Modulation Classification using Convolutional Neural Networks and Spatial Transformer Networks

(Invited Paper)

Moein Mirmohammadsadeghi, Samer S. Hanna, and Danijela Cabric
Electrical and Computer Engineering Department, University of California, Los Angeles
Email: moeinmir@ucla.edu, samerhanna@ucla.edu, danijela@ee.ucla.edu

***Abstract*—In this paper, we propose a method for modulation classification based on Spatial Transformer Networks and Convolutional Neural Networks. Spatial Transformer Networks were originally proposed by the computer vision community to make images invariant to spatial transformations. We adopt this model in the modulation classification problem by making the raw baseband I/Q samples invariant for some channel effects. We study the impact of adding STN to CNN classifier in terms of classification accuracy for different oversampling ratios. We also compare the accuracy of proposed classifier with a conventional statistical method based on forth order cumulants. Results show that proposed classifier has better classification accuracy than the one based on cumulants, and its performance is improved for low oversampling ratios.**

***Index Terms*—Modulation Classification; Spatial Transformer Networks; Convolutional Neural Networks;**

I. Introduction

Modulation classification was first invented and used in military application where there was a need to identify adversary transmitted modulation schemes for jamming purposes [1]. It has recently found application in recent advanced wireless communication systems such as cognitive radio and techniques like link adaptation (adaptive modulation) and interference cancelation [1].

In general, there are two categories of existing classification methods: feature-based classifiers and likelihood-based classifiers. In terms of classification error, likelihood-based classifiers are optimal. However, implementation-wise, they requires very large look-up-tables (LUTs). In addition, errors due to phase offsets and low signal-to-noise ratio (SNR) can reduce accuracy of classification significantly. To address these issues, different feature-based approaches such as cumulant-based classifiers [2], cyclostationary-based [3] and Kolmogorov-Simirnov classifier [4] were developed. These features can also be used as an input to a supervised machine learning technique. For example, in [5], Support Vector Machine (SVM) is used to classify various modulation types using cumulants as features. The performance of these techniques relies on the set of features selected. More recently, Convolutional Neural Networks (CNN) were proposed to classify modulation using the raw I/Q samples directly [6]. CNNs have been shown to give a better performance without having to manually determine any features. In [7], an additional layer was introduced prior to the convolutional neural network to normalize the dataset.

In this paper, we propose a modulation classification technique based on supervised machine learning using CNNs and Spatial Transformer Networks (STNs). In this method, we used I/Q samples of received signal as input without any specific feature extraction. The STN helps normalize the input against the random channel effects. The performance of the proposed method against different oversampling ratios is evaluated and compared with cumulant based classifier for various numbers of samples.

This paper organized as follows. In Section 2, we introduce the system model, provide some background about CNNs and STNs, and describe our proposed architecture. The performance evaluation and comparison between the existing methods and the proposed one is discussed in Section 3. Finally, Section 4 concludes the paper.

II. Proposed Classifier

A. System Model

Let us assume that the transmitted modulated signal is

$$s(t) = \sum_k \alpha_k p(t - kT), \tag{1}$$

where α_k is the k-th complex symbol of the input data, $p(t)$ is the pulse shaping filter, and T is the symbol duration. At the receiver side, the signal can be modeled as

$$r(t) = \exp(-j2\pi\Delta f t)s(t-\tau) * h(t) + q(t) \tag{2}$$

where τ is the timing offset, Δf is the frequency error or offset, $h(t)$ is the channel impulse response, and $q(t)$ is the added noise. The receiver samples the received signal $r(t)$ with the sampling rate $1/T_s = M/T$, where M is the oversampling ratio. The purpose of modulation classification is to find the modulation type of the symbols α_k given N samples of $r(t)$.

B. Classifier Structure

Our proposed classifier consists of a Spatial Transformer Network, followed by six convolutional layers, and finally two dense neural network layers. The STN is used to improve the robustness of our classifier against the randomness introduced by the channel. The convolutional network efficiently extracts features from the input, while the dense neural network outputs the classification decision.

Contrast to regular Artificial Neural Networks (ANN), CNN takes advantages of the structure of the input whether it is 2D

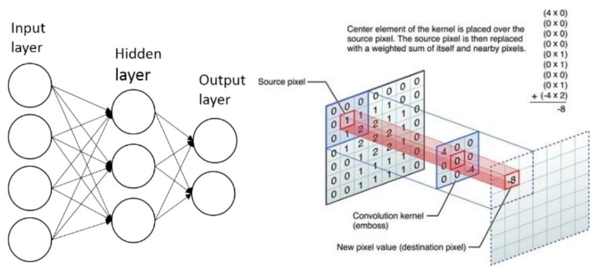

Fig. 1. Regular ANN and CNN.

or 3D. In our problem, the received signal is $1 \times N$ complex vector which is treated as $2 \times N$ real vectors as an input to CNN. A CNN consists of two main blocks: convolutional and linear block. The convolutional block is responsible of extracting features from 2D or 3D input. It includes one or more convolution layers (kernels) followed by a subsampling layer. The second block is a fully connected layer which is responsible for collecting extracted features for classification. In CNN, to learn weights, similar to regular neural networks, back propagation method is used. By using CNN, specific features of each modulated signals from raw samples are learned and then they are used for the classification of future unknown signals. Figure 1 shows a regular ANN and a CNN.

The received signal will be distorted by different channel and real-world effects like timing, phase errors, and multipath fading. To counter these effects, in addition to the CNN, we used STN at the input of the CNN to make network spatially-invariant. STNs are made from three parts which are shown in Figure 2. The first part is the *localisation network* which is responsible for generating the parameters of spatial transformer. These transformation parameters are then used to create a sampling grid. The sampler uses this grid to generate the transformed output [8].

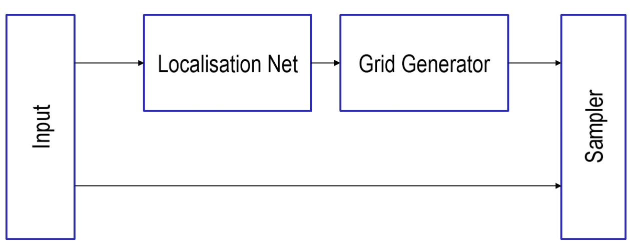

Fig. 2. Spatial Transformer Network.

Our network structure was inspired by networks used to classify the MNIST dataset in computer vision domain. Figures 3 and 4 show the proposed CNN and STN architectures. To prevent overfitting, we used several kinds of regularization. Dropout is used after first dense layer in both STN and CNN. In addition, L1 norm penalty is used in cost definition. Cross entropy loss function is defined for loss function and Adam/SGD [9] solver is used for training.

As shown in Figure 3, we used a CNN composed of 6

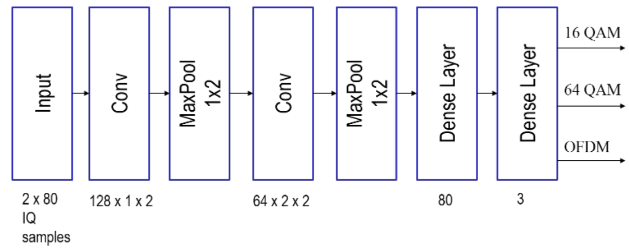

Fig. 3. Convolutional Neural Network structure.

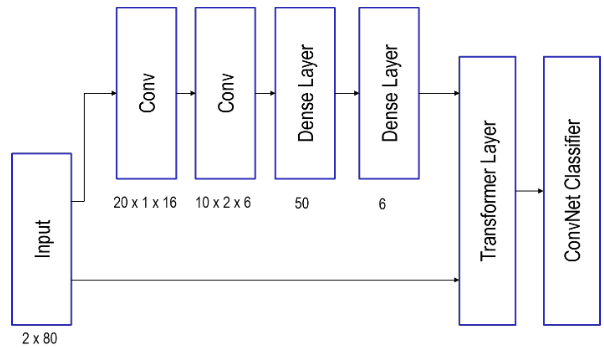

Fig. 4. Spatial Transformer Network structure.

layers; two convolutional neural layers, each followed by a max pooling layer, and two dense layers at the final stage. First convolutional layer includes 128 kernels and each kernel has dimension of 1×2 and second convolutional layer is consisted of 64 kernels of 2×2. All layers use the rectified linear (ReLu) activation function except the output layer which uses a SoftMax activation function.

For STN, as shown in Figure 4, we used a 5 layers network composed of 2 convolutional layers followed by two dense layers. All these layers use a ReLu activation function.

III. Results

The proposed classifier has been evalauted on a dataset that includes realistic channel impairments: timing, frequency offsets, and multipath fading. In particular, the radioml dataset [10] provides labeled samples using various modulation techniques based on a realistic channel model. It is built using GNU Radio and its source code is available online. Unfortunately, radioml does not consider OFDM signals, and it is limited only to an oversampling factor of 8 for all signals. Hence, we generated new datasets based on the same code to include OFDM and oversampling factors of 2,4,6, and 8. We also modified the channel model by increasing the the timing errors. The timing error model follows a clipped random walk process to simulate sample clock drift. We increased the standard deviation of the process to 1 and its maximum value to 10% of the sampling rate.

In this work, we perform experiments designed to better understand the effect of the STN. The first experiment is based on CNN without STN, and the second one uses CNN with

TABLE I
TRAINING AND TEST ACCURACIES

Oversampling	Training Accuracy	Test Accuracy
2	0.72	0.65
4	0.66	0.63
6	0.70	0.63
8	0.61	0.6

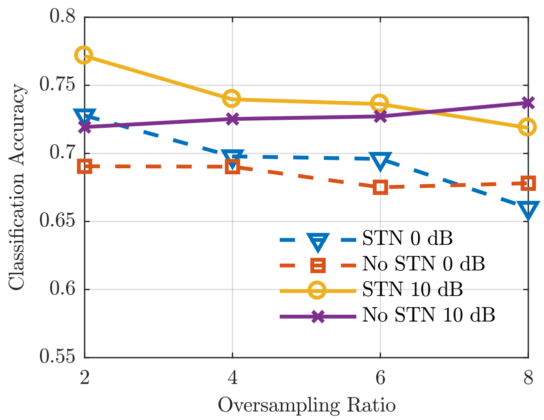

Fig. 5. Overall Classification Accuracy against oversampling factor for SNR 0 and 10 dB.

STN. Additionally, we compare our method against cumulants based classifier. The comparision is divided into two parts. First, we try to understand the effect of the insertion of STN under various oversampling ratios. Second, we analyze the performance of our proposed method for different SNRs and compare it against cumulants.

A. Impact of STNs

The overall training and test accuracies for each oversampling ratio dataset are shown in Table I. From this table, we observe that the overall test accuracy does not exceed 66%, which is acceptable given that the SNRs range starts from -20dB and timing error is high. Tables II – V show the confusion matrices obtained for each oversampling ratio for the entire dataset. From these tables, we can see that most of misclassifications occur between QAM16 and QAM64, while the OFDM confusion is relatively low.

The test accuracy obtained from the proposed classifier was compared against the convolutional neural network without STN for different SNRs. The results are shown in Figure 5. From this figure, we can observe that the insertion of STN improves performance for low oversampling ratios. This is because under low oversampling ratios, the timing errors add a significant distortion to the signal. Hence, the STN is able to counter its effect, thus providing improved performance.

B. SNR Performance Comparison

To better understand the effect of SNR on the performance, we show the accuracy of the test samples against SNR. In

TABLE II
CONFUSION MATRIX OVERSAMPLING RATIO 2

		Predicted			
		16-QAM	64-QAM	OFDM	All
Actual	16-QAM	11796 (0.48)	8400	4004	24200
	64-QAM	8264	11895 (0.47)	4741	24900
	OFDM	634	48	25218	25900 (0.97)
	All	20694	20343	33963	75000

TABLE III
CONFUSION MATRIX OVERSAMPLING RATIO 4

		Predicted			
		16-QAM	64-QAM	OFDM	All
Actual	16-QAM	11386 (0.47)	9094	3720	24200
	64-QAM	9242	11246 (0.45)	4412	24900
	OFDM	652	381	24867 (0.96)	25900
	All	21280	20721	32999	75000

TABLE IV
CONFUSION MATRIX OVERSAMPLING RATIO 6

		Predicted			
		16-QAM	64-QAM	OFDM	All
Actual	16-QAM	9284 (0.38)	10899	4017	24200
	64-QAM	7412	12761 (0.51)	4727	24900
	OFDM	222	468	25210 (0.97)	25900
	All	16918	24128	33954	75000

addition the ”No STN” convolutional neural network, we compare our work against cumulants. It is known that for cumulants to give good results, large number of samples might be required, especially in low SNR. Given that our CNN classifiers use only 80 samples, for cumulants test we extended our dataset to 160 and 320 samples. Figures 6 – 9 show the results. From these figures, we can see that both ”No STN CNN” and ”STN+CNN” outperform cumulants for all input sizes at low SNR. Only at a high oversampling ratio and when using the large samples cumulants performance become comparable to that of neural networks as seen in Figure 9. Comparing neural network with and without STNs, we see that for low oversampling ratio the added STN improves the performance. This improvement is more significant at high SNR for low oversampling ratios. The higher the oversampling ratio, the less the effect of the STN, and at an oversampling ratio of 8, we see that the STN seems to add more distortion to the signals, thus making the performance fluctuate with SNR.

IV. CONCLUSION

A modulation classification technique based on supervised learning using convolutional neural network and spatial transformer network is proposed. The STN was shown to enhance the performance of a CNN for low oversampling ratio datasets. In comparison to the cumulants technique, the proposed clas-

TABLE V
CONFUSION MATRIX OVERSAMPLING RATIO 8

		Predicted			
		16-QAM	64-QAM	OFDM	All
Actual	16-QAM	17974 (0.74)	1676	4550	24200
	64-QAM	17678	1808 (0.07)	5414	24900
	OFDM	202	0	25698 (0.99)	25900
	All	35854	3484	35662	75000

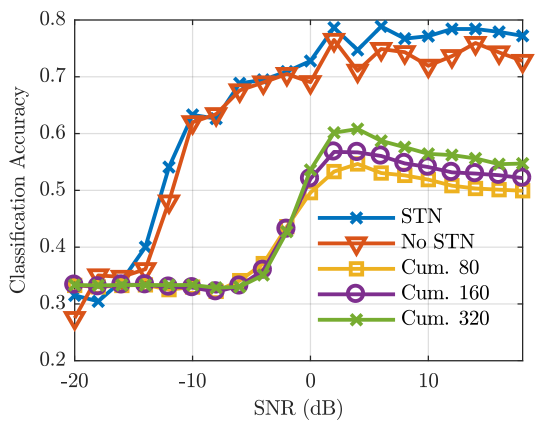

Fig. 6. Classification Accuracy against SNR for an oversampling factor of 2.

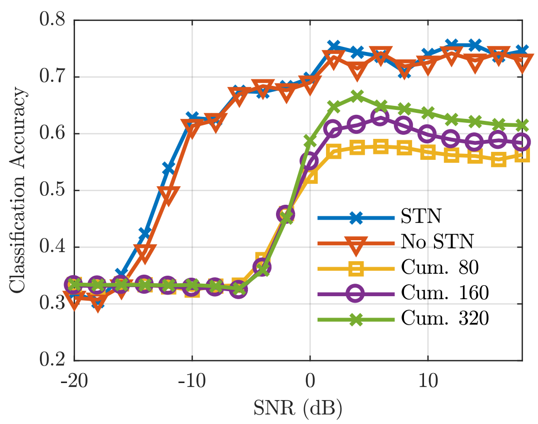

Fig. 7. Classification Accuracy against SNR for an oversampling factor of 4.

sifier is able to achieve a higher recognition accuracy using a smaller number of samples accross wide range of SNRs.

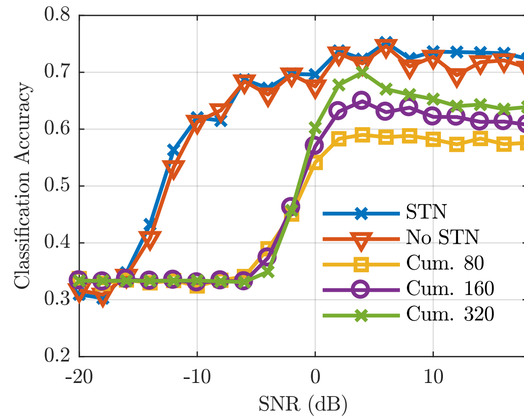

Fig. 8. Classification Accuracy against SNR for an oversampling factor of 6.

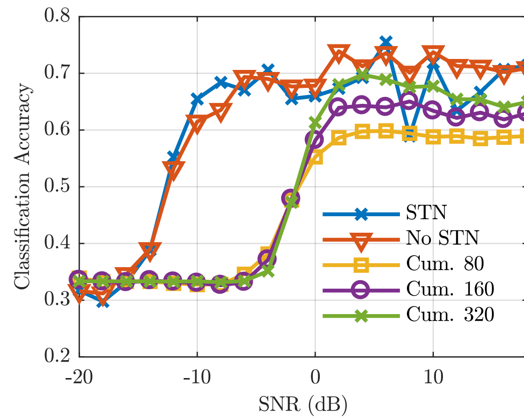

Fig. 9. Classification Accuracy against SNR for an oversampling factor of 8.

REFERENCES

[1] Z. Zhu and A. K. Nandi, *Automatic Modulation Classification: Principles, Algorithms and Applications*, 1st ed. Chichester, West Sussex, United Kingdom: Wiley, Feb. 2015.

[2] A. Swami and B. M. Sadler, "Hierarchical digital modulation classification using cumulants," *IEEE Transactions on Communications*, vol. 48, no. 3, pp. 416–429, Mar. 2000.

[3] O. A. Dobre, M. Oner, S. Rajan, and R. Inkol, "Cyclostationarity-Based Robust Algorithms for QAM Signal Identification," *IEEE Communications Letters*, vol. 16, no. 1, pp. 12–15, Jan. 2012.

[4] P. Urriza, E. Rebeiz, and D. Cabric, "Optimal Discriminant Functions Based on Sampled Distribution Distance for Modulation Classification," *IEEE Communications Letters*, vol. 17, no. 10, pp. 1885–1888, Oct. 2013.

[5] M. Petrova, P. Mhnen, and A. Osuna, "Multi-class classification of analog and digital signals in cognitive radios using Support Vector Machines," in *2010 7th International Symposium on Wireless Communication Systems*, Sep. 2010, pp. 986–990.

[6] T. J. OShea, J. Corgan, and T. C. Clancy, "Convolutional Radio Modulation Recognition Networks," in *Engineering Applications of Neural Networks*, ser. Communications in Computer and Information Science. Springer, Cham, Sep. 2016, pp. 213–226. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-44188-7_16

[7] T. J. O'Shea, L. Pemula, D. Batra, and T. C. Clancy, "Radio Transformer Networks: Attention Models for Learning to Synchronize in Wireless Systems," *arXiv:1605.00716 [cs]*, May 2016, arXiv: 1605.00716. [Online]. Available: http://arxiv.org/abs/1605.00716

[8] M. Jaderberg, K. Simonyan, A. Zisserman, and K. Kavukcuoglu, "Spatial Transformer Networks," *arXiv:1506.02025 [cs]*, Jun. 2015, arXiv: 1506.02025. [Online]. Available: http://arxiv.org/abs/1506.02025

[9] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," *arXiv:1412.6980 [cs]*, Dec. 2014, arXiv: 1412.6980. [Online]. Available: http://arxiv.org/abs/1412.6980

[10] T. O'Shea and N. West, "Radio Machine Learning Dataset Generation with GNU Radio," *Proceedings of the GNU Radio Conference*, vol. 1, no. 1, 2016. [Online]. Available: https://pubs.gnuradio.org/index.php/grcon/article/view/11