https://doi.org/10.2352/EI.2023.35.3.MOBMU-366

An RF Modulation Recognition Method using Machine Learning

Rahul Mundlamuri[1], Devasena Inupakutika[1], Ganesh Reddy Gunnam, Thinh Ngo, David Akopian; The University of Texas at San Antonio; San Antonio, Texas

Abstract

Deep learning has been successfully applied in many fields to optimize decision making including self-driving cars, healthcare, machine translation, image recognition, among others. In wireless communication, deep learning has been used in channel estimation, signal classification, massive MIMOs, heterogeneous networks, energy harvesting, device-to-device (D2D) communications, and so on. Machine learning (ML) and deep learning (DL) neural networks to RF signal recognition are applied. Specifically, we built, trained, and tested two ML models, SVM and XGBoost, and two DL models, Convolutional Network (ConvNet), and Residual Network (ResNet). We leverage the use of ConvNets and ResNets to the complex-valued temporal radio signal domain. We utilized the online dataset 2018.01A from DeepSig. This work explores the scientific application of ML/DL in terms of dataset processing, deep neural network construction, training, testing, fine-tuning, and reporting the analyses and results. An in-depth analysis of the mentioned models was performed for RF communication signals classification.

Introduction and Background

Deep learning has proved to surpass human capacity in some areas such as gaming, image recognition, natural language processing. Next-generation wireless networks have already embraced deep learning paradigms. Data-driven decision making, instead of knowledge-driven or deterministic modeling, is naturally more appropriate for wireless communication systems as propagation medium or wireless channels are constantly changing due to mobility, interference and noises. Additionally, the humongous amount of generated data (e.g. big data) should be dynamically taken advantage of. There are many on-going researches on deep learning for next-generation wireless networks. The task of RF signal classification or symbol detection is a prime candidate for deep learning as shown by [2]-[6]. In this project, we adapted ML and DL to RF signal classification based on *Over-the-air Deep Learning Based Radio Signal Classification*[2]. In addition, we tried various neural network architectures (e.g. types, number of layers, number of nodes per layer, activation functions) and study their performances.

Traditional knowledge-driven RF signal classification

Traditional methods of RF signal classification (e.g. 4G LTE/LTE Advanced) are knowledge-driven using maximum likelihood and Bayesian estimation including Least Square (LS), Minimum Mean Square Error (MMSE), Parametric Model (PM) and Iterative Channel Estimation (ICE). For 5G and future generations, more robust, dynamic and higher capacity RF signal modulations and classifications are demanded. Adapting and incorporating data-driven approaches using ML and DL neural networks are currently in active pursuit.

Data-driven RF signal classification

Wireless propagation is complex and dynamic with various channel impairments such as mobility, noises, interference, shadowing, multipath and under the effects of carrier frequency offset and symbol rate. Consequently, accuracy of knowledge-driven modeling of RF signal classification is limited. Alternatively, data-driven modeling (e.g. ML and DL) would be a proper candidate. The prime difference between ML and DL is that ML uses processed data (e.g. feature extractions) and DL uses more accurate higher degrees of freedom models from raw data using end-to-end feature learning. DL modeling has shown to outperform ML modeling in many tasks as it can extract better features from raw data than ML does with high-order statistical parameters.

In the current work, we used the real time-series radio data captured over the air with the realistic simulation of the wireless propogation environment, real propagation effects and new methods for signal classification. We also adapted ConvNet models for RF based device fingerprinting in cognitive communication networks [8] for the RF signal classification. We investigated the system parameters as well as the effect of training parameters and the Signal to Noise Ratio (SNR) on the performance and the accuracy of the RF signal classifiers.

The main contributions of this work are the following:

1. Understanding the time-series radio signal data, checking the modulation labels, associated Signal to Noise Ratios (SNRs) and the primary impairments present in the wireless channel. Primary impairments present in any wireless channel consist of Carrier frequency offset i.e. due to the disparate local oscillators (LOs) and motion, Symbol rate offset i.e. symbol clock offset and time dilation due to disparate clock sources and motion, Delay Spread i.e. non impulsive delay spread due to delayed reflection, diffraction and diffusion of emissions on multiple paths and Thermal noise which is the additive white Gaussian Noise present due to receiver device sensitivity.
2. Implementing the feature extraction methods for RF signal classification based on the latest over the air capture data. Initially through expert feature extraction and then feeding to ML algorithms mentioned in the above sections. This is called Baseline Classification Approach. Secondly, providing with a windowed input of the raw radio time series r(t)

[1]These authors contributed equally.

to the mentioned DL models.

3. Adapting the general ConvNet models and the ones for RF fingerprinting device classification to the RF signals classification and checking the performance and accuracy by varying the configurations (varying SNRs and training dataset sizes etc.).
4. Further evaluating the performance of ResNet in identifying the signals based on their modulation types and varying SNRs.

Preliminaries

Dataset details and pre-processing

Dataset that we used for the current work includes both synthetic simulated channel effects and over-the-air recordings of 24 digital and analog modulation types. Data are stored in Hierarchical Dsta Format (hdf5) as complex floating point values [1], with 2 million examples, each 1024 samples long. Further details about the data that are extracted after processing further are as below:

1. X dataset consists of the Complex floating point values, with 2 million examples, each 1024 samples long with each sample having in-phase (I) and Q-phase (Q) components.
2. Y dataset consists of the same 2 million signal examples with 24 length and indicated by '1' at the index (order specified by file classes.txt) corresponding to the modulation type.
3. Z dataset consists of the Signal to Noise Ratio (SNR) in dB for the same 2 million signal examples. This column usually will have low SNRs ranging from -20 dB to +30 dB Es/N0.

Hence, the I/ Q time domain plot of the signals with a chunk of 256 out of 1024 complex-valued samples are as shown in Figure 1.

Expert feature extraction: Baseline classification approach

Before feeding the data into the ML models i.e. SVM and XGBoost for this work, we extracted statistical modulation features. These include higher order statistics (HoSs) and cyclostationary moments (HoMs) [9] for the digital modulation techniques. These are the standard and widely used features to compactly sense and detect the signals with dense periodic components that are created by the structure of the carrier, symbol timing, and symbol structure for certain modulations. Expected values of peaks from the auto-correlation and spectral correlation functions usually successfully aid in robust signal classification even with completely unknown and random data. We further calculated mean and standard deviation of the signal segments. These are calculated even for analog modulations where symbol timings doesn't produce the compact HoSs and HoMs. For ease of data handling due to such a huge dataset, we store and manage them as numpy array files called *.npy*. Cumulants are calculated from the moments. We hence, could extract 27 of such features that most closely distinguish signals to be passed as input to SVM and ensembling method of gradient boosting called XGBoost when mapping our features to a class label for the decision criterion. XGBoost clearly outperformed SVM as it combines the collection of classifiers for performance improvement. These models have been directly trained on the OTA dataset. In the section, therefore, we chose to display only XGBoost while comparing the correct classification probablility or rates (plotted usually vs SNR (dB)) for signal classification problems.

We also proceeded, initially with dataset which consists of 11 classes that are all relatively low information density and are commonly seen in impaired environments. These 11 signals represent a relatively simple classification task at high SNR in most cases. Secondly, we performed analyses with the dataset that contains all the 24 modulations and are difficult for the classification task. These include a higher order modulations (QAM256 and APSK256 etc.) that are used in the real world in very high-SNR, low-fading channel environments.

Decision criterion for the models in the current work

Short-time classification is challenging but is unavoidable when decision processes cannot wait to acquire more data to increase certainty. This is quite usual in many real world systems when dealing with short observations (such as when rapidly scanning a receiver) or short signal bursts in the environment. Under these effects, with low SNR examples (from 20 dB to +30 dB Es/N0), achieving near 100% classification rates on the full dataset is not usual. Hence, it can be seen as a good benchmark for comparisons.

DL classification approach

For DL based approaches, ConvNet and ResNet, we considered the complex valued input (actual raw data) as an input dimension of 2 real valued inputs and use r(t) as a set of 2xN vectors into a narrow 2D Convolutional Network where the orthogonal synchronously sampled In-Phase and Quadrature (I & Q) samples make up this 2-wide dimension.

ConvNet Model:

For the ConvNet, we adapted VCGNet architecture to 1D CNN and improved upon the network for the 11 modulations initially. The features into this CNN are the raw I/ Q samples of each RF signal example and they are normalized to unit variance. As discussed in earlier sections, we don't perform any expert feature extraction here and incorporate end-to-end learning. The structure of the model evaluated on the simple set of modulation classes details are as below:

```
_________________________________________________
Layer (type)          Output Shape
     Param #
=================================================
reshape_1 (Reshape)   (None, 2, 128, 1)
     0
_________________________________________________
zero_padding2d_1      (None, 2, 132, 1)
     0
_________________________________________________
conv1 (Conv2D)        (None, 2, 130, 256)
     1024
_________________________________________________
dropout_1 (Dropout)   (None, 2, 130, 256)
     0
```

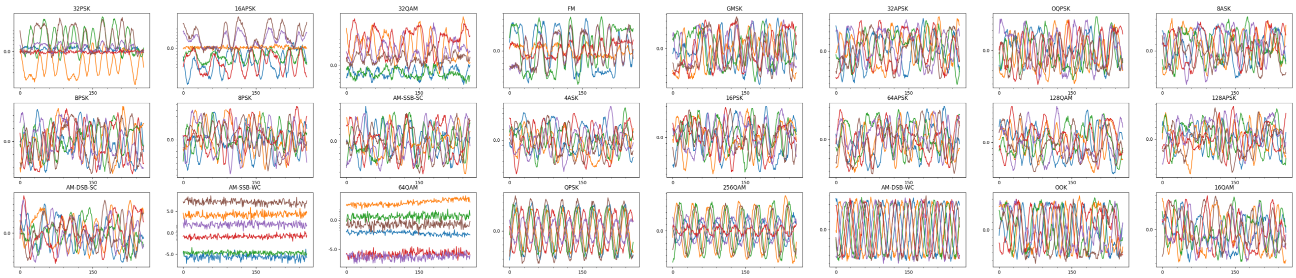

Figure 1. *I/ Q time domain examples of 24 modulations for the over the air capture at SNR=10dB with number of samples=256*

```
_________________________________________________________________
zero_padding2d_2     (None, 2, 134, 256)
       0
_________________________________________________________________
conv2 (Conv2D)       (None, 1, 132, 80)
       122960
_________________________________________________________________
dropout_2 (Dropout)  (None, 1, 132, 80)
       0
_________________________________________________________________
flatten_1 (Flatten)  (None, 10560)
       0
_________________________________________________________________
dense1 (Dense)       (None, 256)
       2703616
_________________________________________________________________
dropout_3 (Dropout)  (None, 256)
       0
_________________________________________________________________
dense2 (Dense)       (None, 11)
       2827
_________________________________________________________________
activation_1 (Activation) (None, 11)
              0
_________________________________________________________________
reshape_2 (Reshape)       (None, 11)
              0
=================================================================
Total params: 2,830,427
Trainable params: 2,830,427
Non-trainable params: 0
_________________________________________________________________
```

ResNet Model:

With the improvement in network algorithms and architectures since AlexNet, the effective training of the deeper networks using more and wider layers has been made possible. In the ConvNet, a smaller ConvNet with several layers was improved over the state-of-art CovNets as mentioned in previous subsection. The skip or bypass connections are common in ResNets. For current work, we tested starting from 2 to 6 residual stacks and it shows improvement in classification accuracy over ConvNets. The longer structure of ResNet is omitted. It also trains in lesser number of epochs as compared to ConvNets. However, both the models are trained over 100 epochs, but we used Keras model checkpoints [10] to observe the trend in accuracy or its saturation point. We also observed that for 1024 samples on the 2 million dataset, Self normalizing Neural Networks [11] in the fully connected layers, and using Scaled Exponential Linear Unit (SELU) activations for the last 2 FCs with Alpha dropout provide improvement over conventional ReLU with dropout configurations. The model evaluated on the complete set of modulation classes 24 for L=4 residual stacks details are as below (This is when the network was trained on subset of data i.e. 75000 signal examples. It improved certainly when increased the training set size from 60000 examples to 75000). We also faced issue when training with 2 million examples pertaining to the storage issue and not enough memory to run on SHAMU GPU. Hence, we limited to 75000 examples set. Table 1 lists the ConvNet and ResNet specifications for training.The results for the discussed approached and models are discussed in the results section.

Table1: Models Configuration

Models	Layers/ Stacks	Trainable Parameters
ResNet	5 Stacks, 2 FC layers, Output Softmax	4,344,984
ConvNet	7Conv+MaxPool Layers, 2 FC layers, Output Softmax	2,830,427

Experimental Results

We conducted sensing performance analysis based on the models discussed and the details of the computing resources are as below:

1. We leveraged UTSA SHAMU cluster [12]: compute and GPU nodes for all our training and analysis.
2. ConvNet and ResNet on GPU runs in minutes to an hour as compared to several hours on CPU (varies with training set size from actual dataset).

We trained on synthetic dataset of 1 million dataset initially and then evaluated on OTA dataset of 2 million examples, each 1024 samples long. We can observe a decent performance after fine tuning the models using transfer learning and observed that accuracy for ResNet is higher than the ConvNet and Baseline methods. It falls in the range of what authors claim in the paper that it is between 64% and 80% (shown in Figure 2). The one reason we see it at the lower end of the range was because there were some modulations (out of 24 classes) missing from the dataset when we started using them for analyses from their website.

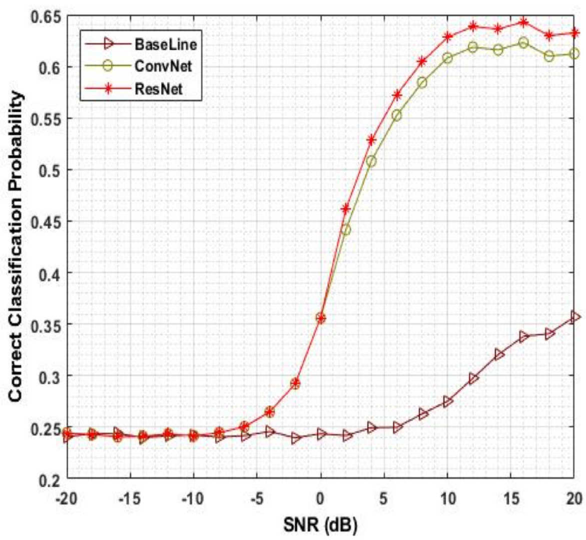

Figure 2. Performance comparison ML (XGBoost) Vs DL

For ConvNet, we trained initially on 11 modulation classes data and at SNR of 10dB, it performs pretty decently in recognizing the signals but along with some multiple classifications as in Figure 3 The ConvNet model after improving upon this was

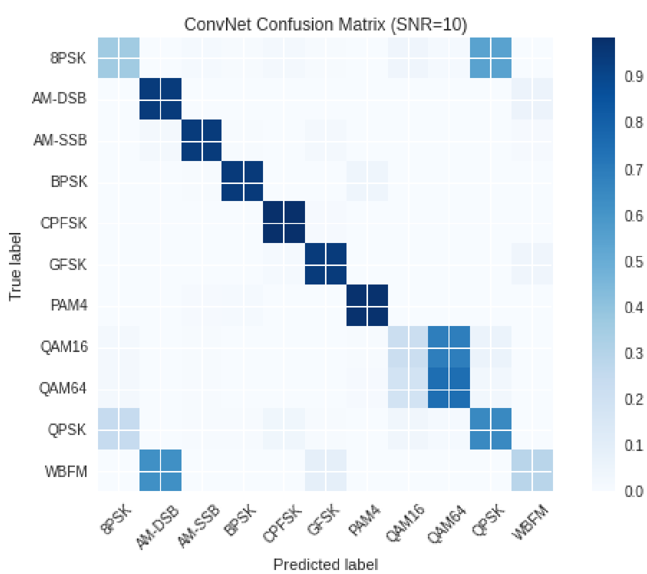

Figure 3. 11 modulation AWGN dataset ConvNet at SNR = 10dB

trained on synthetic 1 million example data and the result of classification evaluation shows that it is only performing better at higher SNRs and poorly at lower SNRs. 4 and 5.

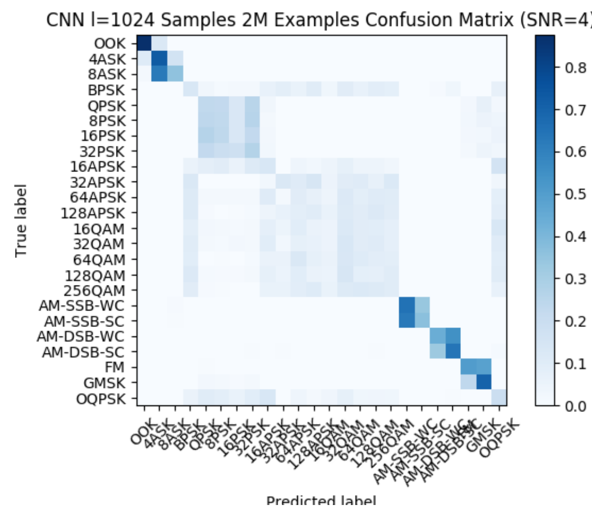

Figure 4. 24 modulation confusion matrix for ConvNet trained and tested on real time capture at SNR = 10dB

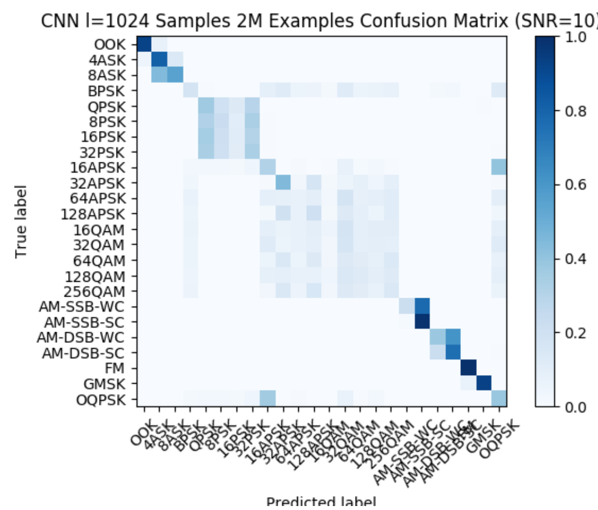

Figure 5. 24 modulation confusion matrix for ConvNet trained and tested on real time capture at SNR = 10dB

Further fine tuning of the ResNet model (with L=6 residual stacks) and increasing the training dataset to complete OTA data and evaluating on the fresh synthesis of the radioML data shows an improvement in the classification probabilities. These are shown based on the SNRs in the form of confusion matrices as in Figure 6. We observed around 88% accuracy in this scenario.

Conclusion and Future work

The ResNet approach achieves the state-of-art classification performance on the complex 24 modulation dataset thus making use of the network depth effectively. When comparing the performance of machine learning and deep learning models, the order is: *ResNet* > *ConvNet* > *XGBoost* > *SVM*. We observed that the performance depends on various parameters for signals such as impairments, training dataset size, propagation effects, observation window and modulation types. Hence, these need to be con-

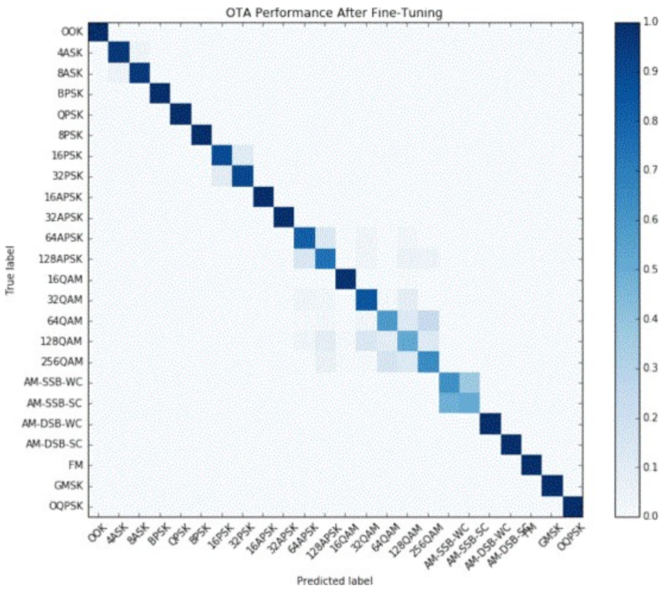

Figure 6. *24 modulation confusion matrix for ResNet trained and tested on real time capture at SNR = 10dB*

sidered while designing DL based model solution for RF signals. Due to the expanding presence of wireless technology and the continuously increasing spectrum, deep learning being applied to RF signals is becoming more useful in Cognitive Radio Networks (CRNs). In future, we intend to investigate the potential of Recurrent Neural Networks (RNN) for these kind of applications.

References

[1] DeepSig OTA Dataset, [Online]: `https://www.deepsig.io/datasets`, Accessed on: January 2019.

[2] T. J. O'Shea, T. Roy, T. C. Clancy. *Over-the-air Deep Learning Based Radio Signal Classification*. IEEE Journal of selected topics in signal processing, Vol. 12, No. 1, February 2018.

[3] T. J. O'Shea, J. Corgan, T. C. Clancy. *Convolutional radio modulation recognition networks* Proc. Int. Conf. Eng. Appl. Neural Netw., pp. 213-226, 2016.

[4] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou. *Modulation Classification Based on Signal Constellation Diagrams and Deep Learning*. IEEE Transaction on Neural Networks and Learning Systems, Jul. 2018.

[5] N. E. West, T. O'Shea. *Deep architectures for modulation recognition* Proc. IEEE Int. Symp. Dyn. Spectr. Access Netw. (DySPAN), pp. 1-6, Mar. 2017.

[6] T. Yucek, H. Arslan. *OFDM signal identification and transmission parameter estimation for cognitive radio applications* Proc. IEEE Glob. Telecommun. Conf. (IEEE GLOBECOM), pp. 4056-4060, Nov. 2007.

[7] Y. LeCun, Y. Bengio, G. Hinton. *Deep learning* Nature, vol. 521, no. 7553, pp. 436-444, May 2015.

[8] K. Merchant, S. Revay, G. Stantchev, B. Nousain. *Deep Learning for RF Device Fingerprinting in Cognitive Communication Networks* IEEE Journal of Selected Topics in Signal Processing, Vol. 12, No. 1., pp. 160-167, February 2018.

[9] Robert M. Gray, Lee D. Davisson. *An Introduction to Statistical Signal Processing* January 2011.

[10] Keras Models, [Online]: `https://keras.io/models/about-keras-models/`, Accessed on: January 2019.

[11] Klambauer, G., Unterthiner, T., Mayr, A., & Hochreiter, S. *Self-Normalizing Neural Networks* In Advances in Neural Information Processing Systems (NIPS), September 2017.

[12] UTSA SHAMU HPC Cluster. [Online]: `https://hpcsupport.utsa.edu/foswiki/bin/view/Main/QuickStartGuide`, Accessed on: January 2019.