UC San Diego

Title
Deep Modulation Recognition in an Unknown Environment

Permalink
https://escholarship.org/uc/item/8164d3jj

Authors
Li, Lei
Peng, Qihang
Cosman, Pamela C
et al.

Publication Date

DOI

Deep Modulation Recognition in an Unknown Environment

Lei Li[*], Qihang Peng[*†], Pamela, C. Cosman[†], and Laurence B. Milstein[†]

[*]*University of Electronic Science and Technology of China*
Chengdu, China, Email: anniepqh@uestc.edu.cn
[†]*Electrical and Computer Engineering Department, University of California, San Diego*
San Diego, California 92093-0407, Emails: {pcosman, lmilstein}@ucsd.edu

Abstract—Deep modulation recognition has demonstrated high classification accuracy when a neural network is trained on large-scale datasets. However, when applied in an unknown environment where there are not any ground-truth labels in collected data, its performance can be significantly degraded. In this paper, we propose incorporating an adversarial discriminative neural network to adapt the deep modulation recognition to an unknown environment. Results show that, when the neural network is trained under an AWGN channel but applied under a frequency-selective Rayleigh fading channel, the adversarial network based domain adaptation can achieve comparable performance with that of the network trained with sufficiently large labeled data.

Index Terms—modulation recognition, unknown environment, frequency-selective fading, neural networks, domain adaptation.

I. Introduction

Deep neural networks emerge as increasingly powerful tools for end-to-end classification tasks in areas such as computer vision and natural language processing. Motivated by this remarkable success, deep neural networks have been applied in modulation recognition, and have been shown to achieve higher classification accuracy than conventional modulation recognition [1] [2].

However, this is obtained by assuming that the collected radio signal samples (also called data in this paper) used for training the network are independent and identically distributed with those used for testing. That is, there is a sufficiently large amount of training data with ground-truth labels for the scenario where the deep modulation recognition is applied. When the test data follows a different distribution from the training data, the performance of the deep modulation recognition can be significantly degraded [3].

Wireless propagations can differ significantly in time, frequency, and space. Accordingly, received radio signals vary in distributions under different propagation conditions [4]. It is difficult to obtain labeled data to train the network for all possible communications environments. So there naturally arises a question: how to ensure the deep modulation recognition performance when it is operating in an unknown environment where there are no labeled training data.

In this paper, we propose unsupervised domain adaptation [5] with an adversarial network [6] [7] [8] to address this problem. For brevity, we use *source domain* to represent the dataset used for training the neural network, and *target*

domain to denote the dataset collected for testing, i.e., the environment that the deep modulation recognition is applied. We are interested in the scenario where the signal in the target domain is collected in an unknown environment, and is different from the source domain. In other words, there is a domain shift between the source and target domain. The adversarial domain adaptation architecture consists of a discriminator, a source encoder and a target encoder [8]. The source encoder is a convolutional neural network (CNN) that is pre-trained by labeled source domain. The target encoder is initialized with the same parameters as the source encoder, and fine-tuned such that the discriminator could not reliably distinguish between the encoded source and target data. In this way, the target domain is adapted to the shared feature space with the source by the target encoder. Note that we use different CNN structures for modulation recognition here than those used in [8] for image classification. Results show that when the source domain is under an AWGN channel and the target domain is under frequency-selective fading channel, the proposed unsupervised domain adaptation can achieve comparable performance with that of the network trained with a sufficiently large labeled dataset in the target domain.

The rest of this paper is organized as follows. Modulation recognition with deep neural networks is described in Section II. In Section III, unsupervised domain adaptation is given. Simulation results of deep modulation recognition in an unknown environment where there are not any ground-truth signal labels are presented in Section IV. Finally, conclusions are discussed in Section V.

II. Modulation Recognition Using Deep Neural Networks

Modulation recognition can be formulated as a classification problem, where the number of modulation types corresponds to the number of classes [9]. Recently, deep neural networks have emerged as powerful tools for classification in image processing [10], and also shown to exhibit potential in modulation recognition [11]. In this paper, modulation recognition is achieved using a deep neural network, which is described in the following.

The received radio signals are filtered and down converted to baseband with a carrier frequency roughly centered on the carrier of interest, and then the in-phase and quadrature

TABLE I
NEURAL NETWORK LAYOUT

| Layer	Kernel Size	Stride	Output Shape
Input			[1, 2, 1024]
Conv1	3 * 7	2	[32, 1, 512]
Conv2	1 * 3	1	[64, 1, 512]
Max Pooling	1 * 2	1	[64, 1, 256]
Conv3	1 * 3	1	[128, 1, 256]
Max Pooling	1 * 2	1	[128, 1, 128]
Conv4	1 * 3	1	[256, 1, 128]
Max Pooling	1 * 2	1	[256, 1, 64]
Conv5	1 * 3	1	[512, 1, 64]
Max Pooling	1 * 2	1	[512, 1, 32]
Conv6	1 * 3	1	[512, 1, 32]
Average Pooling	1 * 32	1	[512, 1, 1]
FC			[256]
FC			[11]

Fig. 1. The deep domain adaptation process, the aim is to train target encoder to make the source domain feature distribution and target domain feature distribution as similar as possible.

components are sampled at discrete time steps to form a $1 \times N$ complex-valued vector [12]. This complex-valued vector is further decomposed as a $2 \times N$ real-valued vector, where the first row corresponds to the in-phase components and the second row as the quadrature components. Let $\mathbf{x}^{(i)}$ denote the vector collected in the i-th observation interval for modulation recognition, and $y^{(i)}$ be its corresponding label denoting the particular modulation type. Then $\{\mathbf{x}^{(i)}, y^{(i)}\}$ forms the input to the deep neural network.

A deep CNN is used for modulation recognition in this paper, which consists of 6 convolutional layers followed by two fully connected (FC) layers. The stride for the first convolutional layer is 2, and 1 for the other convolutional layers. Pooling is used after each convolutional layer except the first one, where max pooling is used for layers from the second to the fifth, and average pooling is adopted for the last convolutional layer. Tanh activation is used for the first convolutional layer, and ReLU activation is used for the other convolutional layers. Batch normalization [13] is used in convolutional layers for faster convergence. Dropout [14] is used in the first FC layer to reduce overfitting, followed by a SeLU activation function. The last FC layer has M neurons corresponding to the M modulation classes. Table I illustrates the artchitecture of the CNN used in this paper.

The CNN outputs the predicted class value $\hat{y}^{(i)}$. Then a loss function (here we use categorical cross-entropy for modulation recognition) can be calculated as

$$\mathcal{L} = -\frac{1}{M} \sum_{i=0}^{M} [y^{(i)} log(\hat{y}^{(i)}) + (1 - y^{(i)}) log(1 - \hat{y}^{(i)})] \quad (1)$$

An Adam optimizer is utilized with a learning rate of 0.001 to update the parameters by back propagation and gradient descent for supervised learning.
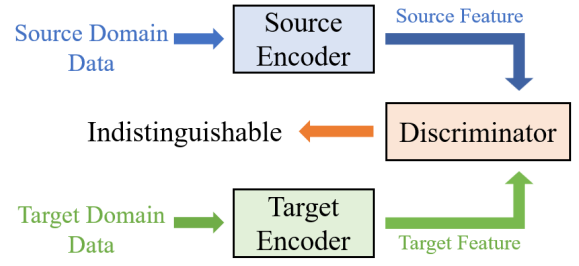
III. DEEP MODULATION RECOGNITION IN AN UNKNOWN ENVIRONMENT WITH UNSUPERVISED DOMAIN ADAPTATION

We are interested in the scenario where the unknown environment is different from the labeled data training the neural network. That is, there is domain shift between the target domain and the source domain. Further, it is assumed that the data in the target domain is unlabeled. Given this scenario, directly applying the neural network, which is trained with a sufficient amount of labled data in the source domain, would lead to modulation recognition performance degradation due to the domain shift.

Unsupervised domain adaptation with an adversarial network, also called adversarial domain adaptaion [7], is proposed in this paper, which aims to adapt the target domain to the same distribution as the source domain. As shown in Fig. 1, the adversarial domain adaptation consists of a source encoder, a target encoder and a discriminator [8].

The source encoder is from the CNN in Table I. Specifically, we pre-train the CNN with sufficient labeled source data, and then split it into two parts as shown in Table II: a source encoder and a classifier, where the source encoder consists of 6 convolutional layers and the first fully connected layer, and the classifier corresponds to the second fully connected layer. The target encoder has the same network structure as that of the source encoder, and is initialized with the same parameters as the source encoder. The discriminator, which is used to discriminate between the source and the target feature distributions, consists of three fully connected layers, where the first and the second layer have 512 hidden neurons each, and the last has 2 neurons.

We train the discriminator to maximize the probability of assigning the correct label to both source encoder and target encoder outputs. We simultaneously freeze the parameter of source encoder and train the target encoder to minimize the mapping loss from target domain to source domain. In this way, the discriminator eventually could not distinguish whether the input is from the source or the target encoder.

For better convergence of the adversarial training process, the learning rate of the discriminator is set to 0.0001 and that of the target encoder is 0.00005. Besides, the parameters of the target encoder are updated every five updates of the

TABLE II
DOMAIN ADAPTATION ARCHITECTURE

	Layer	Output Shape
Source/Target Encoder	Conv Layers +1st FC Layer	[256]
Classifier	FC	[11]
Discriminator	FC	[512]
	FC	[512]
	FC	[2]

discriminator. It is expected that, after sufficient training, the target feature representation and the source will not be distinguishable.

Note that normalization is important for this unsupervised domain adaptation, since the average receiving power of signals due to different propagation conditions would vary. In this case, we employ a data pre-processing algorithm to normalize average power of input vector $\mathbf{x}^{(i)}$. Specifically, each sample in $\mathbf{x}^{(i)}$ is divided by the square root of the average power of $\mathbf{x}^{(i)}$.

IV. RESULTS AND ANALYSIS

In this section, we present the simulation results for deep modulation recognition in an unknown environment where there are sufficient labeled source data but not any labels for the target data. Specifically, the source domain corresponds to the scenario where the signals are in the presence of an AWGN channel, and the target domain is the scenario that the signals experience different delay spread of multipath fading.

The radio signals for both the source domain and the target domain are generated in the same way as those in the GNU Radio ML dataset RML2016.10a [15], and 11 different modulations including both analog and digital modulation types are generated: BPSK, QPSK, 8PSK, PAM4, QAM16, QAM64, GFSK, CPFSK, WBFM, AM-DSB, and AM-SSB. A square root raised cosine filter is used for pulse shaping, and $2 \times N$ real samples are collected to form one example, where the first and second column correspond to the in-phase and quadrature samples for the received signal, respectively, and N is set to 1024. We generate 1000 examples for each SNR and for each modulation format in both source and target domains.

For clarity of descriptions, notations used hereafter are elaborated in the following:

• "Source domain model" corresponds to the performance when the network is trained by the source domain but applied in the target domain.

• "Target domain model" means the network is trained using the target domain dataset with sufficient ground-truth labels and also tested in the same domain.

• "Domain adapted model" represents the performance when the network is trained using the source domain dataset
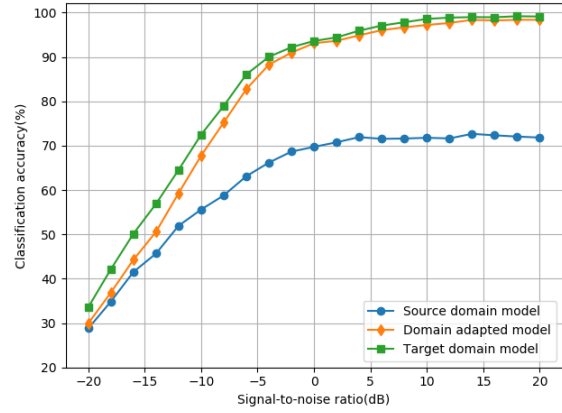

Fig. 2. Deep modulation recognition accuracy under different SNRs in flat fading channel. Source domain: AWGN channel, target domain: flat fading channel.
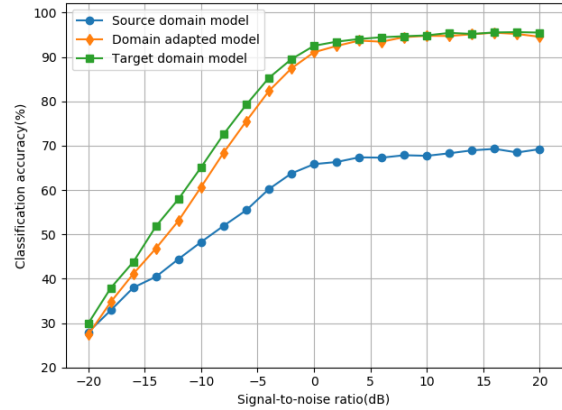

Fig. 3. Deep modulation recognition accuracy under different SNRs in frequency-selective fading channel. Source domain: AWGN channel, target domain: frequency-selective fading channel.

and adapted for the target domain using our proposed adversarial discriminative domain adaptation. This adaptation does not require any labels for the target domain dataset.

The classification accuracy versus different SNR is plotted in Fig. 2, where the source domain signal is under AWGN and the target domain signals experience flat fading with 3 discrete paths. It is shown that directly using the network trained by the source domain leads to the worst performance, and the classification accuracy is below 75%, even when the SNR is higher than 10dB. The performance using the proposed unsupervised domain adaptation method achieves performance very close to the target domain model, i.e., the upper bound on what the deep neural network can achieve.

We increase the domain shift by increasing the delay spread to be much larger than that in Fig. 2 with the same number of discrete paths to simulate a frequency-selective fading environment, and the results are shown in Fig. 3. Similar observations can be made that the performance with domain adaptation can also approach that of the target domain

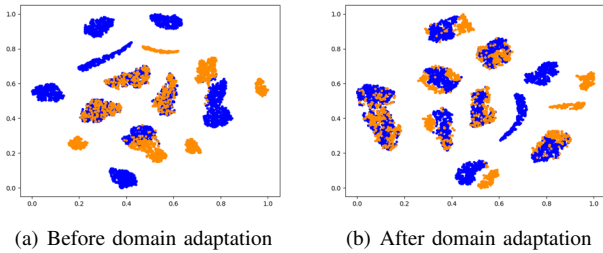(a) Before domain adaptation   (b) After domain adaptation

Fig. 4. t-SNE embedding of source and target domain data for before and after domain adaptation, where the blue dots represent the source domain feature embedding and the orange dots represent the target domain feature embedding. (a) source and target data visualized by t-SNE before domain adaptation (b) source and target data visualized by t-SNE after domain adaptation.

model, and outperforms the source domain classifier results. Meanwhile, it is shown that directly using the network trained by source domain leads to the worst performance.

Deep Modulation recognition seems to be a black box due to the use of a deep neural network. To better understand the domain adaptation results for deep modulation recognition, we visualize both the source domain data and the target domain data using t-distributed stochastic neighbor embedding(t-SNE) [16]. t-SNE is commonly used in machine learning for visualization and is a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data in a low-dimensional space of two or three dimensions.

For a more clear visualization, we choose the source data and the target data with an SNR of 20dB for the 11 different modulation types. The t-SNE embedding is illustrated in Fig. 4, where the blue and orange dots represent the embedded source domain and the target domain data, respectively. Fig. 4(a) illustrates the t-SNE embedding before domain adaptation. It is seen that, without domain adaptation, the target domain feature distribution is not aligned with the source domain. In this way, the network trained with source domain data results in significant performance degradation when tested in the target domain. With the proposed domain adaptation method, as illustrated in Fig. 4(b), the blue dots and the orange ones are significantly more aligned together than that in Fig. 4(a). This coincides with the results in Fig. 3 that the modulation recognition performance with domain adaptation is significantly improved as compared with the "source domain model".

V. Conclusions

An unsupervised domain adaptation with an adversarial network is proposed to address the problem of deep modulation recognition performance degradation, which arises when the neural network trained using source domain data is tested in a new radio propagation environment, and when there are only unlabeled data in the new environment. Results have shown that, with the proposed method, the modulation recognition performance is comparable with that of the network trained with a sufficiently large labeled dataset.

References

[1] N. E. West and T. O'Shea, "Deep architectures for modulation recognition," in *2017 IEEE International Symposium on Dynamic Spectrum Access Networks (DySPAN)*. IEEE, 2017, pp. 1–6.

[2] T. J. OShea, T. Roy, and T. C. Clancy, "Over-the-air deep learning based radio signal classification," *IEEE Journal of Selected Topics in Signal Processing*, vol. 12, no. 1, pp. 168–179, 2018.

[3] B. Luo, Q. Peng, P. C. Cosman, and L. B. Milstein, "Robustness of deep modulation recognition under awgn and rician fading," in *2018 52nd Asilomar Conference on Signals, Systems, and Computers*. IEEE, 2018, pp. 447–450.

[4] A. Goldsmith, *Wireless communications*. Cambridge university press, 2005.

[5] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," *arXiv preprint arXiv:1412.3474*, 2014.

[6] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," in *Advances in neural information processing systems*, 2014, pp. 2672–2680.

[7] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *The Journal of Machine Learning Research*, vol. 17, no. 1, pp. 2096–2030, 2016.

[8] E. Tzeng, J. Hoffman, K. Saenko, and T. Darrell, "Adversarial discriminative domain adaptation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 7167–7176.

[9] A. K. Nandi and E. E. Azzouz, "Algorithms for automatic modulation recognition of communication signals," *IEEE Transactions on communications*, vol. 46, no. 4, pp. 431–436, 1998.

[10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in neural information processing systems*, 2012, pp. 1097–1105.

[11] T. J. OShea, J. Corgan, and T. C. Clancy, "Convolutional radio modulation recognition networks," in *International conference on engineering applications of neural networks*. Springer, 2016, pp. 213–226.

[12] K. W. Martin, "Complex signal processing is not complex," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 51, no. 9, pp. 1823–1836, 2004.

[13] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *arXiv preprint arXiv:1502.03167*, 2015.

[14] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *The journal of machine learning research*, vol. 15, no. 1, pp. 1929–1958, 2014.

[15] T. J. O'shea and N. West, "Radio machine learning dataset generation with gnu radio," in *Proceedings of the GNU Radio Conference*, vol. 1, no. 1, 2016.

[16] L. v. d. Maaten and G. Hinton, "Visualizing data using t-sne," *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.