

시맨틱 통신을 위한 확산 모델의 응용 기법

강정호*, 서효운†, 최완*

서울대학교 전기정보공학부*, 성균관대학교 전자전기컴퓨터공학과†

ttomas11@snu.ac.kr*, hyowoonseo@skku.edu†, wanchoi@snu.ac.kr*

Approaches to Applying Diffusion Models in Semantic Communication

Jeongho Kang*, Hyowoon Seo†, Wan Choi*

*Seoul Univ., †Sungkyunkwan Univ.

요약

본 논문은 확산 모델을 시맨틱 통신 시스템에 적용한 최신 연구들을 종합적으로 분석한다. 특히 통신 과정에서 발생하는 잡음 제거, 채널 상태 추정, 정보 복원 등 다양한 영역에서 확산 모델이 통신 시스템에 적용되는 방식을 중점적으로 다룬다. 또한, 확산 모델을 통신 시스템에 적용할 때 발생하는 문제점과 향후 연구 방향에 대해 고찰한다.

I. 서론

확산 모델은 점진적으로 노이즈를 주입하고, 다시 점차 노이즈를 제거하는 과정을 통해 정보를 복원하는 과정으로 학습을 진행한다. 이 과정을 통해 데이터의 분포를 잘 학습할 수 있고, 최종적으로는 강력한 생성능력을 가질 수 있게 된다. 확산 모델은 기존의 적대적 생성 신경망, 변분 오토인코더 등의 생성 모델들에 비해 데이터 분포를 안정적으로 학습하며 우수한 성능을 보이는 모델로, 다양한 응용분야에 활용되고 있다. 확산 모델 성능 개선에 대한 연구는 Denoising Diffusion Probabilistic Model(DDPM)[1]을 시작으로 하여 모델의 구조에 변화를 주거나 학습 및 샘플링 속도를 빠르게 하기 위한 기법에 관해 많이 진행되고 있다. DDPM[1]에서 제시한 잡음 제거를 통한 데이터 복원의 개념은, 잡음에 따라 수신신호의 품질이 크게 달라지는 통신분야에서도 큰 메리트를 가질 수 있다. 이에, 시맨틱 통신에서도 확산 모델을 적용하는 여러 방안이 제시되고 있으며, 본 논문에서는 확산 모델을 시맨틱 통신에 적용한 여러 기법에 대하여 소개한다.

II. 본론

시맨틱 통신 시스템은 정보의 의미를 추출하여 그 의미를 전송 및 수신하여 정보 전달을 하고자 하는 시스템으로 다양한 연구가 진행되고 있다. 대표적인 방향으로, 인간의 의사소통 방식을 모티브로 하여 맥락추론을 바탕으로 의미정보를 송수신하는 시스템에 대한 연구가 진행되고 있다[2], [3]. 또한, 기계학습의 발전에 따라, 학습된 인공신경망을 통해, 수행하고자 하는 과업에 연관성이 높은 정보의 특정 추출을 하는 것이 가능해졌다. 이를 통신분야에도 적용하여, 인공신경망을 일종의 부호기로 사용하게 되었을 때, 정보 복원 및 과업 수행에 알맞은 특정 정보를 추출할 수 있게 되었다. 특히 인공신경망을 학습시킬 때 채널에 대한 요소 또한 반영되게 학습시킬 수 있으며, 이는 정보와 채널을 동시에 고려하여 복호 및 부호화를 하는 딥러닝 기반 소스-채널 통합 부호화(Joint

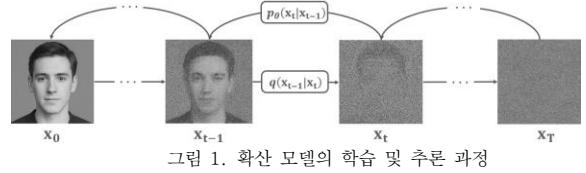


그림 1. 확산 모델의 학습 및 추론 과정

Source Channel Coding, JSCC)기법으로 생각할 수 있다. 기계학습의 발전에 따라 JSCC는 시맨틱 통신의 주요 연구분야이자 프레임워크로 자리잡았다. 대표적으로 이미지 전송의 경우, DeepJSCC[4] 프레임워크를 사용해 이미지를 전송하면, 낮은 SNR에서도 충분한 품질을 유지하며 이미지의 송수신이 가능했다. 이는 통신 자원이 제한되어 있거나 열악한 통신 환경 속에서 효율적, 효과적인 통신의 가능성이 제시된 것으로 볼 수 있다. 해당 프레임워크를 기본으로 하면서 확장성과 일반화를 고려해 시스템을 개선하려는 연구 또한 많이 이루어지고 있다[5]. 또 다른 방향으로, 통신 성능 향상을 위해 기존 JSCC 프레임워크에 확산 모델과 같은 생성형 AI 기법을 적용할 수 있다. [6]은 JSCC 복호기를 거쳐 복원된 정보를 다시 확산 모델의 입력으로 넣는다. 즉, 확산 모델을 사용한 추가적인 잡음 제거 과정을 거치는 것으로, 이를 통해 더 개선된 품질의 데이터를 복원할 수 있다. 또한 확산 모델을 채널 추정과정에 사용하는 연구 또한 존재한다. [7]에서는 채널 추정과정에서 추정된 채널 상태 정보에 대한 채널 보정 과정을 확산 모델을 통해 수행하는 것으로, 이는 더 정확한 채널 정보를 바탕으로 통신 성능을 향상 시킬 수 있다.

최근 연구되고 있는 DDPM[1] 등의 확산 모델은 잡음 제거 과정을 기반으로 학습과 추론을 수행한다. 이는 곧, 통신 과정에서 발생한 잡음을 확산 모델로 제거함으로써 성능을 효과적으로 향상시킬 수 있는 가능성을 시사한다. [8]에서 제시된 Channel Denoising Diffusion Model (CDDM)은 채널을 통과한 신호에 대해서 확산모델 기반으로 설계된 CDDM을 통해 복호화 전에 심볼 별 잡음 제거 과정을 수행한다. 덧붙이자면, CDDM은 채널 추정 과정을 통해 나온 채널 정보와, 이를 바탕으로 통신 과정을 거친 수신 신호를 입력으로 받는다. 이후 심볼 별 잡음 제거 과정을 거친 후 복호화를 수행한다. 이 경우 기

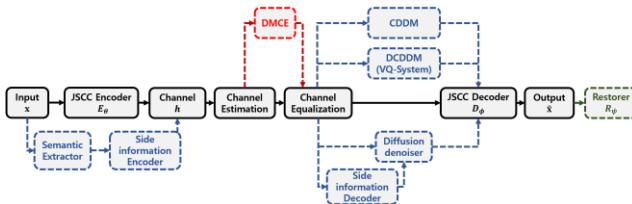


그림 2. 시맨틱 통신 시스템(실선)에 다양한 형태로 적용된 확산 모델(점선)

존 JSCC 프레임워크에 비해 PSNR 등에서 성능이 향상된 것을 확인할 수 있다. 또한 디지털 신호의 잡음 제거를 위한 연구 또한 존재한다. 벡터 양자화된 시맨틱 정보를 전송하는 디지털 시맨틱 통신 프레임워크에서는 전송된 이산 신호에 대해 검출과정을 거친다. 이후 이산 신호에 대한 확산 모델을 이용하여 검출된 벡터 신호의 오류를 정정할 수 있고, 채널 및 잡음에 강인한 디지털 시맨틱 통신 시스템을 구축할 수 있다[9]. 또한 통신환경이 열악한 상황에서도 확산 모델이 유용하게 사용될 수 있다. [10]에서는 대역폭 제한 상황에서 통신을 할 때, 시맨틱 통신과 확산 모델을 사용한다. 학습과정에서 다양한 대역폭 제한에 대해 적응적 학습을 한 결과, 대역폭 제한이 큰 경우에도 좋은 성능을 보일 수 있음을 시사하였다.

확산 모델 중 Latent Diffusion Model(LDM)[11]은 기존 확산 모델의 문제점을 크게 개선한 모델이다. 기존 확산 모델의 경우, 고화질의 이미지를 생성하는 데 있어 성능이 좋지 않았으며, 샘플링 시간이 너무 긴 문제점이 존재했다. LDM에서는 픽셀 공간에서의 확산과정을 수행하지 않고, 잠재 공간에서 정보를 표현하고, 그에 대한 확산 및 샘플링 과정을 거친다. 이 과정을 통해 확산 모델은 데이터의 분포를 잘 학습하며, 동시에 낮은 차원에서 확산 및 샘플링 과정을 진행하므로 학습 및 추론 시간이 줄어든다는 장점을 가진다. 또한 LDM은 샘플링 과정에 서 추가적인 조건부 정보를 결합할 수 있도록 설계되었는데, 이를 통해 사용자는 원하는 방향과 품질에 더 가깝게 데이터를 생성할 수 있게 되었다. 이러한 LDM 구조 또한 시맨틱 통신에 적용될 수 있다. JSCC 프레임워크의 시맨틱 통신 시스템 또한 원본 데이터를 전송하는 것이 아니라, 의미 정보를 내포하는 잠재 공간 상의 데이터로 표현하고 이를 전송하는 구조로, 의미 정보를 복원하는 과정에서 LDM의 잠재 공간 상 확산 및 샘플링 과정을 이용할 수 있다. [12]에서는 송신부에서 이미지와 같은 원본 정보로부터 엣지 맵, 텍스트 등의 의미론적 부가 정보를 추출한다. 이후 원본 정보에 대해서는 JSCC 를, 부가 정보에 대해서는 별도의 부호화를 거쳐 전송한다. 이후 복호화 된 부가 정보를 조건부 정보로써 활용하고 LDM 구조를 이용해 원본 정보를 복원한다. 즉, 확산 모델을 이용해 복원하는 과정에서, 부가 정보를 일종의 지침으로 사용한다고 볼 수 있다. 이 외에도 JSCC 프레임워크에 확산 모델을 통합하여 통신 성능을 개선하기 위한 연구가 꾸준히 진행되고 있다.

III. 결론

JSCC 프레임워크에 확산 모델을 통합하는 연구들을 통해, 시맨틱 통신에서 잡음으로 인해 발생할 수 있는 의미적 모호성, 정보 왜곡 등에 대해 강인해질 수 있음을 확인할 수 있었다. 이는 시맨틱 통신이 더 정확한 의미정보를 송수신 할 수 있는 방향성을 제시했다고 볼 수 있다. 다만 확산 모델의 성능이 뛰어난 반면, 해결해야 할 단점들이 존재한다. 그 중에서도 샘플링 속도가 느리다는 것은 실시간성 및 전송 효율이 중요한 통신에 치명적일 수 있다. 이를 해결하기 위해, 샘플링

과정을 축소하거나 모델 구조를 변형하는 등, 성능 개선을 위한 다양한 연구가 진행되고 있다. 본 논문에서 분석한, 통신 시스템에 적용된 확산 모델의 여러 응용 사례들을 바탕으로, 차세대 통신 시스템의 성능을 확산 모델을 통해 비약적으로 개선시킬 수 있을 것이라 기대한다.

ACKNOWLEDGMENT

본 연구는 한국연구재단을 통해 과학기술정보통신부의 「한-핀란드 공동연구사업」의 지원을 받아 수행되었음(RS-2024-00464570)

참 고 문 헌

- [1] Ho, et al. "Denoising diffusion probabilistic models." *NeurIPS*, (2020).
- [2] H. Seo, et al. "Semantics-Native Communication with Contextual Reasoning," *IEEE TCCN*, (2023).
- [3] H. Seo, et al. "Bayesian Inverse Contextual Reasoning for Heterogeneous Semantics-Native Communication," *IEEE TCOM*, (2024).
- [4] E. Bourtsoulatze, et al. "Deep joint source-channel coding for wireless image transmission." *IEEE TCCN*, (2019).
- [5] Y. Huh, et al. "Universal joint source-channel coding for modulation-agnostic semantic communication." *IEEE JSAC*, (2025).
- [6] S.F.Yilmaz, et al. "High perceptual quality wireless image delivery with denoising diffusion models." *IEEE INFOCOM*, (2024).
- [7] Zeng, Youcheng, et al. "DMCE: Diffusion model channel enhancer for multi-user semantic communication systems." *IEEE ICC*, (2024).
- [8] T.Wu, et al. "CDDM: Channel denoising diffusion models for wireless semantic communications." *IEEE TWC*, (2024).
- [9] G. He, et al. "Residual Vector Quantized Diffusion Model-Based Joint Source-Channel Coding for Task-Oriented Agent Communication." *IEEE TWC*, (2025).
- [10] Guo, Lei, et al. "Diffusion-driven semantic communication for generative models with bandwidth constraints." *IEEE TWC*, (2025).
- [11] R. Rombach, et al. "High-resolution image synthesis with latent diffusion models." *IEEE/CVF CVPR*, (2022).
- [12] M. Zhang, et al. "Semantics-Guided Diffusion for Deep Joint Source-Channel Coding in Wireless Image Transmission." *arXiv:2501.01138*, (2025).