

시맨틱 통신에서의 영상 복원 성능평가

이영묵[§], 방인규[†], 김태훈[§]

[§]국립한밭대학교 컴퓨터공학과, [†]국립한밭대학교 지능미디어공학과

lee_ym@edu.hanbat.ac.kr, {ikbang, thkim}@hanbat.ac.kr

Performance Evaluation of Image Reconstruction in Semantic Communications

Yeongmuk Lee[§], Inkyu Bang[†], Taehoon Kim[§]

[§]Department of Computer Engineering, Hanbat National University

[†]Department of Intelligence Media Engineering, Hanbat National University

요약

최근 시맨틱 통신이 차세대 통신 패러다임으로 주목받는 가운데, 본 연구에서는 세그멘테이션 마스크 기반의 시맨틱 복원 시스템을 구현하고, GAN과 Diffusion 기반 모델의 복원 성능을 비교하였다. 동일한 마스크 입력에 대해 두 모델을 학습하고 복원된 이미지를 정성적으로 평가한 결과, Diffusion 모델이 시각적 자연스러움과 의미 보존 측면에서 더 우수한 성능을 보였다. 본 연구는 시맨틱 통신 환경에서 단순한 정량 지표를 넘어, 의미 기반 평가의 필요성을 제시한다.

I. 서론

기존의 이미지 통신 시스템은 입력 이미지를 압축하여 전송하고, 수신 측에서 이를 복원하는 구조를 기반으로 한다. 이러한 방식은 Shannon의 채널 용량 이론에 따라 주어진 대역폭과 전력 제약 하에서 전송 효율 극대화를 목표로 해왔으며, 비트 오류율(BER), PSNR, SSIM 등의 지표를 통해 통신 품질을 측정해왔다. 그러나 이러한 접근은 데이터의 의미(semantic meaning)나 상대적 중요도를 고려하지 않기 때문에, 특히 위성과 같은 제한된 자원을 가진 환경에서는 의미 보존과 전송 효율을 동시에 달성하는 데 한계를 드러낸다.

최근 인공지능(AI) 및 딥러닝 기술의 발전에 힘입어, 단순한 비트 단위 복원보다는 의미 기반 정보 전달을 중시하는 시맨틱 통신(Semantic Communication)이 주목받고 있다. 시맨틱 통신은 전통적인 통신 방식에서 벗어나, 전송 대상 정보의 의미적 중요성에 따라 데이터를 선택적으로 압축, 전송, 복원하는 새로운 접근이다. 이는 특히 저궤도 위성, UAV, IoT 기기 처럼 전송 에너지 및 시간 자원이 제한된 시스템에서 효과적인 통신 방법으로 주목받고 있다. 이러한 시맨틱 통신은 다양한 방향으로 연구가 진행되고 있다. 위성 이미지 전송에서 불필요한 구름 영역을 제거하고 변화 감지 영역만 전송하는 구조를 통해 에너지 효율을 극대화하는 프레임워크가 제안되었으며[1], 객체의 중요도에 따라 비트 할당을 조절하고 시맨틱-시각적 손실을 동시에 최소화하는 압축 및 복원 방식 또한 연구되고 있다[2]. 더 나아가, 대규모 생성 모델을 활용하여 채널 노이즈가 심한 환경에서도 의미 있는 객체 복원이 가능한 시맨틱 통신 연구가 진행중이다[3].

본 연구는 세그멘테이션 마스크 기반 시맨틱 정보 전달 시스템을 구현하고, Pix2Pix와 Diffusion 기반 모델을 활용하여 의미 기반 복원이 실제로 사람의 인지 수준에서 타당한지를 정성적으로 평가함으로써 시맨틱 통신의 실제 효용성을 분석하고자 한다.

II. 시스템 모델

본 논문에서는 그림 1과 같이, 의미 기반의 시맨틱 통신 파이프라인을

구성하여 세그멘테이션 마스크로부터 RGB 이미지를 복원하는 시스템을 구현하였다. 전체 구조는 송신단에서 의미 추출 및 전송을 수행하고, 수신단에서 의미 정보를 기반으로 RGB 이미지를 재구성하는 구조로 이루어져 있으며, Pix2Pix 및 Diffusion 기반 두 가지 모델을 적용하여 복원하였다.

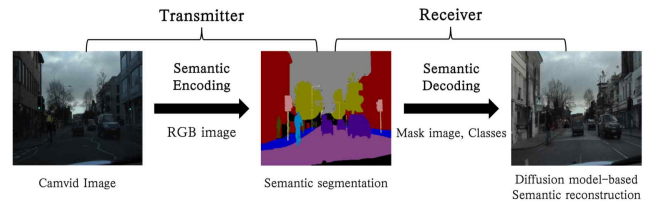


그림 1 시스템 모델

송신단에서는 Camvid 데이터셋의 원본 RGB 이미지로부터 시맨틱 세그멘테이션 마스크를 생성한다. 이 마스크는 각 픽셀이 12개 클래스 중 하나의 정수 값(0~11)으로 라벨링된 단일 채널의 정수형 이미지로, 장면 내 의미 정보를 압축적으로 표현한다.

수신단에서는 전송된 시맨틱 마스크를 조건 입력으로 활용하여 RGB 이미지를 복원한다. 이를 위해 두 가지 방식의 생성 모델을 구성하였다. 첫 번째는 Pix2Pix 모델로, 조건부 GAN 구조와 U-Net 기반의 Generator를 통해 마스크로부터 RGB 이미지를 생성하며, L1 손실과 GAN 손실을 함께 활용한다. 두 번째는 Stable Diffusion 기반의 모델로, U-Net, VAE, ControlNet으로 구성된 구조를 통해 조건 생성 과정을 수행하며, 마스크는 클래스별 색상으로 재매핑된 형태로 입력된다. 두 모델 모두 동일한 시맨틱 마스크를 입력으로 사용하지만, 생성 방식과 결과물의 특성은 다르게 나타난다.

본 시스템 모델은 복원된 이미지의 픽셀 단위 정합성보다 의미 기반의 시각적 타당성을 중시하며, 시맨틱 통신 관점에서 중요한 요소인 의미 보존과 자연스러운 복원의 가능성을 정성적으로 평가하는 데 중점을 둔다.




Original Image	Pix2Pix model Semantic reconstruction	Diffusion model-based Semantic reconstruction
		
	PSNR: 22.99 dB SSIM: 0.6472 LPIPS: 0.4088	PSNR: 20.40 dB SSIM: 0.5742 LPIPS: 0.3846

그림 2 Pix2Pix와 Diffusion 기반 시맨틱 복원 결과 비교

III. 성능평가

본 연구에서는 세그멘테이션 마스크를 기반으로 복원된 이미지의 품질을 평가하기 위해, 전통적인 정량적 지표뿐만 아니라 시맨틱 통신의 목적에 부합하는 정성적 기준을 중심으로 비교 분석을 수행하였다. 일반적인 이미지 복원에서는 PSNR(Peak Signal-to-Noise Ratio), SSIM(Structural Similarity Index), LPIPS(Learned Perceptual Image Patch Similarity)와 같은 수치 기반의 평가 지표가 널리 사용되지만, 시맨틱 통신의 목적은 원본 데이터를 수치적으로 완벽하게 재현하는 것이 아니라, 의미를 보존한 채 유용한 정보를 효과적으로 전달하는 것에 있다. 따라서 본 논문에서는 정량 지표는 보조적으로 활용되되, 사람이 인지적으로 수용 가능한 수준에서 의미와 구조가 얼마나 잘 보존되었는가에 초점을 맞춘 정성적 평가를 주된 분석 방법으로 채택하였다.

그림 2와 같이 두 생성 모델을 비교한 결과, Pix2Pix 모델은 복원 이미지의 품질 측면에서 전반적으로 Diffusion 모델에 비해 부족한 성능을 보였다. Pix2Pix는 U-Net 기반 구조를 활용해 세그멘테이션 마스크로부터 빠르게 RGB 이미지를 생성할 수 있는 장점이 있으나, 실제 복원 결과는 경계가 날카롭고 인위적인 느낌이 강하며, 객체의 색감이나 질감이 과도하게 단순화되는 경향을 보였다. 특히, 작은 객체의 묘사가 불완전하거나 형태가 왜곡되는 경우가 자주 관찰되었다. 반면, Diffusion 기반 모델은 복원된 이미지의 시각적 일관성과 자연스러움 측면에서 우수한 결과를 보였다. 본 모델은 다단계 denoising 과정을 통해 마스크로부터 점진적으로 이미지를 생성하며, 색상, 질감, 명암 등 다양한 요소를 자연스럽게 복원하였다. 특히 객체 간의 의미적 관계나 전반적인 구성에 있어 인간 시각 관점에서 수용 가능한 수준의 결과를 안정적으로 생성하는 것이 확인되었다. 정량적 지표를 참고하면, Pix2Pix 모델은 평균 PSNR 15.02dB, SSIM 0.4619, LPIPS 0.4671을 기록하였고, Diffusion 모델은 PSNR 12.68dB, SSIM 0.3866, LPIPS 0.4675를 기록하였다. PSNR과 SSIM에서 Pix2Pix가 다소 우세한 수치를 보였지만, 이는 픽셀 단위 복원만을 기준으로 하기 때문에 실제 인지 품질과는 차이가 있다. 특히 LPIPS 지표에서 두 모델 간 수치 차이가 거의 없거나 오히려 Diffusion 모델이 낮은 값을 보이는 경우도 있어, 인간의 시각적 유사성 관점에서는 Diffusion 기반 복원이 더 효과적임을 보여준다.

결과적으로, Pix2Pix는 상대적으로 단순한 구조와 생성 속도라는 장점에 불과하고, 의미 기반 복원 성능에서는 제한적이었으며, 이는 시맨틱 통신 시스템에서 의미 보존 및 전달 능력보다 단순한 복원 속도나 학습

효율만을 기준으로 복원 모델을 선택하는 것이 적절하지 않다는 점을 시사한다. 반면, Diffusion 기반 복원 방식은 생성 속도가 느리고 연산량이 많은 단점에도 불구하고, 의미 보존과 시각적 자연스러움을 보다 충실히 만족시켜 향후 시맨틱 통신 시스템 설계에서 중요한 고려 요소로 활용될 수 있다.

IV. 결론

본 연구에서는 시맨틱 통신 기반의 이미지 복원 시스템을 구현하고, Pix2Pix와 Diffusion 모델의 복원 성능을 정성적으로 비교하였다. 실험 결과, 두 모델은 동일한 시맨틱 마스크를 입력으로 사용했음에도 생성 방식의 차이로 인해 복원 품질과 의미 보존 측면에서 뚜렷한 차이를 보였으며, 특히 Diffusion 모델이 시각적 자연스러움과 의미 일관성에서 우수한 결과를 보였다. 이러한 결과는 시맨틱 통신에서 단순한 정량 지표만으로 복원 품질을 평가하는 것은 한계가 있으며, 의미 기반의 정성적 평가 기준이 병행되어야 함을 시사한다. 향후 연구에서는 시맨틱 복원 시스템에 적합한 새로운 평가 기준을 정의하고, 의미 보존과 인지적 타당성을 정량적으로 측정할 수 있는 지표를 설계할 것이다. 이를 통해 향후 의미 기반 정보 전달 시스템의 성능 비교 및 표준화에도 기여할 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

이 논문은 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원 SW중심대학사업 및 대학 ICT 연구센터(ITRC)의 지원을 받아 수행된 연구임 (2022-0-01068, 50%, IITP-2025-RS-2024-00437886, 50%)

참고 문헌

- [1] Bui, Van-Phuc, *et al.* "Semantic Image Encoding and Communication for Earth Observation with LEO Satellites." *IEEE Transactions on Cognitive Communications and Networking* (2024).
- [2] Huang, Danlan, *et al.* "Toward semantic communications: Deep learning-based image semantic coding." *IEEE Journal on Selected Areas in Communications* 41.1 (2022): 55-71.
- [3] Jiang, Peiwen, *et al.* "Semantic satellite communications based on generative foundation model." *IEEE Journal on Selected Areas in Communications* (2025).