# Efficient Deep Learning Framework for MIMO Detection Using Transformers

Saleem Ahmed, Sooyoung Kim*

Electronic Eng. Dept., Jeonbuk National University, Jeonju, Korea

saleem.ahmed@jbnu.ac.kr, *sookim @jbnu.ac.kr

# 트랜스포머를 이용한 효율적인 딥러닝 MIMO 검출 프레임워크

아흐메드 살림, 김수영*

전북대학교

## Abstract

Signal detection in multiple-input multiple-output (MIMO) systems can employ deep learning (DL) to develop efficient receiver for future communication systems. In this paper, we evaluate the performance of a transformer-based MIMO detection method by using the encoder block of the transformer architecture. The proposed method performs a classification task based on the real and imaginary parts of the transmitted symbols, with each part representing bit-level information due to the use of the quadrature phase shift keying (QPSK) modulation scheme. We demonstrate that the choice of training signal-to-noise (SNR) value plays an important role in the overall performance and generalization of the model. Simulation results show that the proposed method outperforms the conventional detection methods.

## Ⅰ. Introduction

Multiple-input multiple-output (MIMO) systems employing the maximum likelihood (ML) detection method achieve optimal performance; however, ML detection method is computationally inefficient [1]. Alternatively, suboptimal detection algorithms are often employed to achieve a better balance between detection accuracy and computational complexity. Linear detectors, such as zero forcing (ZF) and minimum mean-squared error (MMSE), are computationally efficient but tend to exhibit poor detection performance.

Recently, deep learning (DL) has been applied to the MIMO detection problem. The DL-based detection methods have the potential to outperform traditional detection algorithms in certain complex environments.

In general, existing DL-based detectors can be categorized into two types. The first type, known as data-driven DL detectors which are based on deep neural networks (DNNs), while second type consists of model-driven DL detection methods [2]. Data-driven DL detectors utilize data to learn the detection problem through NN architectures. In contrast, model-driven DL detectors are built upon traditional iterative detection algorithms by unfolding each iteration into a network layer with learnable parameters. One example of a model-driven DL method with promising results is the orthogonal approximate message passing network-2 (OAMP-Net2) [3]. However, OAMPT-Net2 is computationally inefficient due to the need for an inverse operation in each iteration.

## Ⅱ. Transformer Learning based MIMO Detection

Let us consider a MIMO system with $M$ transmit and $N$ receiving antennas. The information bits are modulated through quadrature phase shift keying (QPSK) modulated scheme. These modulated symbols are then transmitted through a MIMO channel, and the received signal is represented as:

$$\mathbf{y} = \mathbf{H}\mathbf{s} + \mathbf{n}\,, \tag{1}$$

where $\mathbf{y}$ represent the received signal, $\mathbf{s}$ is the modulated symbol, $\mathbf{H}$ is the channel matrix, and $\mathbf{n}$ is the additive white Gaussian noise.

To facilitate DL processing, we convert the complex-valued model into an equivalent real-valued domain, as given by:

$$\bar{\mathbf{y}} = \bar{\mathbf{H}}\bar{\mathbf{s}} + \bar{\mathbf{n}}\,, \tag{2}$$

where $\bar{\mathbf{y}} \in \mathbb{R}^{2N}, \bar{\mathbf{H}} \in \mathbb{R}^{2N \times 2M}, \bar{\mathbf{n}} \in \mathbb{R}^{2N}$ and $\bar{\mathbf{s}} \in \mathbb{R}^{2M}$ .

Table I. Parameters for transformer learning-based MIMO detection

| Parameter | value |
|---|---|
| Encoder blocks | 8 |
| Attention heads | 8 |
| Sequence length | $2N$ |
| Embedding dimension | 128 |
| Learning rate | 0.0001 |
| Batch size | 5000 |
| Epochs | 150 |
| Training SNR | 20 dB, 25 dB |

The proposed transformer-based MIMO detection receiver has parameters listed in table I. The main advantage of the proposed method is that it only uses encoder block of the transformer architecture. The received signal is modified by applying the QR decomposition. After applying QR decomposition the received signal can be transformed as $\mathbf{y}' = \bar{\mathbf{R}}\bar{\mathbf{s}} + \mathbf{n}'$ The input to the network are the received signal and the channel matrix as, $[\mathbf{y}', \bar{\mathbf{R}}]$. Initially, the input is passed through an embedding layer, followed by the positional embedding. The resulting sequence is then fed into a transformer encoder blocks comprising multi-head attention, add & norm, and feed-forward sub-layers. The total of eight encoder blocks processes the input sequentially, and the final output is used for classification. The model performs a classification task to detect the transmitted bits by representing real and imaginary parts of transmitted symbol as target outputs, where each part corresponds to bit-level information due to QPSK modulation. The loss function used is binary cross entropy with logits, which combines a sigmoid activation with binary cross-entropy. Training is performed using the Adam optimizer with a learning rate of 0.0001, and a batch size of 5000 is used to improve gradient estimation and overall training stability.

## Ⅲ. Simulation results

Fig. 1 shows the training and test loss convergence behavior of the proposed method for a 4 × 4 MIMO system at two different SNR values of 20 dB and 25 dB. At 20 dB, the training and test loss stabilize after 130 epochs. The gap between training and test loss indicates that the model does not reliably approximate the target outputs. In contrast, for the 25 dB case, the model stabilizes after approximately 80 epochs.

Fig. 2 shows the bit error rate (BER) performance for the 4 × 4 MIMO system. The proposed method trained at 25 dB achieves a BER performance approaching the ML method. The conventional methods like OAMP-Net2 with total layers $T$, perform better than the MMSE detection method but still show a larger performance gap compared to the proposed method. The proposed method trained at an SNR of 20 dB outperforms both MMSE and OAMP-Net2 methods, but its performance is lower than the method trained at an SNR of 25 dB.

## Ⅲ. Conclusion

We propose a transformer-learning based MIMO detection method. The proposed method is trained on various SNR values and tested across range of SNR values. It outperforms conventional methods and demonstrates capabilities for both generalizability and scalability.

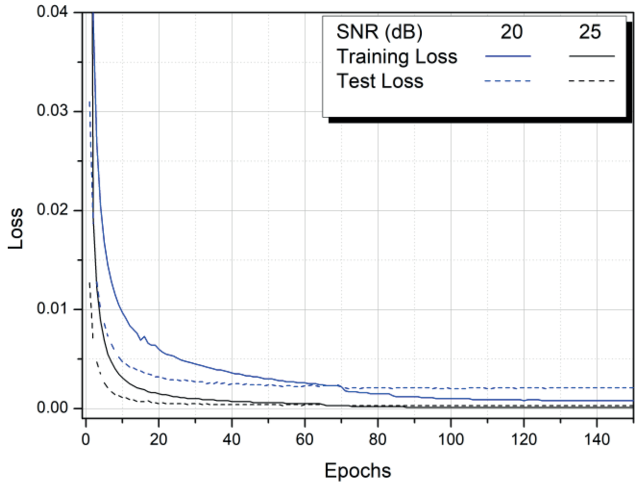


Fig. 1. Convergence behavior comparison for a 4 × 4 MIMO system with QPSK modulation.

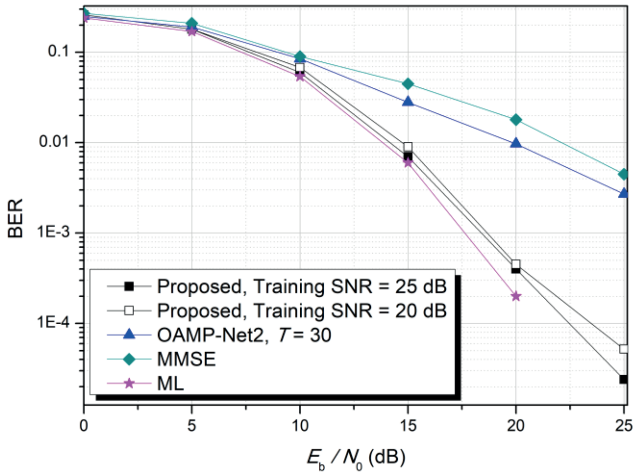


Fig. 2. BER performance comparison for a 4 × 4 MIMO system with QPSK modulation.


## ACKNOWLEDGMENT

This work was supported by National Research Foundation of Korea with Brain Pool (BP) program, Grant No. RS-2023-00217283.


## REFERENCES


[1] Zheng, K. Zhao, L. Mei, J. Shao, B. Xiang, W. and Hanzo, L. "Survey of large-scale MIMO systems," IEEE Commun. Surv. Tutorials, vol. 17, no. 3, pp. 1738- 1760, 2015.

[2] Burera, Ahmed, S. and Kim, S. "Transformer learning-based efficient MIMO detection method" Physical Comm., vol. 70, 2025.

[3] He, H. "Model-driven learning for MIMO detection," IEEE Trans. Of Dig. Process. vol. 68, pp.1702-1715, 2020.