PROCEEDINGS OF SCIENCE

Stochastic vs. BFGS Training in Neural Discrimination of RF-Modulation

Maria Dima[a]*, Mihai-Tiberiu Dima[a], Madalina Mihailescu[b]

a *Joint Institute for Nuclear Research,*
Joliot Curie 6, Dubna, Moscow region, Russia-141980

b *Hyperion University,*
Calea Calarasi 169, Bucharest, Romania-030615

E-mail: mmdima@jinr.ru

Neuromorphic classification of RF-Modulation type is an on-going topic in SIGINT applications. Neural network training approaches are varied, each being suited to a certain application. For exemplification I show the results for BFGS (Broyden-Fletcher-Goldfarb-Shanno) optimization in discriminating AM vs FM modulation and of stochastic optimization for the challenging case of AM-LSB vs. AM-USB (upper / lower sideband) discrimination. Although slower than BFGS, the stochastic training of a neural network avoids better local minima, obtaining a stable neurocore.

The 6th International Workshop on Deep Learning in Computational Physics (DLCP2022)
6-8 July 2022
JINR, Dubna, Russia

*Speaker

1. Introduction

Signals Intell (SIGINT) is a branch of the military and civilian intell services monitoring (mainly) the RF communications. Automated Modulation Classifiers (AMC's) are of 2 types: likelihood classifiers (LC's) [2] and feature classifiers (FC's) [3]. LC's use a likelihood function on the received signal, while FC's neuromorphic software for feature extraction. LC methods have high CPU demand and need prior information from transmitters. FC's do not require this, however perform relatively well. They consist of (i) feature extraction - parameters constructed from amplitude, frequency, and phase distributions [4]. Features from advanced processing, such as Fourier (FFT) and wavelet transforms [5], or high-order statistical cumulants [6] require longer signal samples and are CPU intensive, for instance noise jammed signals can be analysed with the FFT of the cyclic autocorrelation function [7] and decrypted. Secondly, (ii) classification - such as: linear, k-means [8], clustering algorithms, neural software [9] and support vector machine (SVM) with kernels [10]. Typical identification purities are 95% [11] (S/N of 0dB) for a variety of deep-learning methods, and 90% [12] (S/N of -10dB). Existing methods assume equal signal-to-noise (S/N) in the training and witness sets.

2. Signal conditioning and feature creation

The reconstruction of signal we have presented before [1] and here we will just briefly review it. We consider to have as fundamental wave:

$$u(t) = p + A\sin(2\pi f t + \Phi) \qquad (1)$$

where *f*, *p*, *A* are slowly varying functions of time.

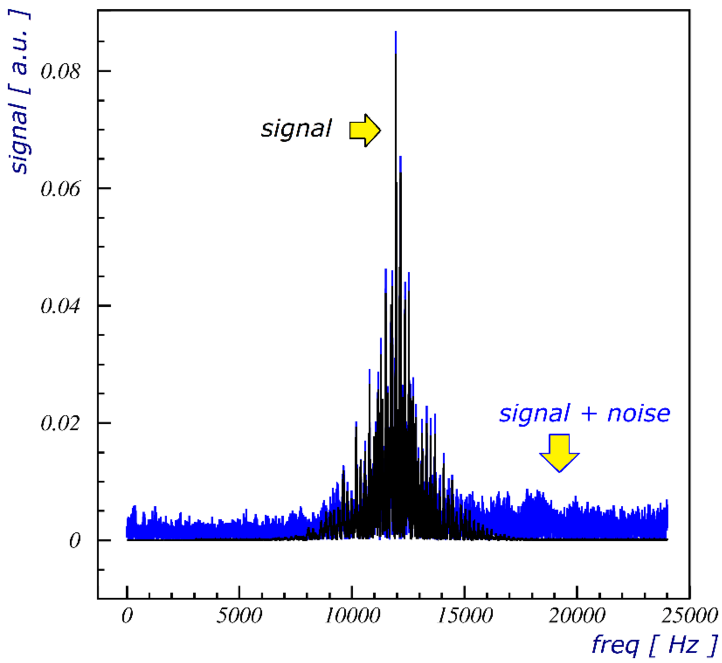

Figure 1. Spectrum of signal with noise (blue) compared to just signal without noise (black), around the intermediate frequency of 12 kHz.

Pedestal reconstruction - for this the simplest method is to take an between t_i and t_f:

$$\langle u \rangle = p + A_e \, sin\left(2\pi f \frac{t_i + t_f}{2} + \Phi\right) sinc\left(\pi f \left(t_i - t_f\right)\right) \tag{2}$$

where $sinc(x) = sin(x) / x$ and $A_e = Asinc(f)$ - with   the duration of one sample. Since the *sin* term does not vanish, we try to zero the sinc term. For t_f-t_i = $n\Delta$, the n ≅ m/fΔ, condition must be met, with m ∈ N. Basically m is scanned until the relation gives a close-enough integer (in our case n = 11) - and then $p = \langle u \rangle$.

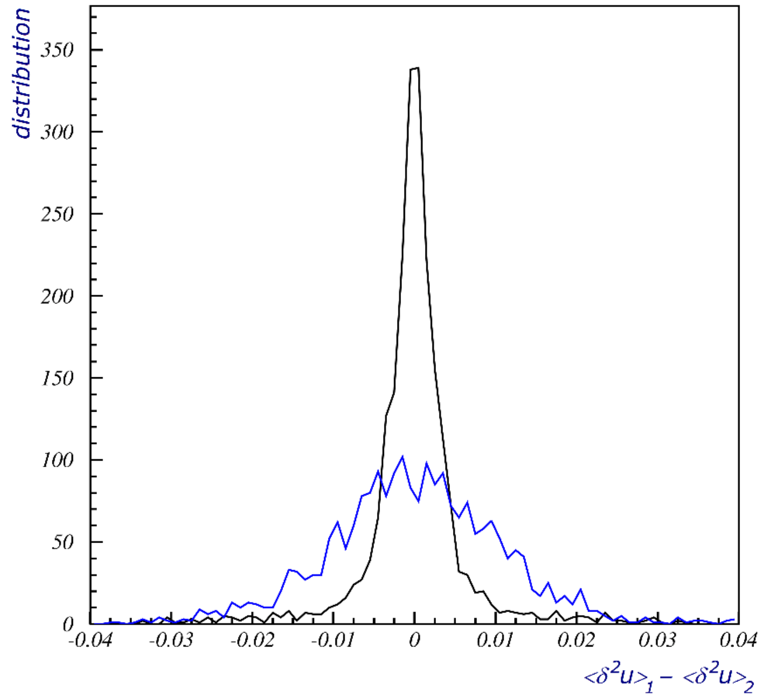

Figure 2. Spectrum of signal with noise (blue) compared to just signal without noise (black), around the intermediate frequency of 12 kHz.

Amplitude reconstruction - following the same idea, we designed a formula for the amplitude:

$$\langle \delta^2 u \rangle = A_e^2 \langle sin^2 \rangle - A_e^2 \langle sin \rangle^2 \simeq \frac{1}{2} A_e^2 \tag{3}$$

Through a coincidence the "magic number" n for double the frequency is very similar to the n = 11 from pedestal determination and we can use the same loop for the averages.

Frequency reconstruction - similar to amplitude, we determined frequency with an O_{II} dipolar moment, which basically differentiates the sine wave, where k indexes u(t-kΔ). We used k=1:

$$\begin{aligned} \langle u(u - u_{k\Delta}) \rangle &= pA_e(\langle sin \rangle - \langle sin_{k\Delta} \rangle) + A_e^2(\langle sin^2 \rangle - \langle sin \cdot sin_{k\Delta} \rangle) \\ &\simeq A_e^2 sin^2(\pi f k\Delta) \end{aligned} \tag{4}$$

Phase reconstruction - similar to frequency we determined phase with an O_{II} dipolar moment taking reference to a fixed phase sine:

$$\langle u \cdot cos \rangle \simeq \frac{1}{2} A_e sin\Phi \tag{5}$$

where for $\Phi = 0$ the error may be significant, however phase is not absolute, rather relative to the previous sample's phase, as such such errors tend to systematically cancel out.

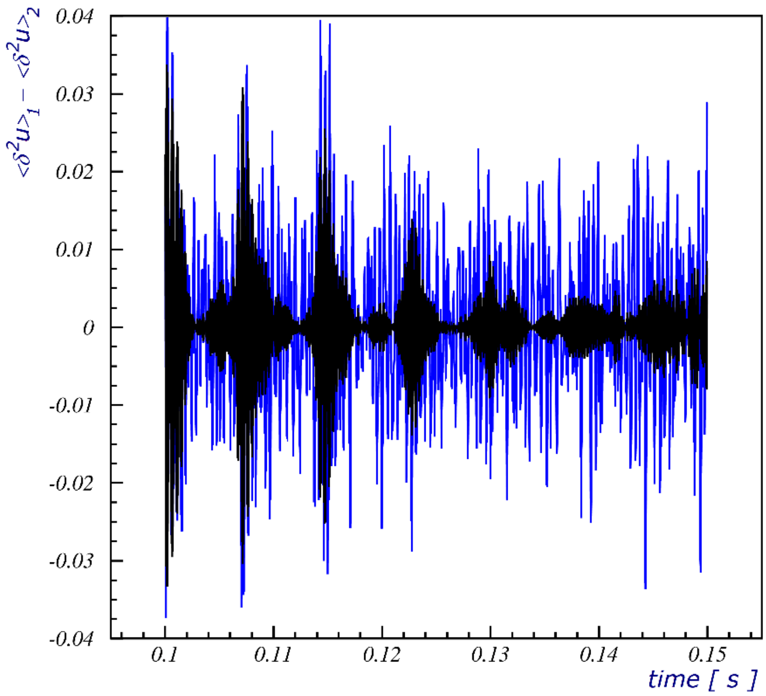

Figure 3. Distribution of $<\delta^2 A>$ vs. time: in blue the signal with noise, compared to just signal without noise (in black).

Feature creation - having pedestal, amplitude, frequency and phase for each sample, we accumulated these quantities in histograms looking for discriminating features. We devised similar other parameters (in number of 12) to capture the differences between various modulation types. As an example, the figure above shows the distribution of the full width at ¼ maximum for the phase distribution histograms - with red, again, AM modulation and with blue, FM modulation.

3. Elimination of perturbance due to noise

The improvement over our previous work [1] was to consider noise in the data:

$$u(t) = p + noise + Asin(2\pi f t + \Phi) \tag{6}$$

A simplistic view would be to eliminate the noise, however this is a difficult, if not impossible task and we can concede, that noise is not that important, what we more specifically need is to discriminate between the different types of modulation in the presence of noise. For this we proceed with the same method used in the no-noise case.

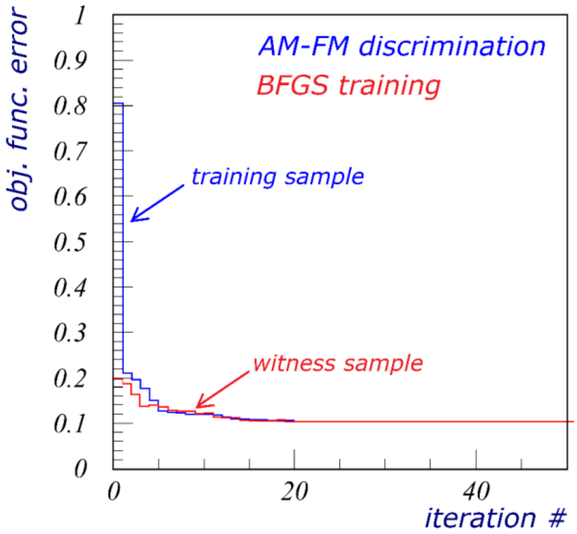

Figure 4. Training evolution of network error versus iteration (epoch) number for Broyden-Fletcher-Goldfarb-Shanno (BFGS) training, in the case of an AM vs. FM binary classifier.

Pedestal reconstruction

$$\langle u\rangle = p + \langle noise\rangle + A_e\, sin\left(2\pi f\frac{t_i + t_f}{2} + \Phi\right) sinc\left(\pi f\big(t_i - t_f\big)\right) \quad (7)$$

Amplitude reconstruction

$$\langle\delta^2 u\rangle = \langle\delta^2 noise\rangle + A^2\langle\delta^2 sin\rangle + 2\langle\delta(noise)\delta A\rangle \quad (8)$$

$$2\langle\delta^2 u\rangle_1 - 2\langle\delta^2 u\rangle_2 = A_1^2 - A_2^2 = \delta A^2 \quad (9)$$

Although for the pedestal itself we have <noise> = 0, in practice the noise induces a sizeable fluctuation of the pedestal, depending on its amplitude.

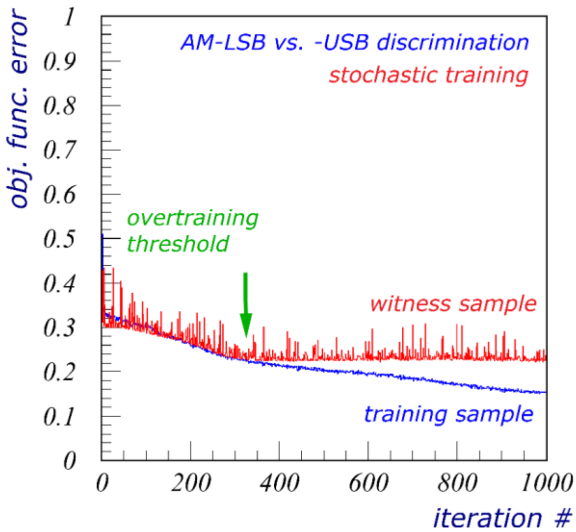

Figure 5. Training evolution of network error versus iteration (epoch) number for stochastic training, in the case of an AM vs. FM binary classifier.

However, the noise term vanishes when we consider the signal variation between two adjacent samples, mainly because the rms of the noise does not vary that rapidly, and by subtraction cancels out.

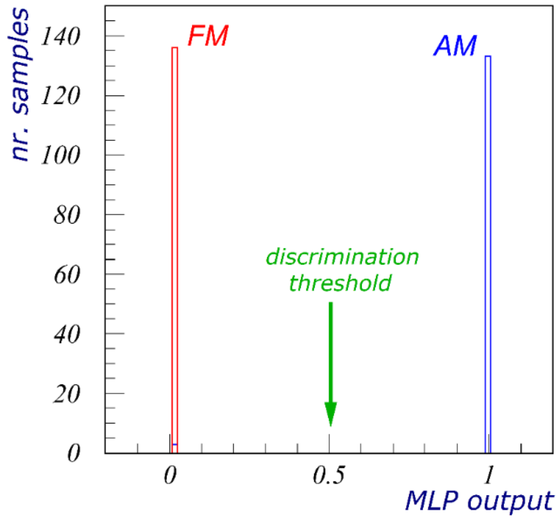

Figure 6. Performance of the BFGS trained classifier, essentially perfect classification.

To study this we add noise to our signal – figure 1, and proceed with the data reconstruction as outlined in equations (7) and (9). The noise has the effect of broadening the $<\delta^2 A>$ distribution, figure 2 (blue with noise, black without).

However, the aim is not to eliminate noise, an impossible task, but to discriminate between RF-modulation types, in the presence of noise.

The better understanding of this aspect is seen from figure 3, where we followed the $<\delta^2 A>$ distribution in time.

It is obvious that in the passages of low volume, the noise will prevail and induce amplitude rms, $<\delta^2 A>$, however, the low volume passages are not characteristic for any of the RF-modulation types under consideration.

Rather, the higher modulation passages confer the characteristics needed to discern between RF-modulation types.

As such, it is an excellent result that with a noise "pedestal" x10-15 times higher than our $<\delta^2 A>$ quantity, we can faithfully reproduce at least the higher-modulation passages in this quantity.

This analysis holds similar relevance for the differences in frequency rms $<\delta^2 f>_1$ - $<\delta^2 f>_2$ and phase rms, $<\delta^2 \Phi>_1$ - $<\delta^2 \Phi>_2$.

4. Deep learning training and conclusion

Using the same set of 12 parameters previously reported in [1], we trained BFGS - Broyden-Fletcher-Goldfarb-Shanno training for AM vs. FM – figure 4, and stochastic training for AM-LSB vs. AM-USB – figure 5, deep learning neural networks for binary sets of modulations.

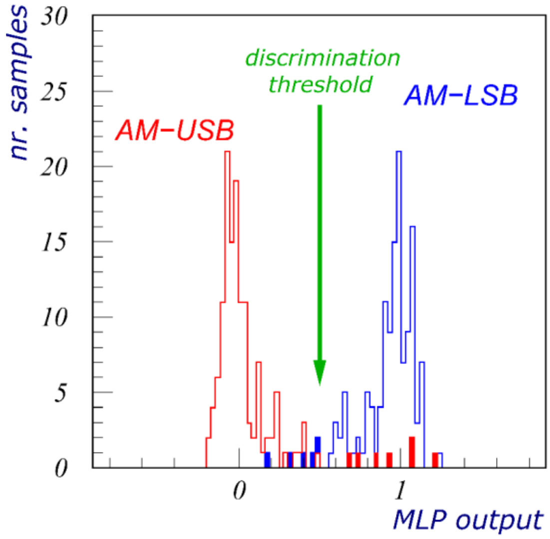

Figure 7. Performance of the statistically trained classifier, with 95% classification purity of AM-LSB and AM-USB modulations (impossible to attain with the classical statistical-moments method in the presence of noise).

The results were again excellent, AM vs. FM confirming that noise plays a negligible role, and the very difficult case AM-LSB vs. AM-USB (essentially the same type of modulation) again marginally sensitive to noise (due to the slanted frequency spectra, LSB rising with frequency, while USB sinking with frequency – slope determined by the high-modulation passages, not the low-modulation ones affected by noise). It is our opinion that this result would be impossible to attain with the classical statistical-moments method in the presence of noise.

We conclude that our noise reduction methods are performing very well and that our previously organization of the training programme [1] into BFGS for structurally different types of RF-modulation and statistical for very similar modulations (the AM family) is the correct procedure also in the case of signal containing noise.

References

[1] M. Dima, T. Dima, *Deep Learning for Automatic RF-Modulation Classification*, Proceedings 9th "Distributed Computing and Grid Technologies in Science and Education" (GRID'2021), Dubna, Russia, p. 498, July 2021.

[2] J.L. Xu, W. Su, M. Zhou, *Likelihood-Ratio Approaches to Automatic Modulation Classification*, IEEE Trans. Syst. Man Cybernet. Part C Appl. Rev. **41**, 455–469 (2011).

[3] A. Hazza, M. Shoaib, S. A. Alshebeili, A. Fahad, *An overview of feature-based methods for digital modulation classification, Communications, Signal Processing, and their Applications*, ICCSPA 1st International Conference IEEE, document 6487244, pp. 1–6, (2013).

[4] EE. Azzouz, AK. Nandi, *Automatic identification of digital modulation types, Signal Process*. **47**, 55–69 (1995); JJ. Popoola, R. van Olst, *A novel modulation sensing method*, IEEE Veh. Tech. Mag. **6**, pp. 60–69, (2011).

[5] J. Liu, Q. Luo, *A novel modulation classification algorithm based on daubechies5 wavelet and fractional fourier transform in cognitive radio*, 14th IEEE ICCT Communication Technology Conference, document/6511199, pp. 115–120, (2012); Y. Lv, Y. Liu, F. Liu, J. Gao, K. Liu, G. Xie,

Automatic modulation recognition of digital signals using CWT based on optimal scales, IEEE Conference On Computer and Information Technology, document/6984692, pp. 430–434, (2014).

[6] D. Das, A. Anand, P. K. Bora, R. Bhattacharjee, *Cumulant based automatic modulation classification of QPSK, OQPSK, π/4-QPSK and 8-PSK in MIMO environment*, IEEE Conference On Signal Processing and Communications, document/7439996, pp. 1–5, (2016); A. Hazza, M. Shoaib, A. Saleh, A. Fahd, *Robustness of digitally modulated signal features against variation in HF noise model*, EURASIP J. Wirel. Commun. Netw. 2011, 24, (2011).

[7] U. Satija, M. Manikandan, B. Ramkumar, *Performance study of cyclostationary based digital modulation classification schemes*, IEEE Conference on Industrial and Information Systems, document/7036609, pp. 1–5, (2014).

[8] M. W. Aslam, Z. Zhu, A. K. Nandi, *Automatic modulation classification using combination of genetic programming and KNN*, IEEE Trans. Wirel. Commun. **11**, 2742–2750, (2012).

[9] K. Hassan, I. Dayoub, W. Hamouda, M. Berbineau, *Automatic modulation recognition using wavelet transform and neural network*, IEEE Conference on Intelligent Transport Systems Telecommunications, document/5399351, pp. 234–238, (2009).

[10] V. Orlic, M. Dukic, *Multipath channel estimation algorithm for automatic modulation classification using sixth-order cumulants*, Electron. Lett. **46**, 1348–1349, (2010).

[11] GJ. Mendis, J. Wei, A. Madanayake, *Deep learning-based automated modulation classification for cognitive radio*, IEEE Conference on Communication Systems, document/7833571, pp. 1-6, (2016).

[12] A. Dai, H. Zhang, H. Sun, *Automatic modulation classification using stacked sparse auto-encoders*, IEEE Conference on Signal Processing, document/7877834, pp. 248–252, (2016).

PoS(DLCP2022)011