PROCEEDINGS OF SCIENCE

SigCLR: A contrastive learning approach to unsupervised modulation recognition and novelty detection

Nicholas Bruce,[a,b,*] **Belaid Moa,**[b] **Stephen Harrison**[a] **and Peter F. Driessen**[b]

[a] *Herzberg Astronomy and Astrophysics Research Centre, National Research Council Canada, Penticton, BC, Canada*

[b] *Department of Electrical and Computer Engineering, University of Victoria, Victoria, BC, Canada*

E-mail: nicholas.bruce@nrc-cnrc.gc.ca

We introduce a contrastive learning architecture for radio datasets called SigCLR. We show how this method of unsupervised machine learning can be applied for a variety of tasks including but not limited to modulation recognition and novelty detection. We show results for a functioning modulation class detector, where an unknown signal can be passed into the trained network and it can be compared to a number of known modulations.

Radio Frequency Interference Conference (RFI2024)
14-18 October 2024
Bariloche, Argentina

[*] Speaker

https://pos.sissa.it/

PoS(RFI2024)025

1. Introduction

As radio environments become more crowded, and radio frequency interference (RFI) monitoring expands to wider bandwidths, a significant amount of RFI data being collected. At the Dominion Radio Astrophysical Observatory (DRAO), efforts have evolved from building a wideband RFI monitor [1], to detecting signals in wideband spectra [2], and extracting baseband time series of these signals [3] while discarding the rest of the wideband spectra. This process reduces the data volume and recovers time-domain characteristics of the interfering signals. These time-domain characteristics are lost in other common stored data types like integrated power spectra and occupancy.

This data collection method supports rich RFI science, such as RF site characterization and signal identification, without storing large data volumes. However, the lack of labels in this dataset complicates tasks like modulation recognition, feature vector extraction, and novelty detection.

We demonstrate a method for modulation detection on an unlabeled dataset containing 53 modulations, generated using TorchSig [4]. Self-supervised methods like contrastive learning [5], known for their accuracy in image detection, require large datasets, which Torchsig can provide.

We introduce SigCLR, a model architecture that adapts image-based contrastive learning to complex-valued radio time-series data from the TorchSig dataset.

2. Overview

Figure 1 illustrates the model's architecture and training process. Each time-series signal $s(t)$ is duplicated and augmented differently using RF-specific augmentations, a_i and a_j (e.g. gain drift, clipping, spectral inversion) from the same family $\mathcal{A}$. The augmented copies $s_1(t)$ and $s_2(t)$ are processed through identical encoders and projection heads. Their representations z_1 and z_2 are compared in a cross-entropy loss function to guide the network towards understanding that the signals should be considered identical. After training, the augmentation pipeline and projection head are removed, leaving the encoder for downstream tasks. We demonstrate its ability to cluster signal types, effectively serving as a modulation detector.

3. Architecture

Each copy of the input time-series signal undergoes a single, randomly chosen augmentation from the following: time-varying additive white Gaussian noise, random phase shift, time reversal, random time shift, gain drift, local oscillator drift, clipping, or spectral inversion.

We use an EfficientNet B4 encoder [6] adapted from image to complex time series data [4]. The encoder classification layer is removed and a neck added which customizes the encoder for a specific downstream task.

The projection head, a small neural network, decouples the latent representation space from contrastive learning. It prevents the encoder from optimizing for the contrastive task, allowing the encoder to focus on creating good input representations. After training, the projection head is discarded and the encoder is used for downstream tasks.

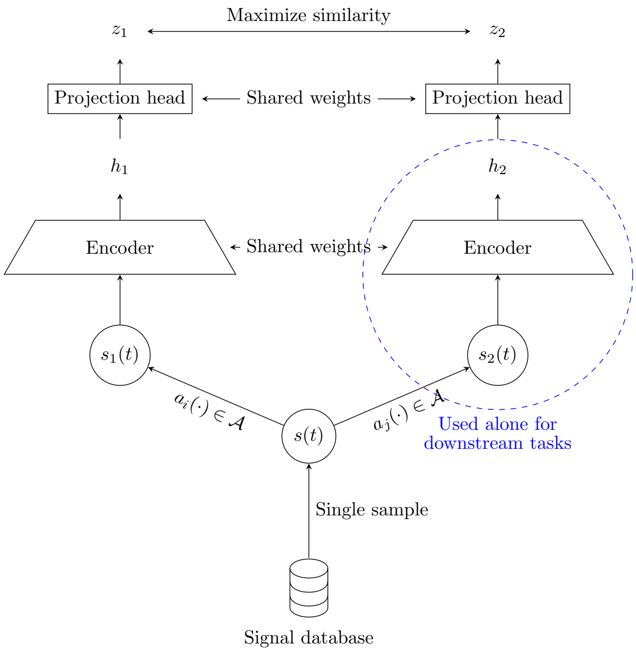


Figure 1: SigCLR model architecture

We use Cosine Similarity to measure similarity between two signal representations, z_i and z_j, at the output of the projection head:

$$sim(z_i, z_j) = \frac{1}{\tau} cos(\theta) = \frac{1}{\tau} \frac{z_i \cdot z_j}{||x_i||_2 \cdot ||x_j||_2}. \tag{1}$$

Here, the L2 norm, $||x||_2$, computes the magnitudes and τ (set to 0.07) acts as a temperature to penalize differing vectors and favor similar ones. We use Cross Entropy Loss on the similarities with summing reduction as the loss criterion, the same loss used by other contrastive methods [5].

4. Method

Using Torchsig we generate a training dataset of 530,000 unique signals evenly split across 53 modulation classes, each with 4096 complex time series samples. The validation dataset contains 106,000 unique signals with an even class divide.

The training is distributed across three compute nodes, each with four 32 GB NVidia V100 Volta GPUs. Each GPU processes a batch size of 64, resulting in a total effective batch size of 768. Training lasts for $\approx$ 800 epochs.

We use the AdamW optimizer with an initial learning rate of 0.001 and a Cosine Annealing Learning rate scheduler to avoid local minima. We experimented with a LARS optimizer but did not see a gain in performance. This was likely because of our relatively small batch size.

5. Results

To validate the training we demonstrate using the trained encoder as a modulation detector. Although the training was unsupervised, Torchsig provides labeled data. We generate two signals

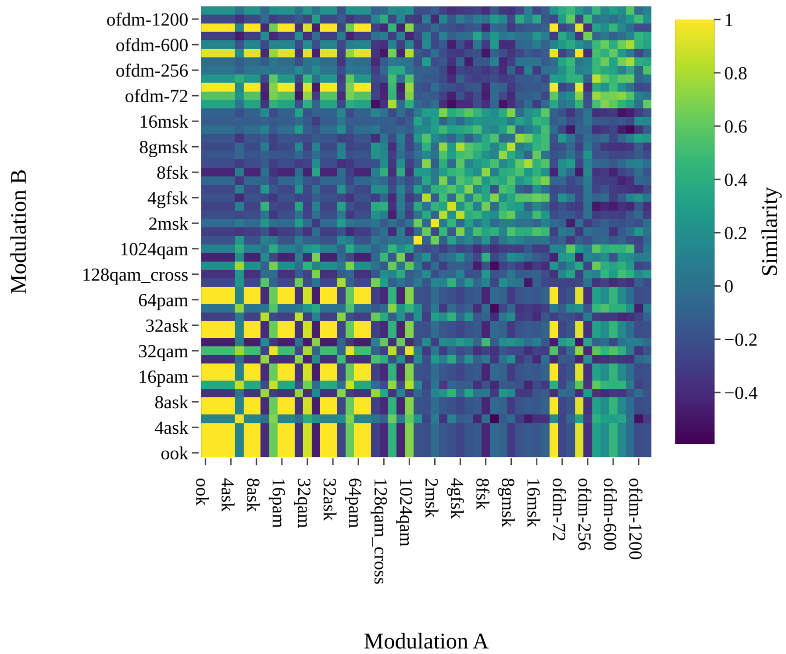


Figure 2: Similarity matrix for every modulation.

for each of the 53 modulations then compute the cosine similarity between each pair without temperature normalization, visualized in a non-symmetric similarity matrix fig. 2. The asymmetry arises because signals are divided into two buckets "A" and "B" and similarities are computed between them. This shows unsupervised modulation recognition but its performance is limited by the small dataset size.

We then group modulations: frequency-modulated signals under FSK, phase-modulated under PSK, amplitude-modulated under ASK, and OFDM signals kept distinct. The average similarity of each group is shown in fig. 3a. The simplified modulation detector performs better (fig. 3b), especially when OFDM is removed, as it is often confused with amplitude-modulated signals.

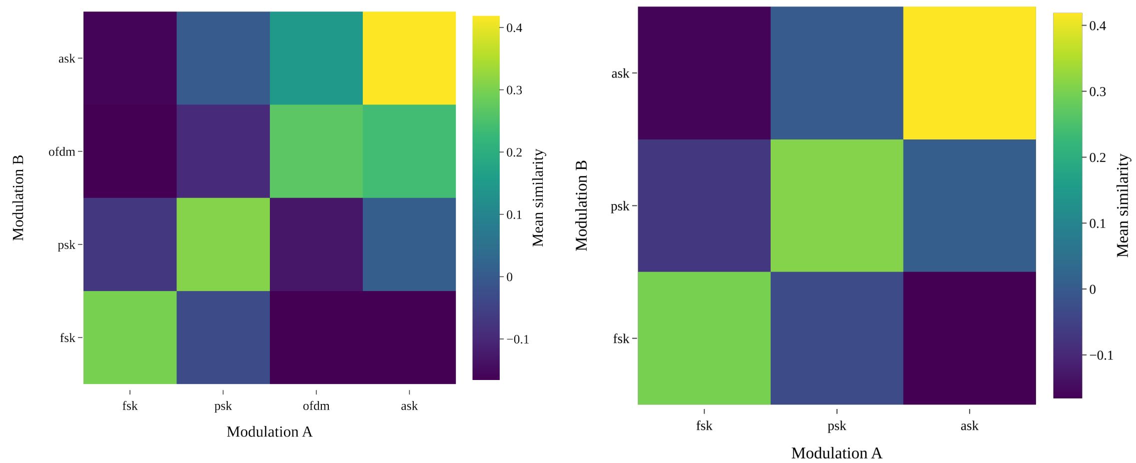
6. Conclusion

We demonstrate a method for training large models on unsupervised radio data using a contrastive learning architecture to build a simple modulation detector. We introduce the architecture and show results for classifying unknown signals as frequency, phase, or amplitude modulated.

To the best of our knowledge, this is the first completely unsupervised method for clustering similar signals. For modulation detection, this method can learn to group any signal type by only providing the signals itself, without containing a priori knowledge about any particular modulations.

Future plans include expanding the training dataset to 500,000 signals per modulation class, reducing the number of complex samples per signal from 4096 to 512, and using a ResNet-50 encoder adapted for time-domain data. We also plan to experiment with a DirectCLR architecture to prevent dimensional collapse during training [7].

Additionally, we aim to explore other architectures like SWaV [8] for unsupervised model development on large radio datasets.



(a) Modulation groups including OFDM.

(b) Modulation groups excluding OFDM.

Figure 3: Similarity matrices for simplified modulation detector.

References

[1] S. Harrison, A. Lakhani, J.-P. Lavigne, N. Bruce, T. Robishaw, R. Messing, C. Baard, M. A. Smith, B. Robert, and D. D. Rizzo, "RFI monitor system design for the Dominion Radio Astrophysical Observatory," in *XXXIV General Assembly and Scientific Symposium of the International Union of Radio Science*, 2021.

[2] S. Harrison, R. Coles, T. Robishaw, and D. Del Rizzo, "RFI novelty detection using machine learning techniques," in *RFI2019: Coexisting with Radio Frequency Interference*, 2019.

[3] N. Bruce, S. Harrison, and P. Driessen, "DDC Pool: Efficient down-conversion of signals for RFI monitoring," in *Proceedings for RFI 2022*, 2022.

[4] L. Boegner, M. Gulati, G. Vanhoy, P. Vallance, B. Comar, S. Kokalj-Filipovic, C. Lennon, and R. D. Miller, "Large scale radio frequency signal classification," 2022. arXiv:2207.09918.

[5] T. Chen, S. Kornblith, M. Norouzi, and G. Hinton, "A Simple Framework for Contrastive Learning of Visual Representations," June 2020. arXiv:2002.05709 [cs, stat].

[6] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *CoRR*, vol. abs/1905.11946, 2019.

[7] L. Jing, P. Vincent, Y. LeCun, and Y. Tian, "Understanding dimensional collapse in contrastive self-supervised learning," *arXiv preprint arXiv:2110.09348*, 2021.

[8] M. Caron, I. Misra, J. Mairal, P. Goyal, P. Bojanowski, and A. Joulin, "Unsupervised learning of visual features by contrasting cluster assignments," 2020.