High-Performance Deep Learning Classification for Radio Signals

Ahsen J. Uppal[1], Michael Hegarty[1], William Haftel[1], Phil A. Sallee[2], H. Brown Cribbs III[2], H. Howie Huang[1]

[1] The George Washington University [2]The Raytheon Company

***Abstract*—The ability to classify different types of signal modulations in radio transmissions is an important task with applications in defense, networking, and communications. This process has traditionally been done manually by human analysts. Recent advances have shown that applying deep learning methods to this task is feasible. But existing recognition networks are complex, with heavy computational requirements, and poor accuracy on some modulation types and in noisy environments.**

We have built a robust radio frequency signal classifier with a hybrid approach that uses images derived from signal constellation and spectrogram data, combined with an efficient convolutional neural network. Compared to the state-of-the-art deep learning classifier, our system obtains better accuracy, with lower computational requirements.

I. Introduction

During radio telecommunication, the transmitter varies one or more properties of a carrier signal based on the content of the message. Diverse modulation schemes vary different properties of the carrier. Typically properties include magnitude, frequency, or phase.

The rapid and accurate identification of the modulation of arbitrary radio signals, especially with minimal human intervention is an important task for many real-world applications such as spectrum sharing, dynamic access, regulatory policy enforcement, and many defense and security applications. This task is challenging because RF data can be voluminous, received signals may have poor strength. Traditional approaches often rely on hand-crafted feature extractors and hand-tuned statistical discriminators. Such approaches are costly to develop and have limited accuracy. Furthermore, related variants of modulation types often have similar features that can easily confuse classifiers.

Deep learning (DL) approaches have been applied extensively in big data analytics in areas such as speech recognition[1], language translation[2], bioinformatics[3], drug design[4], medical diagnosis[5], and image recognition and classification in computer vision[6], [7], [8]. But the application of DL techniques to communications systems is less well-developed. Recent advances have shown that that radio signal classifiers based on deep learning (DL) techniques can be effective for signal recognition[9], [10] and can also significantly increase the accuracy of classification[11]. Furthermore, DL-based signal classification approaches can learn features directly from data.

But challenges remain: deep learning frameworks require extensive computational resources to train and classify signals. Our goal is to improve the accuracy and computational performance (particularly during classification) of DL based signal classifiers. We have developed a prototype system which improves the speed and accuracy of DL based classification by exploiting frequency domain information at the front end to simplify processing at subsequent layers. Our system generates both IQ constellation and spectrogram images from time domain input data and uses them as inputs to a convolutional neural network (CNN).

II. Proposed Method

We implemented our classifier using a Convolutional Neural Network (a Neural Network designed to extract features from images with multiple dimensions)and relies on techniques common to image-processing DL methods. It uses images derived from frequency-domain transformations – constellation and spectrum – and processing multiple pixels at once, then concatenating both processing chains, before the final classification decision occurs.

A spectrogram is a 2D representation of the frequency-dependent magnitudes over time. An IQ constellation plot is a 2D radial scatterplot of the FFT bin values at one fixed frequency bin, with each points radius and angle corresponding to the complex magnitude and phase. Our model uses both to simplify the CNN design.

Attempting to classify a signal using just the spectrogram or IQ constellation is ineffective, because certain modulation types have very limited information with each image type as illustrated in Figure 1, Figure 2,. For example, different frequency modulation types (such as FM, FSK, GMSK) have no discernible pattern in the IQ constellation, and different PSK types have no pattern in the spectrogram. But using both types of images as inputs allows for accurate classification, with a less complex neural net structure than in current needed by state-of-the-art methods. Similar approaches using just constellation data [12] are effective, but only for modulation types that vary by amplitude and phase such as PSKs and QAMs. In contrast, our model is capable of accurately classifying modulations that also vary by frequency.

A. Model Comparison

The DeepSig model [11] is illustrated in Table I. This is a traditional CNN configuration with many layers of convolutions and pooling, leading to several rounds of dense layers that classify the input signal. This model takes the 1024 by 2

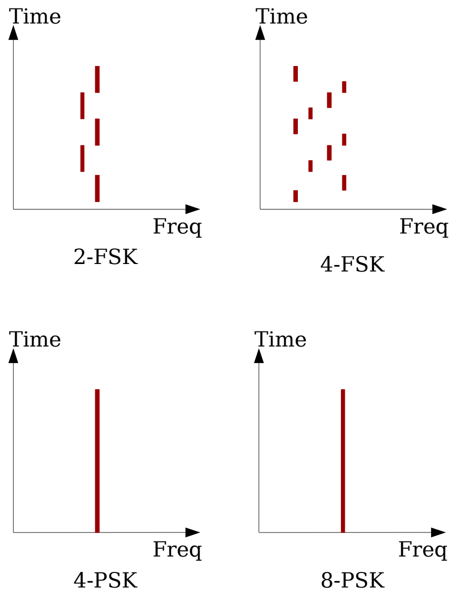

Fig. 1. Spectrogram images can be used to distinguish FSK variants that vary in frequency, but not PSK variants that vary in phase.

TABLE I. BASELINE CNN MODEL LAYOUT

Layer	Output Dimensions
Input	2 x 1024
Conv	64 x 1024
Max Pool	64 x 512
Conv	64 x 512
Max Pool	64 x 256
Conv	64 x 256
Max Pool	64 x 128
Conv	64 x 128
Max Pool	64 x 64
Conv	64 x 64
Max Pool	64 x 32
Conv	64 x 32
Max Pool	64 x 16
Conv	64 x 16
Max Pool	64 x 8
Dense / Selu	128
Dense / Selu	128
Dense / Softmax	24

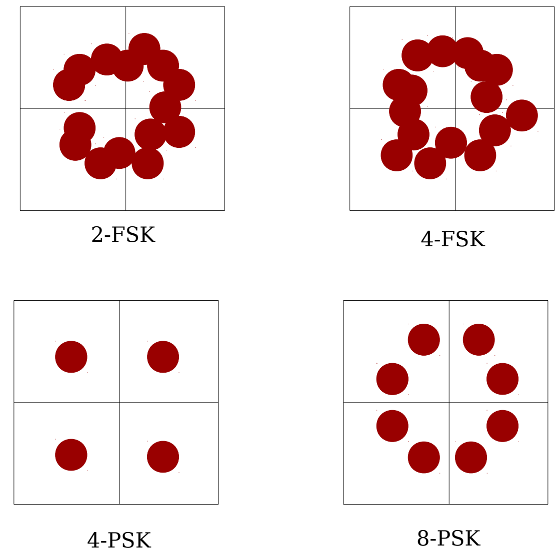

Fig. 2. IQ Constellation images can be used to distinguish PSK variants that vary in phase, but not FSK variants that vary in frequency.

TABLE II. OUR CNN MODEL LAYOUT

Time-domain Input		
Spectrogram		
Depthwise Conv		
Max Pool		Time-domain Input
Conv		IQ Constellation
Max Pool		Conv
Conv		Max Pool
Dropout		Dropout
Flatten	→ Concatenate ←	Flatten
	Dense / Selu	
	Dense / Selu	
	Dense / Selu	
	Dense / Selu	
	Dense / Softmax	

time series data as direct input, and slowly pools the input in half repeatedly.

In contrast, our model shown in Table II takes spectrogram and constellation images of the signal as inputs, and performs similar convolutional processing on the both of them, but with fewer layers and filters, then passes the concatenated data into dense layers to classify the signal. The images are faster to generate using FFTs than doing the equivalent operations using convolutional layers. In our design, we generated 8 non-overlapping 128-point FFTs for each 1024 sample input. From these FFTs we derived an 8x128 magnitude-only spectrogram and a 128x128 constellation image.

III. EXPERIMENTAL DATA AND RESULTS

We implemented our system using Keras[13], a Python deep learning framework with a TensorFlow[14] backend. For comparison we also implemented DeepSig's model using Keras with the same backend.

We evaluated our model using DeepSig's publicly available training data set[11] (improving upon tools described in [15]) of 2M signals, covering 24 different modulation types. The classes are:OOK, 4ASK, 8ASK, BPSK, QPSK, 8PSK, 16PSK, 32PSK, 16APSK, 32APSK, 64APSK, 128APSK, 16QAM, 32QAM, 64QAM, 128QAM, 256QAM, AM-SSB-WC, AM-SSB-SC, AM-DSB-WC, AM-DSB-SC, FM, GMSK, and OQPSK.

This dataset was made to be challenging for a test to clas-

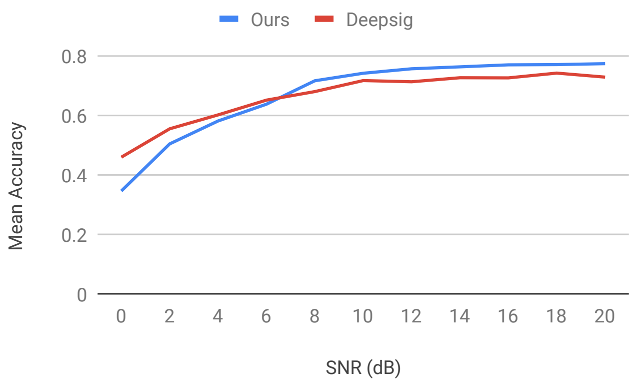

Fig. 3. Model mean accuracy across SNRs vs. baseline Deepsig implementation.

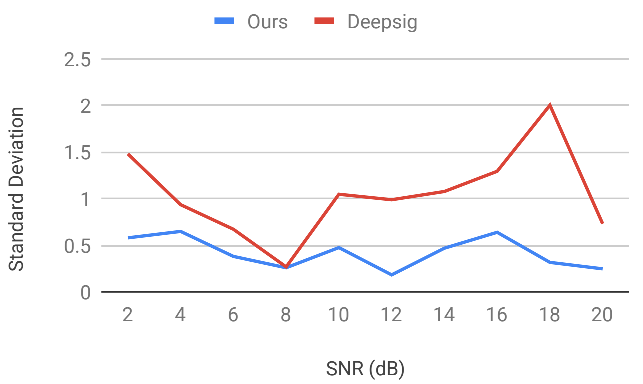

Fig. 4. Model standard deviation of accuracy across SNRs vs. baseline Deepsig implementation.

sify, with some normal class signals, as well as more difficult classes (QAM and PSK flavors beyond 16, AM flavors), that are not easily discerned from each other. Additionally signal durations are very limited in time with 1024 complex samples. Short duration signals are difficult to classify, but useful in cases such as a rapidly scanning receiver.

We created a model using our approach, and compared against Deepsigs established neural network model. We trained our model with training data at a single SNR of 20 dB and we tested with across varying SNRs (from 2 to 20 dB) to compare the accuracy of classification, time to train, and time to classify of each model. These tests were run using a single Nvidia Tesla K40c GPU [16] on a Xeon E5-2683 server with 56 cores and 512GB RAM using Keras version 2.20 and Tensorflow version 1.8.0.

A. Results

Our model performs consistently better at signal to noise ratios of 6 dB and above. Model consistency, as indicated by the accuracy standard deviation is significantly better than DeepSig's.

We also evaluated the time to train each CNN, as well as the time to classify. The time to train with 80k signals, and the total time to evaluate 1k signals are shown in Figure 5

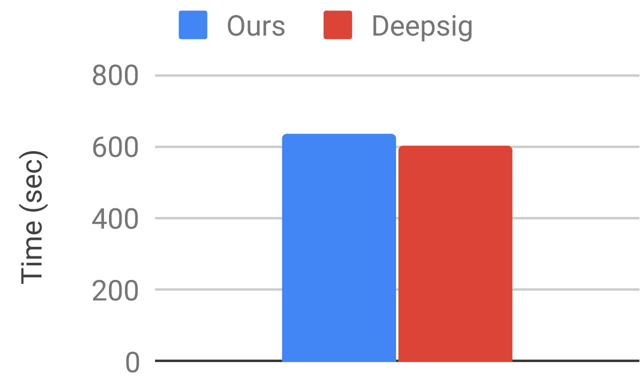

Fig. 5. Time to train our model and reference with 80k signals. Our model took 635.9 seconds compared to DeepSig's 601.1 seconds.

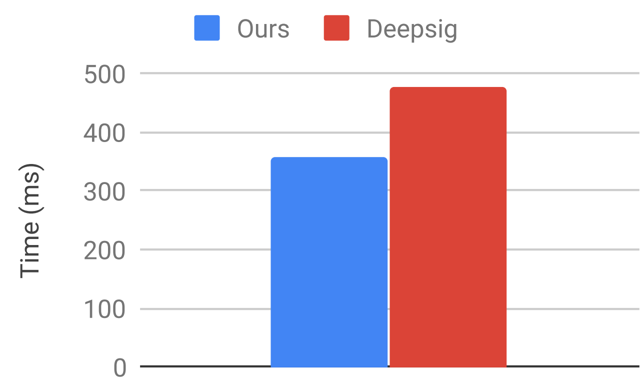

Fig. 6. Time to classify 1k signals using our model and reference model. Our model took 359.27 ms compared to 476.9 ms for Deepsig, yielding 1.11x faster performance for classifications.

and Figure 6. The difference in training performance here is small, but our model obtains a 1.11x speedup in classification time. Note that it should also be possible to operate each branch of our model in parallel. We will explore this and other computational enhancements as future work.

We also obtained confusion matrices shown in Figure 7 and Figure 8 for each implementation. These describe the relative

TABLE III. ACCURACY MEANS AND STANDARD DEVIATIONS FOR EACH MODEL.

	Ours		DeepSig	
SNR	Mean	Std	Mean	Std
0	0.347	0.579	0.460	1.481
2	0.505	0.648	0.556	0.935
4	0.582	0.379	0.602	0.671
6	0.638	0.257	0.652	0.267
8	0.717	0.474	0.681	1.047
10	0.742	0.181	0.718	0.988
12	0.757	0.467	0.714	1.077
14	0.764	0.638	0.727	1.294
16	0.770	0.315	0.727	2.004
18	0.771	0.246	0.743	0.730
20	0.774	0.366	0.729	1.897

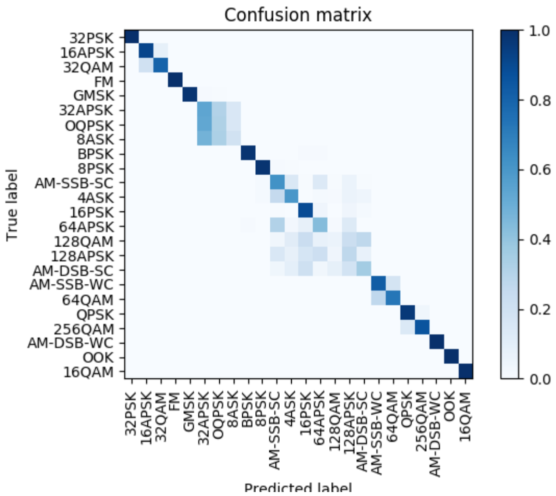

Fig. 7. Confusion matrix from baseline model.

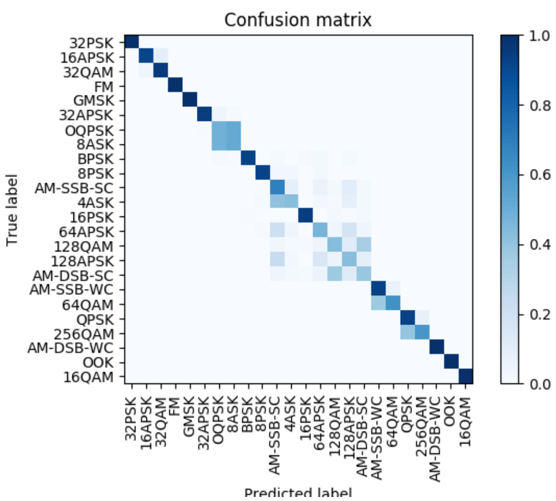

Fig. 8. Confusion matrix from our model.

accuracy of each neural network against each signal type. Notably, our implementation more accurately discriminates between different PSK variants.

IV. Conclusion

We have described our new RF signal classification method. We have developed a prototype that shows significant accuracy and speed gains over the state-of-the-art baseline implementation.

V. Acknowledgment

Funding for this work was supported in part by the Raytheon Company.

References

[1] D. Amodei, S. Ananthanarayanan, R. Anubhai, J. Bai, E. Battenberg, C. Case, J. Casper, B. Catanzaro, Q. Cheng, G. Chen *et al.*, "Deep speech 2: End-to-end speech recognition in english and mandarin," in *International conference on machine learning*, 2016, pp. 173–182.

[2] Y. Wu, M. Schuster, Z. Chen, Q. V. Le, M. Norouzi, W. Macherey, M. Krikun, Y. Cao, Q. Gao, K. Macherey *et al.*, "Google's neural machine translation system: Bridging the gap between human and machine translation," *arXiv preprint arXiv:1609.08144*, 2016.

[3] S. Wang, S. Sun, Z. Li, R. Zhang, and J. Xu, "Accurate de novo prediction of protein contact map by ultra-deep learning model," *PLoS computational biology*, vol. 13, no. 1, p. e1005324, 2017.

[4] E. Gawehn, J. A. Hiss, and G. Schneider, "Deep learning in drug discovery," *Molecular informatics*, vol. 35, no. 1, pp. 3–14, 2016.

[5] W. Sun, B. Zheng, and W. Qian, "Computer aided lung cancer diagnosis with deep learning algorithms," in *Medical imaging 2016: computer-aided diagnosis*, vol. 9785. International Society for Optics and Photonics, 2016, p. 97850Z.

[6] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein *et al.*, "Imagenet large scale visual recognition challenge," *International journal of computer vision*, vol. 115, no. 3, pp. 211–252, 2015.

[7] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[8] Z. Yu and C. Zhang, "Image based static facial expression recognition with multiple deep network learning," in *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*. ACM, 2015, pp. 435–442.

[9] 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.

[10] S. Rajendran, W. Meert, D. Giustiniano, V. Lenders, and S. Pollin, "Deep learning models for wireless signal classification with distributed low-cost spectrum sensors," *IEEE Transactions on Cognitive Communications and Networking*, vol. 4, no. 3, pp. 433–445, 2018.

[11] 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.

[12] S. Peng, H. Jiang, H. Wang, H. Alwageed, Y. Zhou, M. M. Sebdani, and Y.-D. Yao, "Modulation classification based on signal constellation diagrams and deep learning," *IEEE transactions on neural networks and learning systems*, no. 99, pp. 1–10, 2018.

[13] F. Chollet *et al.*, "Keras," 2015.

[14] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard *et al.*, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} Symposium on Operating Systems Design and Implementation ({OSDI} 16)*, 2016, pp. 265–283.

[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] NVIDIA Corporation, "NVIDIA Tesla Supercomputing Solutions," www.nvidia.com/en-us/data-center/tesla/.