

UAV 영상 전송을 위한 U-Net 기반 의미론적 통신

이예림, 신요안*

송실대학교 전자정보공학부

yllee4806@soongsil.ac.kr, *yashin@ssu.ac.kr

(*교신저자)

U-Net-Based Semantic Communication for UAV Video Transmission

Yelim Lee and Yoan Shin*

School of Electronic Engineering, Soongsil University

(*Corresponding author)

요약

본 논문은 U-Net 딥러닝 아키텍처를 기반으로 한 의미론적 통신 방식을 UAV 영상에 적용하였다. 전이 학습을 통해 향상된 특징 추출과 영상 복원 과정을 통해, 실시간으로 고해상도 영상을 효율적으로 처리하고 전송하는 방법을 제안하였다. 제안한 구조는 다양한 채널 잡음 환경에서 여러 성능 지표에 대해 기존 통신 구조보다 높은 성능을 보여준다.

I. 서론

UAV (Unmanned Aerial Vehicle) 기술은 재난 관리 등 인간이 접근하기 어려운 지역을 관찰하고 동시에 대량의 영상 데이터를 수집할 수 있다. 이 데이터는 실시간으로 처리 및 전송되어야 하며 고해상도의 영상은 높은 데이터 전송을 요구한다. 그러나 모든 영상 데이터가 동일한 중요도를 가지는 것은 아니며, 특히 긴급 상황에서는 중요한 의미를 담은 정보를 식별하고 전송하는 것이 중요하다.

본 논문에서는 의미론적 통신 (Semantic Communication)의 개념을 적용하여 영상 데이터의 중요한 의미 정보를 전송하는 방법을 제안한다 [1]. 이때, 딥러닝 기반의 전이 학습 (Transfer Learning) [2]을 활용하고 U-Net [3] 아키텍처를 기반으로 한 의미론적 통신 구조로 구축하였다. 이를 통해 UAV 영상 데이터의 효율적인 처리와 전송 방법을 제안하고 PSNR (Peak Signal-to-Noise Ratio) 및 SSIM (Structural Similarity Index) 지표를 통해 제안 구조의 성능을 분석한다.

II. U-Net 아키텍처 기반 의미론적 통신 구조

본 논문에서 제안한 구조는 기존의 U-Net 아키텍처를 기반으로, 복잡한 영상 데이터를 효과적으로 처리할 수 있는 다중 인코딩 및 디코딩 구조를 참고하여 설계하였다. 그림 1과 같이, 송신기 측에서는 각 인코더에서 영상 데이터에서 중요한 특징들을 추출하고 이를 임시 저장함과 동시에 영상 데이터 압축을 진행한다. 이후 임시 저장된 중요한 특징 추출 정보와 압축된 데이터는 물리 채널을 통해 수신자에게 전달된다. 수신기의 구조는 여러 디코더로 구성되어 있으며, 각 디코더가 전송된 데이터를 바탕으로 복원 과정을 수행하며 이 과정에서 각 단계의 정보를 종합하여 최종 영상을 정밀하게 재구성한다. 신호가 전송되는 물리 채널로는 AWGN (Additive White Gaussian Noise) 채널을 고려하고, OFDM (Orthogonal Frequency Division Multiplexing) 변조 방식을 활용하여 데이터의 전송 도중 발생할 수 있는 잡음과 간섭에 강한 통신을 가능하게 하였다.

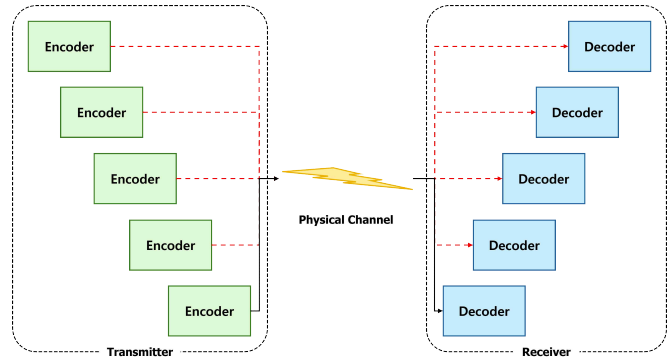


그림 1. 제안된 의미론적 통신 구조

III. 제안한 의미론적 통신의 세부 구조

그림 2는 전송 과정 중 데이터 처리와 전송의 세부적인 단계를 설명한다. 이 과정에서 송신기는 CNN (Convolutional Neural Network) 딥러닝 모델을 활용해 영상의 중요 특징들을 추출한 후, Flattening, Modulation, Normalization의 과정을 거쳐 물리 채널을 통해 수신기로 전송한다. 이때 사용된 CNN은 ResNet152V2 기반의 전이 학습으로 효율적인 학습 단축과 특징 추출이 가능하다 [4]. 수신기에서는 수신 데이터를 Demodulation과 Reconstruction 과정을 거쳐 최종적으로 영상을 복원한다. 여기서 Flattening은 앞서 언급한 CNN에서 추출된 특징들을 일련의 정보로 변환하여 모듈화와 정규화 과정에 적합한 형태로 조정한다. Modulation 과정은 이 정보들을 효과적으로 채널에 적합한 신호로 변환하는 작업이며, Normalization은 전송 신호의 전력을 균일하게 조정하여 채널을 통한 전송 효율을 최적화한다. 이 구조는 인공 신경망을 통해 데이터 압축과 채널 특성을 동시에 학습하는 DeepJSCC (Deep Joint Source-Channel Coding) 원리를 이용하였다 [5]. Demodulation은 수신 신호에서 원래 데이터 신호를 추출하는 과정이며, Reconstruction은 전송된 데이터를 바탕으로 CNN 기반 디코더에서 영상 복원이 가능하게 연산할 수 있도록 재구성한다.

이러한 전체 과정은 영상 데이터의 중요한 의미를 유지하면서 효율적으로 전송할 수 있게 한다.

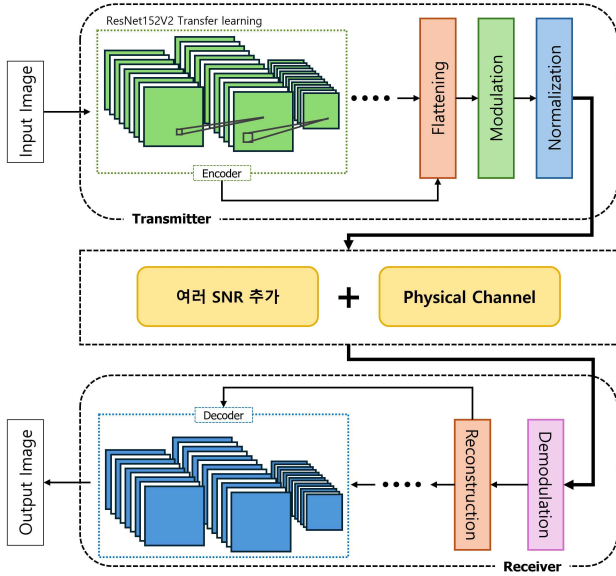


그림 2. 제안된 의미론적 통신의 세부 구조

III. 실험 결과 및 결론

본 연구에서는 UAV의 영상 데이터 셋으로 Tianjin 대학교 기계학습 및 데이터 마이닝 연구소의 AISKYEYE 팀이 수집한 VisDrone2019 데이터 셋을 사용하였고, 낮과 밤, 날씨 등 여러 환경을 담은 영상들로 훈련하여 다양한 데이터를 추출할 수 있도록 실험을 진행하였다 [6].

그림 3에서 볼 수 있듯이 영상 전송 결과에 차이를 시각적으로 직접 확인할 수 있다. 기존 통신 구조에 비해 제안 구조는 배경의 세부 정보와 객체의 경계가 더욱 명확하게 보존되었으며, 원본 영상과 매우 흡사하게 출력되었다. 특히 야간 영상에서는 제안 구조가 불빛 및 반사된 효과가 지도 원본에 가깝게 복원하였다. 또한, 정량적 성능 검증을 위해 PSNR 및 SSIM 지표를 사용하여 원본 영상과 전송된 영상을 비교 분석하였다. 그 결과, 그림 4의 그래프처럼 제안 구조는 SNR이 증가함에 따라 (a) PSNR과 (b) SSIM 모두 기존 통신 구조를 개선하는 것으로 나타났다. 이러한 결과들은 제안된 의미론적 통신 구조가 UAV 영상 전송에 있어서 높은 잡음 환경에서도 효율적으로 중요한 정보를 전송하고 복원할 수 있음을 보여준다. 이처럼 재난 현장 등 실시간 데이터 및 정확한 전송이 요구되는 환경과 의미를 중심으로 전송하는 것을 추구하는 차세대 기술의 통신 과정에도 기여할 수 있을 것이다.

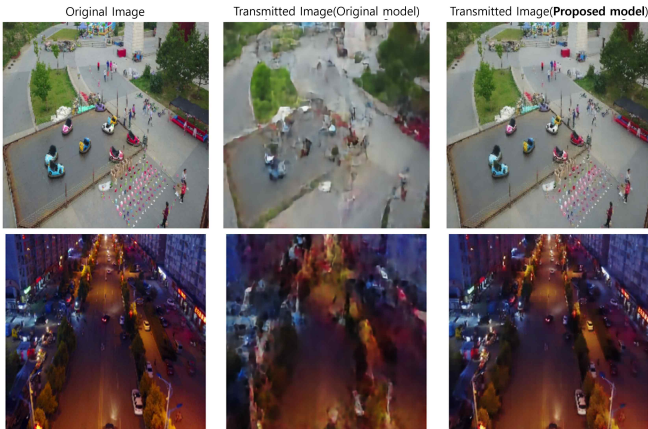
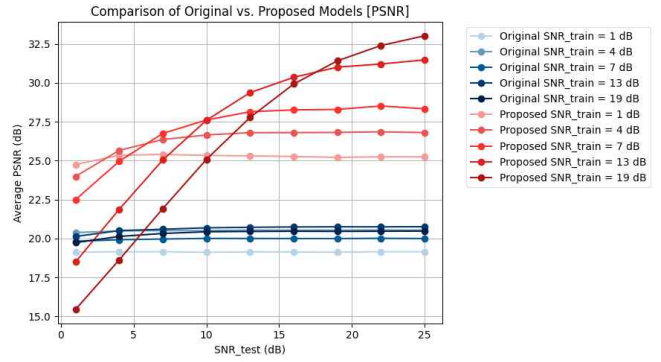
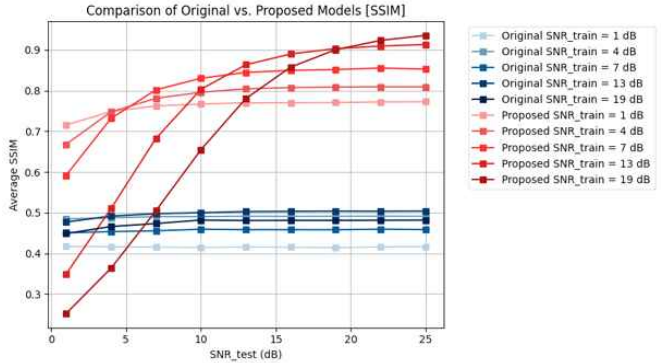


그림 3. 기존 및 제안된 의미론적 통신 구조의 영상 복원 결과



(a) PSNR



(b) SSIM

그림 4. 기존 및 제안된 의미론적 통신 구조의 성능 지표 결과

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학ICT연구센터사업의 연구결과로 수행되었음 (IITP-2024-RS-2023-00258639)

참고 문헌

- [1] L. Chuanhong, G. Caili, Y. Yang, and J. Nan, "Adaptable semantic compression and resource allocation for task-oriented communications," *arXiv preprint arXiv:2204.08910*, Apr. 2022.
- [2] S. J. Pan, and Q. Yang, "A survey of transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345-1359, Oct. 2009.
- [3] O. Ranneberger, A. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," *Lecture Notes in Comp. Sci.*, vol. 9351, Issue Cvd, pp. 234-241, May 2015.
- [4] K. He, X. Zhang, S. Ren, and J. Sun, "Identity mappings in deep residual networks," *arXiv preprint arXiv:1603.05027*, July 2016.
- [5] J. Xu, T-Y. Tung, B. Ai, W. Chen, Y. Sun, and D. Gunduz, "Deep joint source-channel coding for semantic communications," *arXiv preprint arXiv:2211.08747*, July 2023.
- [6] <https://github.com/VisDrone/VisDrone-Dataset>