

Semantic 통신 시스템에서 Transformer 에 기반한 Equalizer 에 관한 연구

최동하, 채찬병
연세대학교 글로벌융합공학과

{elijah1030, cbchae}@yonsei.ac.kr

A Study on the Transformer-based Equalizer in Semantic Communication Systems

Dongha Choi, Chan-Byoung Chae
Yonsei Univ.

요약

본 논문은 인공 신경망에 기반한 통신 기법인 semantic communication 시스템에서 사용될 수 있는 새로운 형태의 equalizer 를 제안한다. 기존 Rayleigh fading channel 에서 사용되었던 zero forcing 에 기반한 equalizer 를 벗어나, transformer 를 사용한 새로운 형태의 equalizer 를 설계하고, 이를 구현하여 비교하고 분석하였다.

I. 서론

최근 5G 와 6G 가 목표로 하는 ultra reliable low-latency communications (URLLC) 기술을 완성하기 위해 (6G 에서는 X-URLLC) 수많은 시도가 지속돼 왔으며, 인공신경망을 기반으로 source coding, channel coding, modulation 을 하나의 블록으로 처리하는 semantic communications 또한 그중 하나이다. Deep JSCC [1] 에서 출발한 이미지 도메인의 semantic communication 은 채널 환경에 adaptive 한 디코딩이 가능하다는 장점을 기반으로, 비트 에러에 따라 아예 디코딩이 불가능할 수 있어 재전송이 필요했던 기존 방식에 비해 더 낮은 지연시간을 구현할 수 있다.

본 논문에서는 Rayleigh fading channel 에서 semantic 통신을 할 때 사용할 수 있는 transformer 에 기반한 equalizer 를 제안한다. 또한 기존의 zero forcing 에 기반한 equalizer 와 제안하는 equalizer 를 각각 적용한 시스템을 설계하고 두 시스템의 성능을 비교한다.

II. 본론

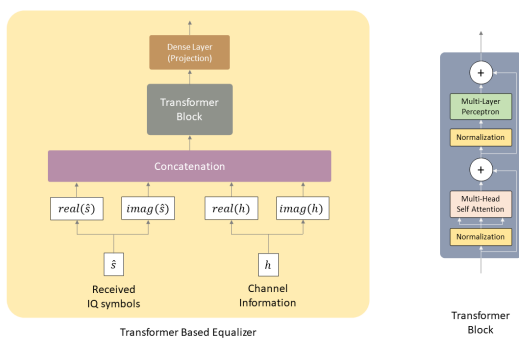


그림 1. Equalizer 와 transformer block 아키텍처

그림 1 의 좌측은 본 논문에서 제안하는 transformer 기반 equalizer 의 아키텍처이다. 먼저 리시버에서 수신한 심볼과 채널 정보를 모두 실수부와 허수부로 분리하였다. 이후 나누어진 데이터를 한 번에 concatenate 해준 후 transformer block 에 통과시켜 주었으며, 마지막으로 zero forcing 기반 모델과 연산 차원을 맞추기 위해 projection layer 를 추가해 주었다. 그림 1 의 우측은 사용된 transformer block 의 아키텍처이다.

Equalizer 부분을 제외한 인코더와 디코더에 사용된 모델은 각각 8 개의 transformer block 과 patch-merging, patch-division 등의 레이어로 구성된 visual transformer 에 기반한 모델을 사용하였으며, 신경망의 총 파라미터 개수는 13.5M 이다. 또 이미지 별로 사용된 심볼의 개수는 512 개로, channel bandwidth ratio (CBR)은 1/6 이다. 신경망 모델의 학습은 50,000 장의 훈련용 이미지로 구성된 32x32 크기의 CIFAR-10 데이터셋으로 진행되었으며, 최적화를 위해 LR=1e-4 의 Adam optimizer 와 mean squared error (MSE) Loss 를 사용하였다. 훈련 과정에서 채널은 SNR = 10dB 의 AWGN 채널과 Rayleigh 채널을 사용하였다.

그림 2 는 각각 AWGN 채널에서 학습된 모델, zero forcing 기반 equalizer 를 사용하여 Rayleigh 채널에서 학습된 모델, transformer 기반 equalizer 를 사용하여 Rayleigh 채널에서 학습된 모델을 사용하여, 10,000 장의 검증용 데이터셋을 Rayleigh 채널에서 전송했을 때의 성능을 비교한 것이다. 제안하는 시스템을 적용한 세 번째 모델이 전체적으로 낮은 SNR 환경에서 가장 좋은 것을 확인할 수 있다. 극단적으로 높거나 낮은 SNR 구간에서는 다른 모델들의 성능이 상대적으로 좋으나, 학습된 10dB SNR 을 기준으로 일반적으로 많이 사용되는 SNR 구간에서는 transformer 를 활용한 모델이 더 좋은 성능을 나타내는 것을 확인할 수 있다. 그림 3 은 실제로 이미지를 전송하여 비교한 것인데, 제안하는 모델이 원본 이미지를 왜곡 없이 가장 잘 전송할 수 있음을 보여준다.

참고 문헌

- [1] E. Boursoulatze, D. B. Kurka, and D. Gündüz, "Deep joint source-channel coding for wireless image transmission," *IEEE Transactions on Cognitive Communications and Networking*, vol. 5, no. 3, pp. 567–579, 2019.
- [2] A. Dosovitskiy et al., "An image is worth 16x16 words: Transformers for image recognition at scale," *arXiv preprint arXiv:2010.11929*, 2020.
- [3] H. Wu, Y. Shao, C. Bian, K. Mikołajczyk, and D. Gündüz, "Vision transformer for adaptive image transmission over MIMO channels," in *ICC 2023-IEEE International Conference on Communications, 2023*, pp. 3702–3707.
- [4] A. Vaswani et al., "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

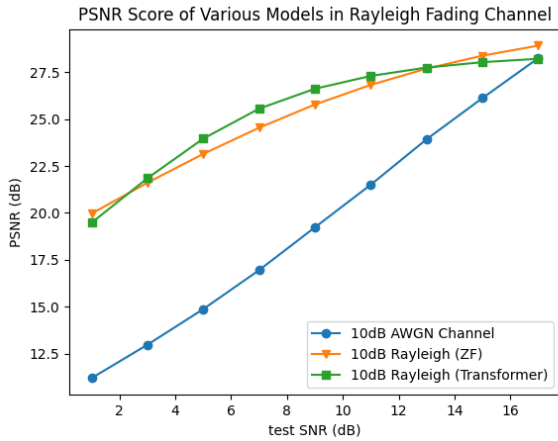


그림 2. 검증용 데이터셋에서의 PSNR 비교

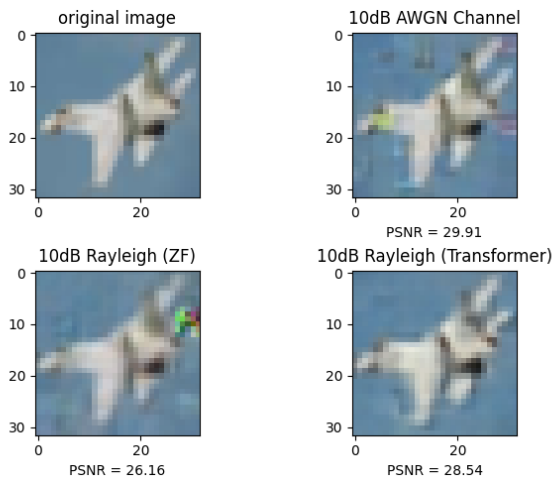


그림 3. 여러 모델을 사용하여 이미지를 전송한 결과

III. 결론

본 논문에서는 semantic communication 시스템에서 Rayleigh fading channel 을 고려하였을 때 사용될 수 있는 인공지능 신경망 형태의 equalizer 를 제안하고, 기존의 zero forcing 을 사용한 것과 비교하고 분석하여 보았다. 이때 제안한 시스템이 전체적으로 좋은 성능을 보였으며, 위와 같은 방법이 추후 더 실용적이고 범용적인 통신 상황에서도 사용될 수 있기를 기대한다.

ACKNOWLEDGMENT

이 논문은 2023 년도 두뇌한국 21 사업(4 단계 BK21 사업)에 의하여 지원되었음. (School of Integrated Technology / BK21 Graduate Program in IST)