Machine Learning I Final Project Report

Yagna Kaasaragadda

April 2024

1 Abstract

As digital communication technologies advance, the complexity and volume of radio frequency (RF) signals increase, necessitating more effective compression and learning methods. This paper introduces an innovative approach that utilizes a hierarchical quantized autoencoder to compress and learn digitally modulated RF signals processed through deep learning models. This methodology addresses key challenges in wireless communications such as high-dimensional data management and real-time signal processing requirements.

Our research builds upon the foundations laid by previous studies on deep learning applications in wireless communications [4, 6]. Extending these concepts, we propose a novel compression framework employing hierarchical quantized autoencoders, drawing from advancements in quantization for deep learning models [1] and hierarchical structures [3]. This approach not only reduces data dimensionality efficiently but also learns a robust codebook that represents the complex characteristics of digitally modulated signals.

The hierarchical structure of our proposed autoencoder allows for multi-layered compression, where each layer captures different levels of abstraction of the signal data, thereby enhancing the compression efficiency. The quantization aspect of the autoencoder discretizes the continuous latent space into a finite set of codes, forming a codebook that can be efficiently used for signal reconstruction.

Experimental validation demonstrates that our approach not only maintains the fidelity of the reconstructed signals but also significantly reduces the computational overhead compared to traditional methods. Results indicate improved performance in terms of compression ratio and signal quality, confirming the potential of hierarchical quantized autoencoders in next-generation wireless communication systems.

2 Methodologies

The Dataset used in this research is a synthetic dataset for digital RF signal modulations called torchsig [2]. The accuracy of RF signal learning is evaluated based on an efficient Net b4 model transfer learnt on the signal modulations which can classify the six modulations of our choice [”4ask”,”8pam”,”16psk”,”32qam_cross”,”2fsk”,”ofdm-256”] with 100% accuracy after transfer learning for 15 epochs. The idea is that a reconstructed signal from the hierarchial model should be classified accurately by the efficient net model suggesting that the signal is infact reconstructed properly. The accuracy of the classification on the reconstructed dataset is the accuracy of reconstruction of the HQARF. More on this in 3. The synthetic dataset that is used for the training has 1024 IQ samples per signal and 8000 samples per class with a total of 24000 samples per training set and 4800 samples for the testing dataset.

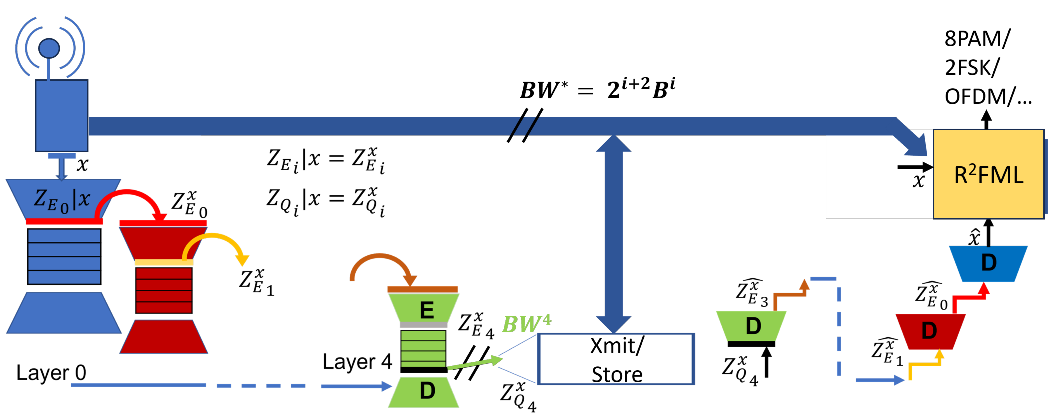


Figure 1: HQARF Architechure from [5]

The hierarchical quantized autoencoder architecture in Figure 1 consists of three main components: an encoder, a quantizer, and a decoder sandwiched as shown in the Figure 1. The encoder takes the input data x and transforms it into a series of latent vectors Z_{E_1}, Z_{E_2}, ..., Z_{E_d}, where d is the number of layers in the encoder. Each Z_{E_i} represents the latent vector at layer i. The quantizer takes the latent vectors Z_{E_1}, Z_{E_2}, ..., Z_{E_d} produced by the encoder and quantizes them to obtain quantized vectors Z_{Q_1}, Z_{Q_2}, ..., Z_{Q_d}. The quantized vectors are essentially discrete representations of the latent vectors.The decoder takes the quantized vectors Z_{Q_1}, Z_{Q_2}, ..., Z_{Q_d} and reconstructs the original input data. Each quantized vector Z_{Q_i} is passed through a corresponding decoder D_i, which produces a reconstructed signal $\hat{Z}_{Ei-1}$ for layer i. If $i = 0$, the reconstructed signal $\hat{x}$ is produced.

There are many issues that I ran into to improve the HAE classification accuracy. The first step is to find the best AutoEncoder architecture so that the representation of the signals in the Z_e latent space is as close as possible to the original signal. To achieve this, the loss function at any layer is given by Equation 1. where $x, \hat{x}$ are the original signal and the reconstructed signal from the auto encoder and Cf is the coefficient of the Cosine Loss. The addition of the Cosine loss to the loss function is due to the fact of the nature of digital modulation signals that any modulation can be represented as combination of Sine and Cosine waves namely IQ signals. From the observations we can also state that the cosine loss is significantly helping the accuracy of the auto encoder in the initial layers. Gradually as we go deep into the layers the cosine loss is hurting the accuracy which is demonstrated in Figure 2.

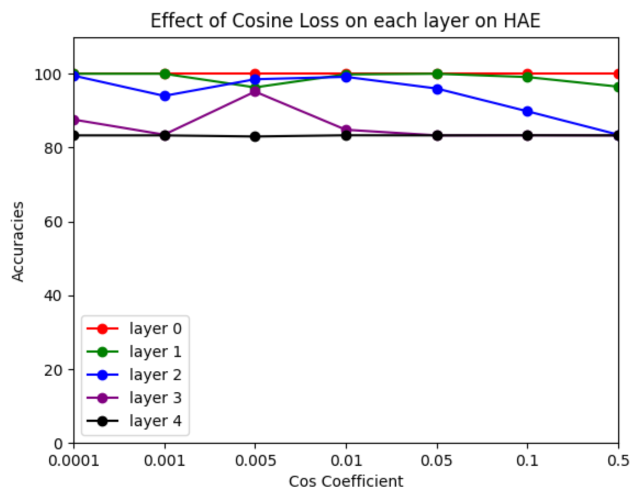


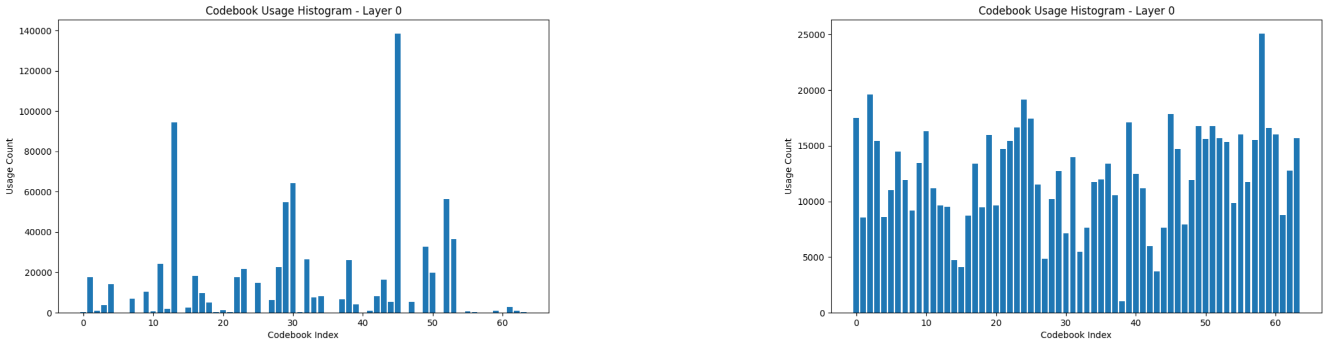
Figure 2: Cosine Loss effect on HAE classification accuracy for layers 0 to 4

$$L(\hat{x}|x) = \text{MSE}(x, \hat{x}) + Cf \times (1 - \cos(x, \hat{x})) \qquad (1)$$

By changing the cosine coefficient according to the layer we were successfully able to compress the signals and get good accuracy at different compression levels for the HAE. The next challenge is to significantly improve the classification accuracy of the quantized version of the Hierarchial auto encoder called HQARF [5]. For this we have used the same cosine losses and freeze the encoder and decoder part with the HAE encoder and decoder. This makes only the codebook as learnable parameters. The loss of function of the HQARF is shown in equation 2. The addition of KL Loss and the CL Loss is to make the distributions of the Z_E and Z_Q as close to each other as possible. Even then we find a difficulty in converging the model as there could be a chance that only first few codewords of the codebook are being used to represent the entire Z_Q space. This is inefficient learning of the codebook. To minimize this we need to reset the least used codewords frequetly by counting the usage of the codebook and reseting the least used codeword after few batches of training. This improves the effective learning of the codebook which is shown in the Figure 3. As we can see the clear usage of the codebook being completly used instead of only few codebook slots mostly being used.

$$L(\hat{x}|x) = \text{MSE}(x, \hat{x}) + Cf \times (1 - \cos(x, \hat{x})) + KL \times (p(z_q^{(i)} = j) \left(\log \left(\frac{p(z_q^{(i)} = j)}{p(z_e^{(i)} = j)} \right) + \log(K) \right)) + CL \times (p(z_q^{(i)} = j) \|\mathbf{z}_e^{(i)} - \mathbf{z}_q^{(i)}\|^2) \qquad (2)$$

Now we have acheived an effective method of learning a codebook for the representation of the signals. There are few important steps to note that the hyper parameters are randomly chosen and the accuracy could be improved significantly by just adjusting the KL loss coefficient and the CL loss coefficient and also the complexity of the model at different layers



(a) Initial 25 batches

(b) After complete training

Figure 3: Comparison of Codebook Usage with codebook reset

should also be experimented. The codebook initialization can also be a big part of the improvement of the model but as of now we only tried with Uniform and Normal distributions to initialize the codebook and found out Normal distribution is better suited for the case of learned representation.

3 Results & Conclusion

To find the learning of the model in representing the complex signals we chose classification accuracy of a pretrained efficient net b4 model on the reconstructed signal. The final accuracy of the HQARF and HAE are shown in Figure 4. This is not the best compression that we can get but with few changes in the model architechure like increasing the number of codebook slots and increasing the number of codebook dimensions and rerunning the entire HAE and HQA training we can acheive even better accuracy. The goal of this paper is to find if we were able to compress the signal and learn from the signal the underlying feature in the codebook and try to generate new signals from the codebook. To do that we have used the torchsig dataset and passed around 8000 samples of each modulation class and observed a histogram of which codebook slots are being used or triggered the most while reconstructing the signal as shown in Figure 6. We can see a clear seperated usage of codewords for clasees 8PAM, 2FSK, 4ASK but the classes 32QAM CORSS,OFDM-256 and 16PSK overlap with each other in few cases which is why the accuracy of those classes is less by looking at the confusion matrix 5. This is because 32QAM can represent 32 bits of information in a single signal and definitely need more codebook slots to represent its entire feature space.

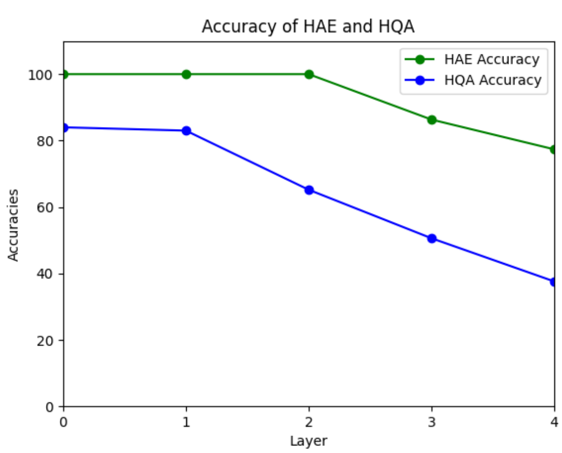


Figure 4: Classification accuracy of the HAE and HQARF

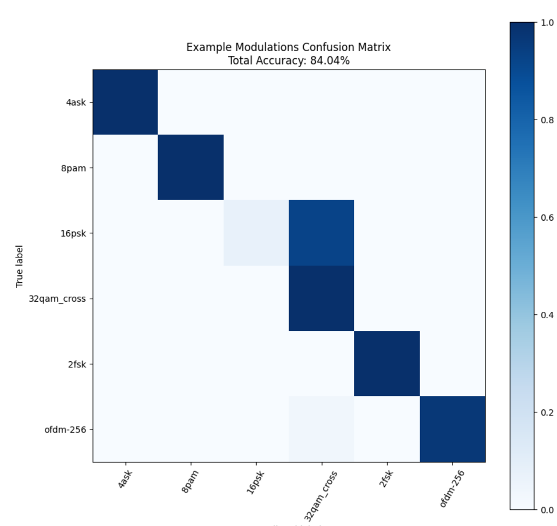
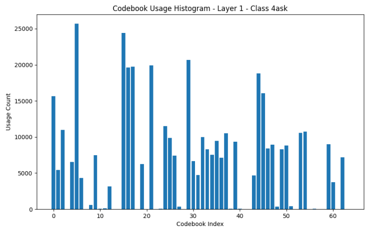
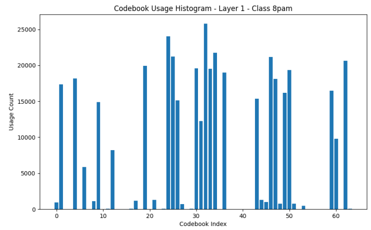


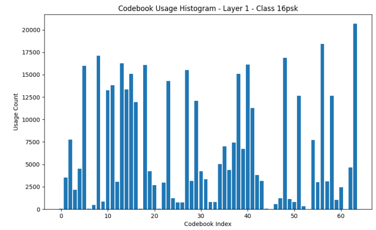
Figure 5: Confusion matrix of HQARF Layer 0 classifications



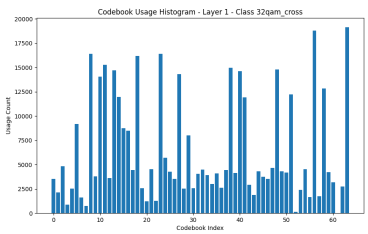
(a) 4ASK



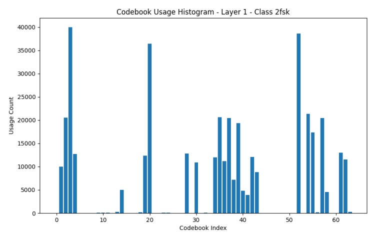
(b) 8PAM



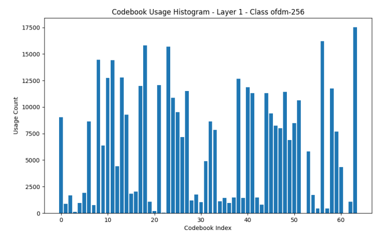
(c) 16PSK



(d) 32QAM CROSS



(e) 2FSK



(f) OFDM-256

Figure 6: Codebook usage Histograms for different modulations

Upon reconstructing each signal from different class for layer 0 and layer 4 for both HAE and HQARF we can see the constellation diagrams as shown in 7, which shows a lot more distortion for the signals reconstructed with the HQARF for the last layer which explains the low classification accuracy between different classes.

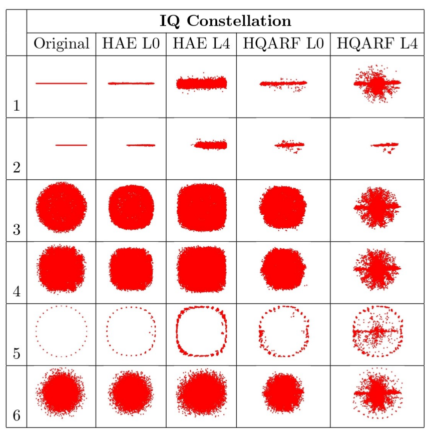


Figure 7: Constellations of Classes

From this paper we conclude that it is possible to learn and represent a complex digital Radio frequency signals in a very short compressed dimensions. This new representation is also useful to generate new signals from the codebooks by passing a random signal through the first layer of the Encoder, Quantizer and Decoder. There are a lot of things that should be further analysed to increase the classification accuracy of each layer at different compression as it is found to get even higher accuracies at different compression rates from the paper [5]. The main focus to implementing a better HQARF and HQA is by changing the architechure by reducing or increasing the number of resnet layers in encoder and decoder block. The initialization of codebook using the means of each class from the train dataset. Also so far the classification is good enough to classify at least few of the RF signals which can indicate that the RF signals can be compressed and reconstructed using HQARF architechure. These signal compression and learning the representations in codebook can be used to even learn the signals from the brain (EEG signals) which can be then represented for better classification of whats going with the signals inside the brain for classification.

References

[1] Eirikur Agustsson, Michael Tschannen, Fabian Mentzer, Radu Timofte, and Luc Van Gool. Soft-to-hard vector quantization for end-to-end learning compressible representations. *Advances in Neural Information Processing Systems*, 30:1141–1151, 2017.

[2] Luke Boegner, Manbir Gulati, Garrett Vanhoy, Phillip Vallance, Bradley Comar, Silvija Kokalj-Filipovic, Craig Lennon, and Robert D. Miller. Large scale radio frequency signal classification, 2022. arXiv:2207.09918.

[3] Jörn-Henrik Jacobsen, Jan van Gemert, Zhao Lou, and Arnold W. M. Smeulders. Revnet: A hierarchical deep learning model for large-scale visual recognition. *Pattern Recognition*, 77:426–437, 2018.

[4] A. M. Khaleghi, K. K. Wong, and F. Chin. Recent advances in the application of deep learning techniques in wireless communications. *IEEE Wireless Communications*, 24(2):145–151, 2017.

[5] Armani Rodriguez, Yagna Kaasaragadda, and Silvija Kokalj-Filipovic. Deep-learned compression for radio-frequency signal classification, 2024.

[6] M. Sun, Z. Huang, Z. Gao, S. Wang, and K. Niu. Deep learning-based modulation recognition for mimo-ofdm systems. *IEEE Access*, 8:5094–5104, 2020.