the constellations are introduced. The statistical features are used for the machine learning models and the hybrid model. The constellations of the modulated signals are utilized for the deep learning models and the hybrid model.

### *2.1. Signal model*

The received signal in baseband is defined with r(t) and it is given by

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

$s(t)$ is the original signal which transmit through the additive white Gaussian noise (AWGN) channel.

With the demand of features calculation and analysis, the raw data should be represented by in-phase and quadrature components

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

In this case, signals simulated in software are composed of 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 hardware is also read as I/Q data. Building the model in this way can simulate the signals expected in reality.

Four kinds of modulated signals, BPSK, QPSK, 8PSK and QAM16 are applied in this work. There are 100 symbols utilized, sampled by the rate of 50 samples per symbol.

SNR is also an important index to describe the noise and imitate the signals in the real world. And it is defined by

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

In this work, signals are considered from -10dB to 20dB SNR. According to the previous works, focus was on the classification above 10dB SNR [6]. The SNR value less than 0dB should also be tested, to improve the performance of distinguishing modulation types. We have developed the models from -10dB to 20dB SNR to provide a comprehensive dataset.

### *2.2. Statistical features for Machine Learning and Hybrid Model*

The traditional statistical methods of machine learning are proposed in the previous research to classify the digital modulations [7]. The features proposed in the literature for artificial intelligence technologies are now considered. The statistical models can obtain the results by capturing the statistical features among the four modulation types from -10dB to 20dB SNR. The useful features are as follows.

Signal power ratio of in-phase and quadrature part $\beta$

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

Standard deviation of the direct instantaneous phase $\sigma_{dp}$

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

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

Standard deviation of the signal instantaneous normalized amplitude $\sigma_{aa}$

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

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

Standard deviation of the signal normalized amplitude of signal $\sigma_v$

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

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

Standard deviation can describe the statistical dispersion, which the above three kinds of standard deviation results are utilized for describing the signals in the main characteristics of phase and amplitude.

Mixed order moment $v_{20}$

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

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, which is associated with the power of signal and noise. This feature can reflect the power in another way [7].

Mean value of the signal samples $X$

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

Normalized square root of signal $X_2$

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

This feature can describe the amplitude scale of the signals.

Maximum value of power spectral density (PSD) $\gamma_{max}$

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

This feature describes the situation of signals in the frequency domain, by using the Discrete Fourier Transform (DFT).

Cumulants

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

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

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

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

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

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

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

$M_{p+q,p}$ could be defined by $E[a[n]^p a[n]^{*q}]$. Cumulants and moments are widely used in variety of classification, especially for the high-order cumulants, which can avoid the influence from the AWGN [9] [10]. Cumulants are proposed to represent the moments in an alternative way and the moments can measure functions quantitively. When the higher order cumulants of the received signals are calculated, the part of the Gaussian noise is zero and the signals are independent [11].

Kurtosis K

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

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

Skewness

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

Skewness can describe the position of the tapering side of the distribution, which is also the third central moment. Both the abilities of kurtosis and skewness are to confirm the shape of the signals [6].

Ratio of peak-to-rms, PR

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

Ratio of peak-to-average, PA

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

These two features aim to describe the shape of different signals in detail, which is not clear to observe.

Several of these techniques are utilized in the traditional machine learning. To develop the hybrid model, the feature variables as supplements should be stable and clear to classify the modulation types.

### *2.3. Constellations of modulation format for Deep Learning and Hybrid Model*

For the Deep Learning model and the first part of the hybrid model, the main principle is image classification. To be applied by the Deep Learning, the data of modulated signals are transferred to the constellation images. The Fig1 shows the constellations as examples of the four modulations at 20dB SNR.

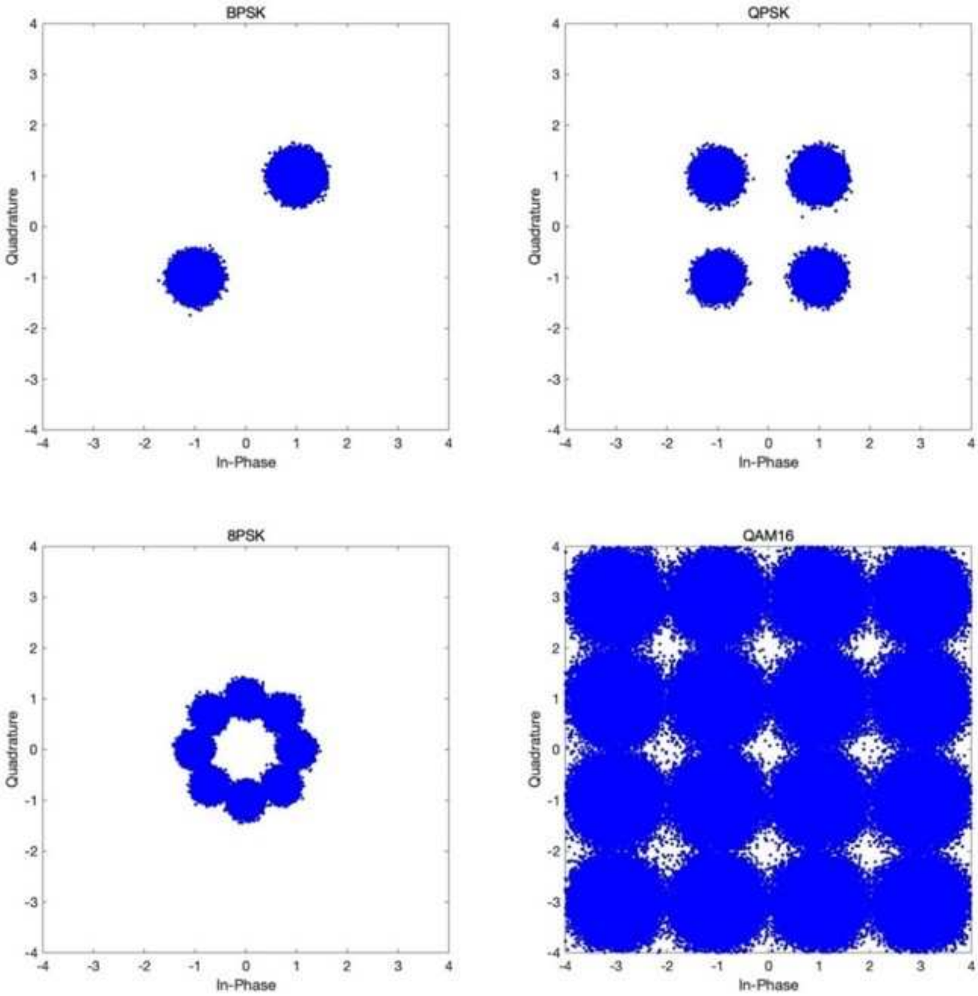



Figure 1 Constellation of Modulated Signals with noise at 20dB SNR

We assume that the systems providing the signals have phase and symbol lock. The SNR effects and constellations are directly observed from Fig 1, which indicates the CNN should give a high accuracy of classification at 20dB SNR. These constellations are captured in JPG format and stored with the labels of the modulation names. To get the efficient constellation images for the Deep Learning, the axes are limited from -4 to 4 of both directions. In this way, the sampled dots can stay in the main part of the images and the constellation images can have the same specification. There are 1200 images captured to feed the Deep Learning models. All the constellation images are used as dataset. 70% of the images are utilized for training, and the rest of images are for the validation process. The models will adjust the parameters to achieve the best performance based on the results of each experiment.

## 3. Classification Method

In this section, the classification methods are introduced. Machine Learning models are based on the statistical features from the modulated signals. The Deep Learning models achieve the classification according to the constellation images. Both of these two methods have limitations of the classification, the hybrid model is developed by combining them both.

### *3.1. Machine Learning*

For the Machine Learning methods, the statistical features needed to be analyzed and selected to work in the model. Based on the formula of features from section2.2, the features from four kinds of modulation types are calculated.

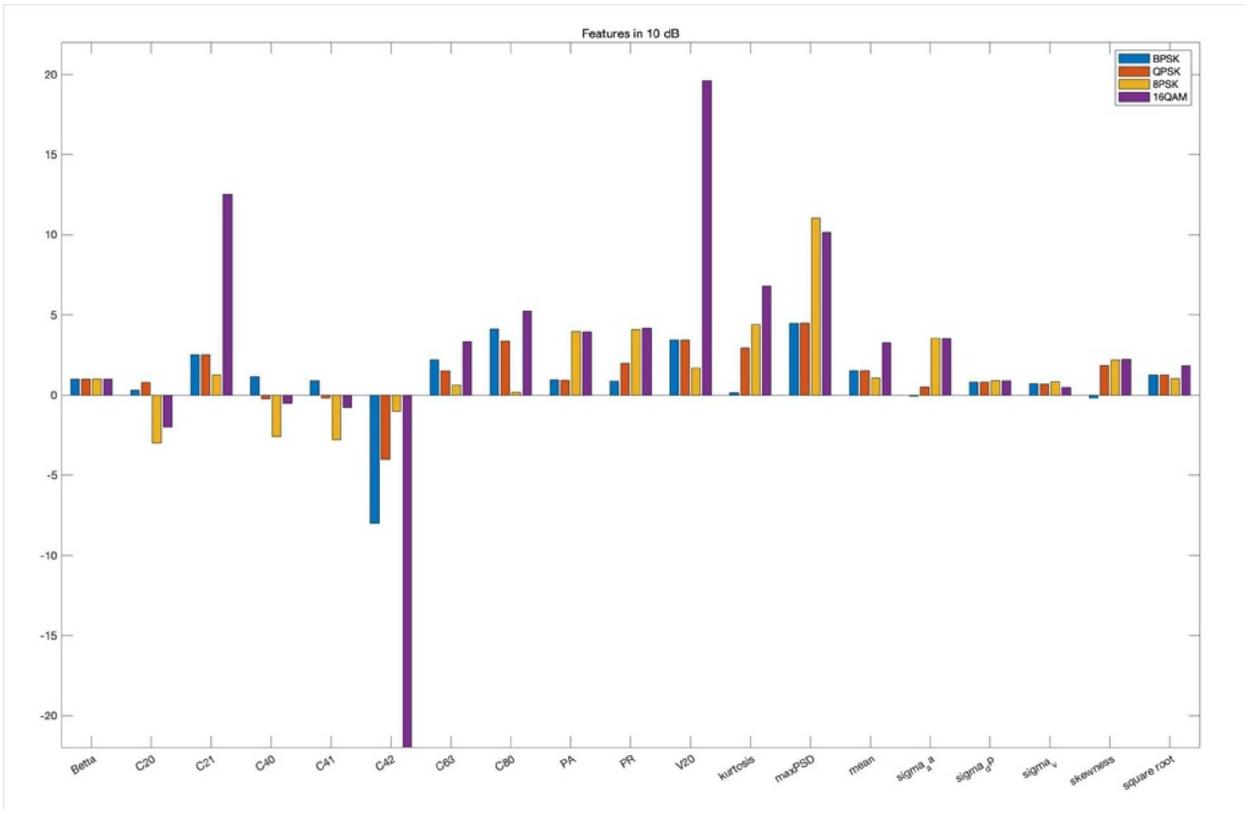



Figure 2 Features of four kinds of modulated signals at 10dB SNR

Fig 2 has shown the statistical features of the four modulated signals at 10dB SNR. To show the characters 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 apply the data with magnitude difference. According to the following graph, 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 the fixed values.

The classification results from SVM, KNN and Decision Tree are compared. SVM is a supervised model that makes the classification between two groups. It can be used to classify the modulation types step by step. KNN is a supervised and non-parametric method, which need to set the value of k at first, the input consists of the k closest training examples in the feature space, the output is the class membership. The decision tree consists of a structure with different nodes which works as a condition and check if the input data are satisfied. All of them are classic statistical methods to make decisions between the four types of modulation using features. But the calculation results from the signals fluctuate when the SNR varies, making decisions difficult.

### *3.2. Deep Learning*

In this session, four CNN models are introduced. A simple CNN developed from the Iris case with several layers, SqueezeNet model, GoogleNet model and Inception-v3 model [12]. The last three models are pretrained networks, which are used for Transfer Learning. As Fig 3 shown, an example of a simple CNN structure can consist of two convolutional layers and two dense fully connected layers. The characters of the signals can be captured by the two convolutional layers and the feature map can be produced by these two layers. After that the features can be learned by the two sequential layers. At last, the results of the classification are outputted.

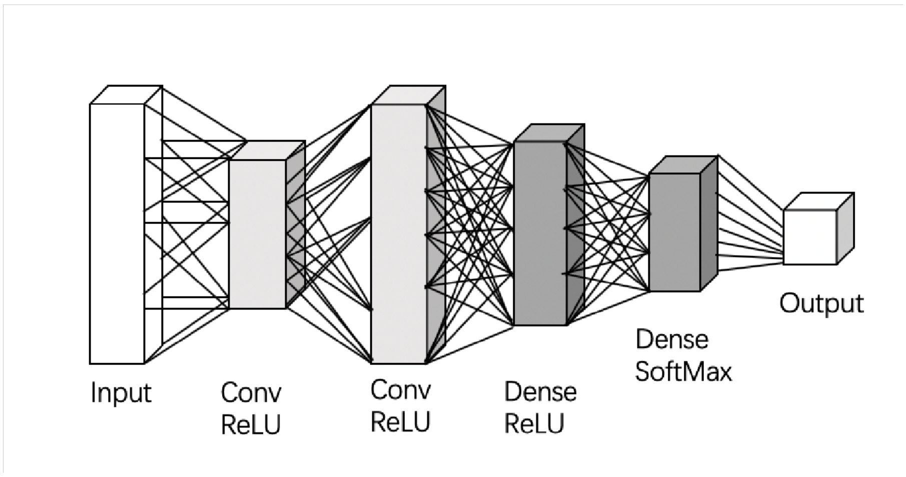



Figure 3 Structure of CNN

*3.2.1. CNN.* This model is created by developing from the Iris recognition case. There are fifteen layers in total. Three convolutional layers are followed by the batch normalization layers and ReLu layers. Two max pooling layers are set between the other three blocks. The convolutional layers can capture the features of the images by the filters. The batch normalization layers can normalize the input channel. The ReLu layers calculate the threshold to the elements.

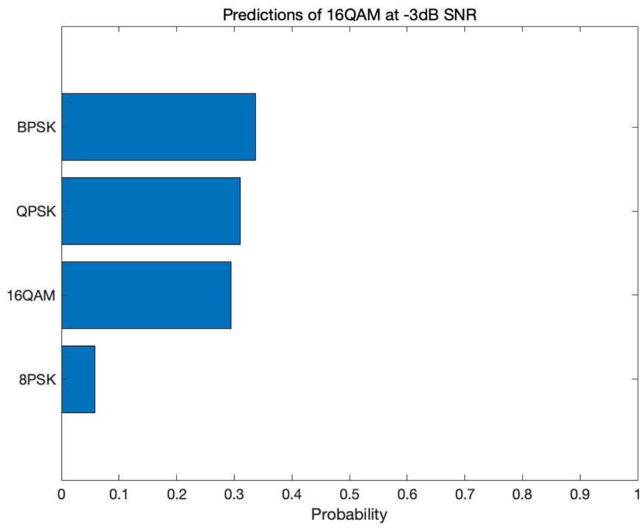
*3.2.2. Transfer Learning Models.* The SqueezeNet has 68 layers and its coefficients of the network were already pretrained to classify 1000 categories. The input image size is 227-by-227. The GoogleNet has 144 layers. It was pretrained to classify the images into 1000 categories and 365 places. The input image size is 224-by-224. The Inception-v3 has 315 layers and it was trained by more than a million images to classify into 1000 categories. The input size is 299-

by-299. Images need to be resized before feeding to models. The last learnable layers are modified, to provide only 4 outputs, corresponding to each of the modulation types.

In these experiments, 70% of images are used as training and the rest are used as validation. The augmented image datastore is applied to rotate the images from -90 degrees to 90 degrees (representing constellation phase errors) and rescale images from 1 to 1.5 randomly, which can help improve the amount of training data and avoid overfitting.

### *3.3. Hybrid Model*

According to previous research, the hybrid method is provided. The deep learning is utilized at first, the statistical features of the signals are applied to the second process for the classification.



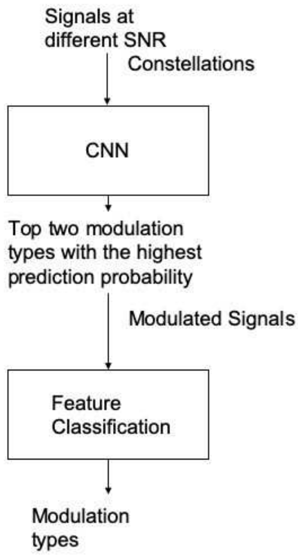




Figure 4 Predictions of 16QAM at -3dB SNR    Figure 5 Structure of Hybrid Model

Figure 4 illustrates the probability of the predictions when the Inception-v3 model detects the signals at -3dB SNR. It is obvious the system cannot classify the QAM16 from the BPSK and QPSK, so the statistical features must be involved to detect the proper detected signals.

The proposed structure of the hybrid model is shown in Figure 5. The structure of deep learning is the model that gets the best classification results from the four CNN. We can see the results at low SNR are not accurate enough. Based on the traditional Machine learning method, the features which can obviously help distinguish the modulation types are chosen to help classify as continue. Feature classification requires the statistical features from the modulated signals.

## 4. Results

This section indicates the results from different classification methods. The Inception-v3 model approves the best results in the deep learning part. It shows the transfer learning can perform better with the pretrained coefficients. To improve the performance of the system, the hybrid model will use the features from the statistical features and be compared to the deep learning model.

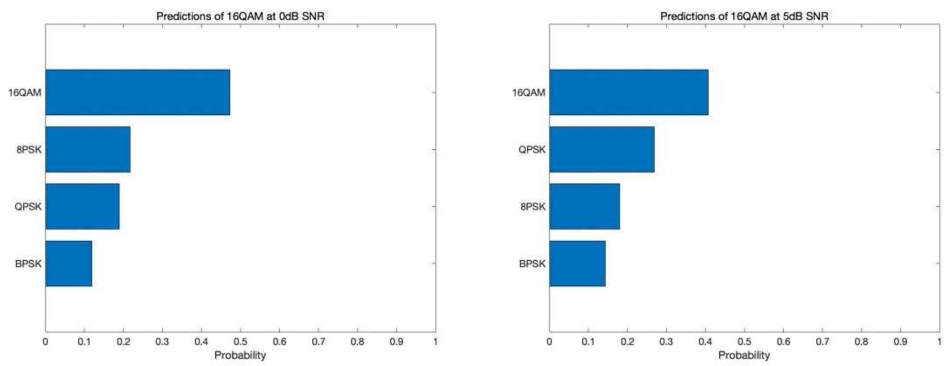
### 4.1. *Deep Learning for images classification*

Form Table 1, the results are shown from the four CNN models. The CNN developed with Iris case performs worse than traditional machine learning methods. It cannot indicate the CNN is unsuitable for detection because the structures and coefficients of the CNN can influence a lot in the classification system. The other three models are CNN as well, but they have more complicated structures and they are pretrained by millions of images. Although the images applied for pretraining are not the constellation of modulated signals, they can also help the models to classify the images by modifying the coefficient again and again.

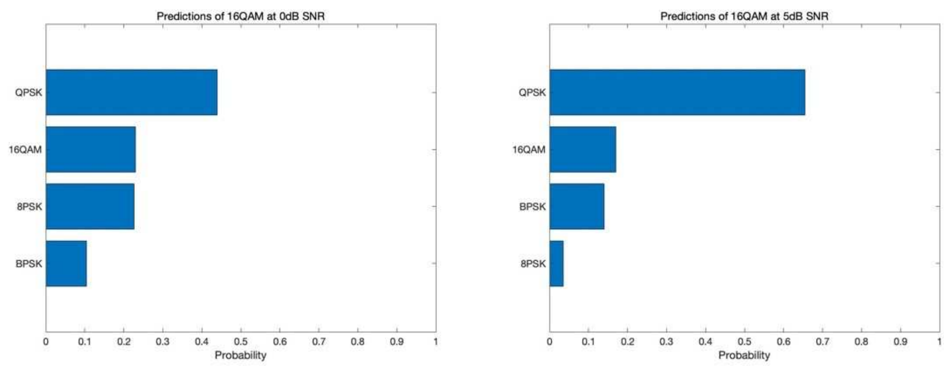
Table 1 Validation Accuracy of Deep Learning Models

| Model Name | Validation Accuracy |
|---|---|
| CNN from Iris case | 36.90% |
| SqueezeNet | 55.83% |
| GoogleNet | 56.11% |
| Inception-v3 | 63.33% |

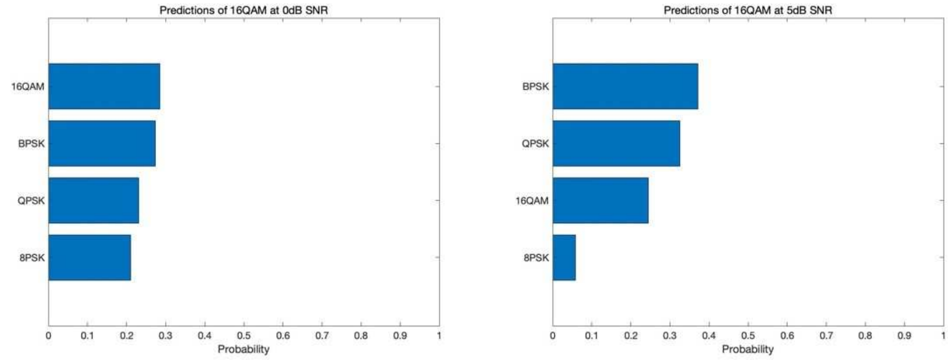
The Inception-v3 compared to the SqueezeNet and GoogleNet shows best accuracy of 63.33%. The training dataset is the images of four modulated signals from -10dB SNR to 20dB SNR. And the Accuracy is calculated from detect the four modulation types from -10dB SNR to 20dB SNR.



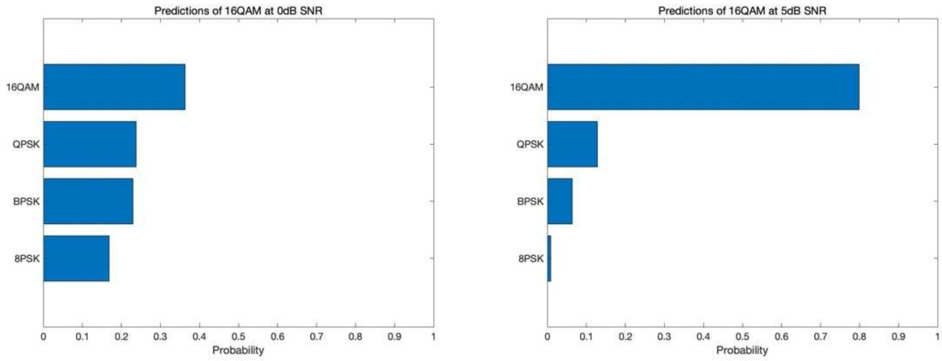

(1) Predictions based on CNN from Iris model




(2) Predictions based on SqueezeNet




(3) Predictions based on GoogleNet




(4) Predictions based on Inception-v3

Figure 6 (1)-(4) Prediction of four deep learning models

Figure 6 is the examples for the QAM16 detection at 0dB SNR and 5dB SNR by the simple CNN model, SqueezeNet, GoogleNet and Inception-v3 model. The validation accuracy of Inception-v3 is only higher than the CNN model based on the Iris case by 10 percent. The CNN model developed from Iris model shows a good classification ability at 0dB SNR, but it cannot promise accuracy at 5dB SNR. For the higher SNR, the Inception-v3 have a robust classification capability. Considering the validation accuracy and stability of the classification at the same time, we decided to choose the Inception-v3 as the most outstanding one from the four models.

When the SNR is more than 10 dB, the predictions of the classification goes nearly 100%. Deep learning has shown its advantages by the results. Based on these, to improve the classification accuracy of low SNR area, we use the feature classification as a supplement.

*4.2. Hybrid model for images classification and feature classification*

According to the section 4.1, we find the Inception-v3 provides the best results from the four kinds of CNN structures over all SNR. From Figure 2 we can find the high order cumulants perform well. The high order cumulants are employed for the hybrid system as the second classification process. The Kurtosis and Skewness can describe the shape of the signals. But as the training data, the signals were varied from -10dB to 20dB SNR, which means the features also changed dramatically. The dynamic changes will increase the complexity of the classification system. So only $C_{20}$, $C_{21}$, $C_{40}$, $C_{41}$, $C_{42}$, $C_{63}$, $C_{80}$ are employed.

Table 2 Compared Accuracy between CNN and Hybrid Model

| SNR Level | Accuracy of Inception-v3 | Accuracy of hybrid model |
|---|---|---|
| -5dB | 52.6% | 75.3% |
| -3 dB | 29.5% | 63.5% |
| 0 dB | 37.5% | 67.7% |
| 5dB | 79.0% | 88.5% |
| 10 dB | 96.7% | 97.3% |
| 20 dB | 99.2% | 99.7% |

Table 2 shows the compared results of Inception-v3 and hybrid model. The accuracy was improved especially at low SNR area. The hybrid model explores signals at the low SNR area and the accuracy is more than 60% by training the whole SNR range data. The accuracy at -5dB SNR is higher than the accuracy at -3dB SNR because we capture the constellation images by the scale of the axis from -4 to 4 and the images are full of the random noise pixels compared to the constellations of other modulation types. At the -3dB SNR, all the four kinds of modulation types have lots of noisy pixels around the focus on the middle part of the constellation, which is more complicated to classify at this period.

## 5. Conclusion

This work compared the different classification methods for AMC, finding the Inception-v3 network provided the best results at SNR levels from -10dB SNR to 20dB SNR. To improve the situation at low SNR, the hybrid model is provided based on the advantages of deep learning and traditional machine learning. The new hybrid model involved the novel image classification to classify the constellation at first and the features classification as a supplement. The accuracy provided by the hybrid model at -3dB SNR increase from 29.5% to 63.5% compared to the Inception-v3. In this way, the hybrid model utilized the constellation to avoid the complex calculation and several features to improve the performance of previous image classification.

The deep learning technology could still be improved further by detecting the images of the Time-Frequency domain at the same time instead of changing the coefficients of the CNN system.

## References

[1] E. Axell, G. Leus, E. G. Larsson, and H. V. Poor, "Spectrum sensing for cognitive radio : State-of-the-art and recent advances," IEEE Signal Process. Mag., vol. 29, no. 3, pp. 101–116, 2012, doi: 10.1109/MSP.2012.2183771.

[2] M. Hamid, S. Ben Slimane, W. Van Moer, and N. Björsell, "Spectrum Sensing Challenges: Blind Sensing and Sensing Optimization," IEEE Instrum. Meas. Mag., no. April, 2016.

[3] J. A. Sills, "Maximum-likelihood modulation classification," pp. 217–220, 1999.

[4] H. Gang, L. Jiandong, and L. Donghua, "Study of modulation recognition based on HOCs and SVM," IEEE Veh. Technol. Conf., vol. 59, no. 2, pp. 898–902, 2004, doi: 10.1109/vetecs.2004.1388960.

[5] A. Abdelmutalab, K. Assaleh, and M. El-Tarhuni, " Automatic modulation classification based on high order cumulants and hierarchical polynomial classifiers," Phys. Commun., vol. 21, pp. 10–18, 2016, doi: 10.1016/j.phycom.2016.08.001.

[6] J. E. Whelchel, D. L. McNeill, R. D. Hughes, and M. M. Loos, "Signal understanding: An artificial intelligence approach to modulation classification," pp. 231–236, 1989, doi: 10.1142/9789814354707_0021.

[7] M. Zhang, M. Diao, and L. Guo, "Convolutional Neural Networks for Automatic Cognitive Radio Waveform Recognition," IEEE Access, vol. 5, pp. 11074–11082, 2017, doi: 10.1109/ACCESS.2017.2716191.

[8] N. An, B. Li, and M. Huang, "Modulation classification of higher order MQAM signals using mixed-order moments and fisher criterion," 2010 2nd Int. Conf. Comput. Autom. Eng. ICCAE 2010, vol. 3, pp. 150–153, 2010, doi: 10.1109/ICCAE.2010.5451214.

[9] C. M. Spooner, "On the utility of sixth-order cyclic cumulants for RF signal classification," Conf. Rec. Asilomar Conf. Signals, Syst. Comput., vol. 1, pp. 890– 897, 2001, doi: 10.1109/ACSSC.2001.987051.

[10] O. A. Dobre, Y. Bar-Ness, and W. Su, "Higher-order cyclic cumulants for high order modulation classification," Proc. - IEEE Mil. Commun. Conf. MILCOM, vol. 1, no. C, pp. 112–117, 2003, doi: 10.1109/milcom.2003.1290087.

[11] J. Lee, B. Kim, J. Kim, D. Yoon, and J. W. Choi, "Deep neural network-based blind modulation classification for fading channels," Int. Conf. Inf. Commun. Technol. Converg. ICT Converg. Technol. Lead. Fourth Ind. Revolution, ICTC 2017, vol. 2017-December, pp. 551–554, 2017, doi: 10.1109/ICTC.2017.8191038.

[12] K. Grm, V. Štruc, A. Artiges, M. Caron and H. K. Ekenel, "Strengths and weaknesses of deep learning models for face recognition against image degradations," in IET Biometrics, vol. 7, no. 1, pp. 81-89, 1 2018, doi: 10.1049/iet-bmt.2017.0083.