Efficient Deep Learning Model for Data-Limited Modulation Recognition

Judith Nkechinyere Njoku, Angela Caliwag, Wansu Lim

Department of Aeronautics, Mechanical and Electronic Convergence Engineering
Kumoh National Institute of Technology, Gumi, South Korea
email: wansu.lim@kumoh.ac.kr

Abstract

Deep learning (DL) has been successfully applied for modulation recognition tasks. Most of the existing applications pay little attention to the volume of data used. In reality, real historic data for modulation recognition could be limited. Thus it is better to train recognition models on small amounts of data. This is however a huge challenge, since the performance of DL models depend on sufficient data. In this paper, we introduce an efficient system model based on CNN and different data augmentation methods, for the purpose of modulation recognition. Our system employs random rotation, flip, zoom, random shift and resize methods for data augmentation. The CNN model employs small filter sizes, pooling layers, and dropout layers to improve the network. We apply a small dataset consisting of 40 constellation images per modulation type, to the system. We further analyze the performance based on three data augmentation intervals. From our experiments, the model achieved an accuracy of 32% without data augmentation and 76.07%, 35.71% and 96.42% on the three data-augmentation intervals.

Ⅰ. INTRODUCTION

Automatic modulation recognition (AMR) is a fundamental step in many wireless communications applications such as unauthorized signal detection and spectrum management. It can be described as the process of classifying and recognizing the types of modulation techniques from unknown received signals [1]. The accuracy of AMR models has been a challenge, due to the numerous types of modulation types and the adverse effect of noise and multipath fading conditions. One viable solution is to employ machine learning (ML) solutions.

Deep learning (DL) is an arm of ML, which has gained immense popularity, due to its ability to learn complex identification/classification tasks. As a result, it has been applied in various application fields including computer vision, economics, and natural language processing, just to mention a few. DL has recently been explored for its ability to solve problems in wireless communication applications such as AMR [1-3]. Different variants of DL have thus been applied and analyzed for the purpose of accurate AMR. For instance, a model based on convolutional neural networks (CNN) was proposed in [4] for the purpose of achieving high accuracy and short computing time. In [5], a long short term memory (LSTM) model based on attention mechanism was introduced and shown to achieve high convergence and accuracy.

While most of these works have been efficient in developing highly latent and accurate models, there is little attention on the volume of data used. The dataset which is usually used to train these models is generated by dedicated software or analytical models. This allows for the generation of sufficient training data. However, in real scenarios, historical data could be limited. Thus it is better to train the model on small amounts of data. This is however a huge challenge because the performance of DL based models depend on sufficient data.

In this paper, our aim is to improve the recognition of modulation signals on small datasets, by using data augmentation techniques. We convert I/Q modulation signals to constellation images and apply data augmentation. Specifically we introduce an efficient DL model based on several data augmentation types and compare its performance for three data augmentation intervals: (i) Train-time augmentation, (ii) Test-time augmentation, and (iii) Train-test-time augmentation.

Ⅱ. PROPOSED SYSTEM MODEL

In this section, the modulation recognition problem is introduced. The proposed data-limited modulation recognition model is also summarized and represented in Fig 1.

A. Problem Formulation

Consider a communication system consisting of a transmitter, receiver and channel. Given that a transmitted signal is affected by additive white Gaussian noise (AWGN), the received signal can be represented as:

$$y(n) = x(n) * h(n) + g(n), \qquad (1)$$

where, $y(n) = [y_1(n), \dots, y_i(n)]$ represents the received signal that corresponds to the input signal $x(n) = [x_1(n), \dots, x_i(n)]$, $h(n)$ is the channel impulse response and $g(n)$ is the AWGN. Modulation recognition can be seen as an i-class classification problem, where the information about the received signal $y(n)$ helps in the recognition of i modulation types.

Several papers have proposed to transform raw I/Q signal data into constellation images before feeding to the CNN image classifier. As represented in Fig 2, constellation diagrams are a 2-D representation of the modulated signal. It is created by mapping signal samples into scatter points, located on the complex plane. In [6], it was established that training CNN models on RGB constellation resulted in a higher accuracy than grayscale images.

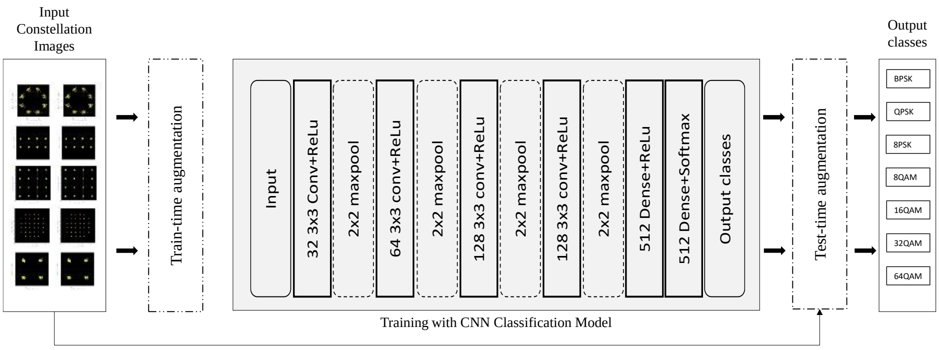


Figure 1: Proposed Modulation Recognition Model

B. Data Augmentation Methods

Data augmentation is the process of synthesizing new data from the available data. It has been extensively used to solve the problem of limited data for training DL models in many applications including image recognition, natural language processing, just to mention a few. In essence data augmentation helps to improve the generalization ability of the model and resolve overfitting issues. In this paper, we adopted five data augmentation methods including:

(a) Random rotation: The data sample is rotated by an angle (e.g. 40 degrees). Each rotated sample is a unique one to the model.
(b) Flip: The data sample can be flipped vertically generating upside down views or horizontally, generating left and right side views.
(c) Zoom: The data sample can be zoomed in or out within a specified range.
(d) Random shift: The pixels of the sample can be shifted horizontally to the left or right randomly.
(e) Resize: The sample can be scaled. Either the height and width can be scaled, or only the shortest or longest end is scaled.

Consequently, the entire dataset is expanded by a scale factor of N.

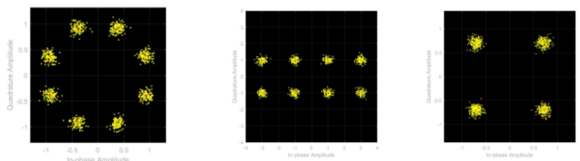


Figure 2: RGB Constellation images of 8PSK, 8QAM and QPSK modulation signals

C. CNN Architecture

We propose to use an efficient CNN model, which is made up of four 2D convolutional layers, followed by dropout and max pooling layers. The input to the model is designed to receive and process modulation constellation images. As shown in Fig. 1, the architecture employs 3×3 filters in its convolutional layers to precisely learn generic features from the images. It also uses 2×2 filters in its max pooling layers to reduce the spatial dimension of the input image. Furthermore, it uses dropout layers with a drop rate of 0.1, to regularize the network. The convolutional layers are activated with the ReLu activation function, while the final dense layer is activated using the Softmax activation function.

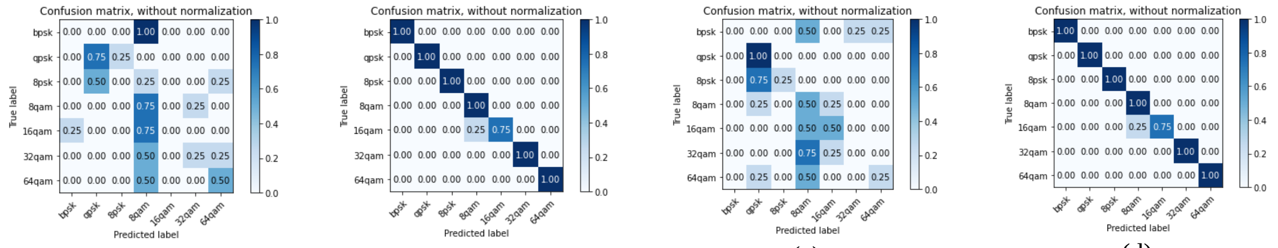
Ⅲ. Experimental Results and Discussion

A. Implementation Details

In this section, the MATLAB software was used to generate the constellation images for the following modulation signals: binary phase shift keying (BPSK), quadrature-PSK (QPSK), 8PSK, 8-qadrature amplitude modulation (8QAM), 16QAM, 32QAM and 64QAM signals. All signals were generated with both Gray and Binary symbol mapping with an output symbol rate of 1000 sym/s. The normal raised cosine and root raised cosine filtering were applied to randomly selection signals. Several channel impairments were also applied to the data generated. They include: AWGN, phase offset, frequency offset, phase noise, DC offset, and IQ imbalance. 40 constellation images were generated for each signal type. The data set was split into a train, test and validation set of ratio 70:20:10.

The CNN model was trained in Keras with GPU 1xTelsa K80, having 2496 CUDA cores and 12 GB GDDR5 VRAM in Google Colaboratory. The model was trained for 75 epochs.

Since deep learning models include the training and inference stages, data augmentation can be applied in both stages, leading to three possible data augmentation intervals: train-time augmentation, test-time augmentation and train-test-time augmentation. We have thus applied and analyzed the performance of data augmentation in these three intervals. In an end-to-end process, the proposed system receives the input constellation image, performs data augmentation and forwards the augmented samples to the CNN model for training and testing. During the testing stage, data augmentation is also applied to the test set, and the model is evaluated on the accuracy of modulation recognition.



(a) (b) (c) (d)

Figure 3: Confusion matrices for modulation recognition task with a) no augmentation b) train-time augmentation c) test-time augmentation and d) train-test-time augmentation

B. Results and Discussion

In Fig 3, we numerically study the performance of the data augmentation at different intervals. Where scale factor N =10. We compare with the baseline without augmentation. The model achieves a recognition accuracy of 32%, when there is no data augmentation applied. The train-test time augmentation achieved the best performance with an accuracy of 96.42%. While the train-time augmentation performs better than the test-time augmentation with an accuracy of 76.07% and 35.71% respectively.

Ⅲ. CONCLUSIONS

In this paper, we proposed a system for the efficient recognition of modulation types, based on limited data, consisting of modulation constellation images. The proposed system employs different data augmentation methods to enlarge the dataset and applies this dataset for the training of a CNN model. We further analyze the performance based on three different data augmentation intervals. Results show that the train-time-test data augmentation system is efficient in recognizing modulation types.

For future works, we will study the effect of each data augmentation type on the classification accuracy. We will also explore DL based methods of data augmentation, as they have been successfully applied in other domains.

ACKNOWLEDGMENT

This work was supported by the Technology development Program (S2829065, S3010704) funded by the Ministry of SMEs and Start-ups (MSS, Korea), and by the Basic Research Program through the National Research Foundation of Korea (NRF) funded by the MSIT (2020R1A4A101777511).

REFERENCES

[1] G. B. Tunze, T. Huynh-The, J. -M. Lee and D. -S. Kim, "Sparsely Connected CNN for Efficient Automatic Modulation Recognition," in *IEEE Transactions on Vehicular Technology*, Dec 2020, vol. 69, no. 12, pp. 15557-15568.

[2] N. J. Nkechinyere, M. E. Morocho-Cayamcela, and W. Lim, "Automatic Radar Waveform Recognition using the Wigner-Ville distribution and AlexNet-SVM," Proceedings of the KICS Summer Conference, Pyeongchang, South Korea, 2020.

[3] J. N. Njoku, M. E. Morocho-Cayamcela and W. Lim, "CGDNet: Efficient Hybrid Deep Learning Model for Robust Automatic Modulation Recognition," in *IEEE Networking Letters*.

[4] A. P. Hermawan, R. R. Ginanjar, D. Kim, and J. Lee, "Cnn-based automatic modulation classification for beyond 5g communications," IEEE Communications Letters, vol. 24, no. 5, pp. 1038-1041, 2020.

[5] Y. Chen, W. Shao, J. Liu, L. Yu and Z. Qian, "Automatic Modulation Classification Scheme Based on LSTM With Random Erasing and Attention Mechanism," in *IEEE Access*, vol. 8, pp. 154290-154300, 2020.

[6] B. Tang, Y. Tu, Z. Zhang, and Y. Lin, "Digital signal modulation classification with data augmentation using generative adversarial nets in cognitive radio networks," IEEE Access, 2018, vol. 6, pp. 15713-15722.