

무선 이미지 전송을 위한 딥러닝 기반의 향상된 결합 소스-채널 부호화



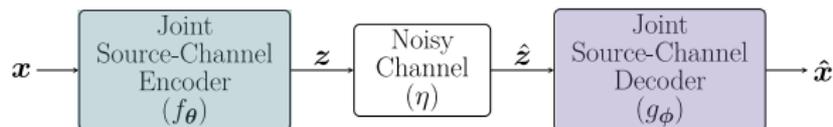
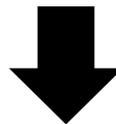
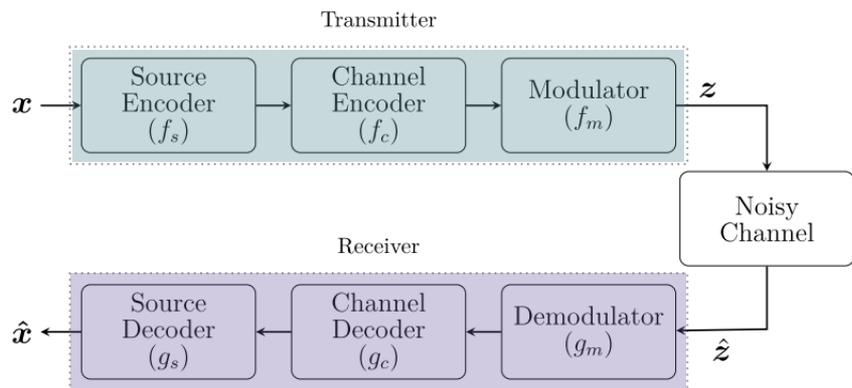
임수빈^o, 김정현^o, 송홍엽^{*}

순천향대학교^o, 연세대학교^{*}

qlsl0519@sch.ac.kr, kimjh@sch.ac.kr, hysong@yonsei.ac.kr

연구 배경

■ 결합 소스-채널 부호화의 사용



<소스, 채널 부호화의 분리>

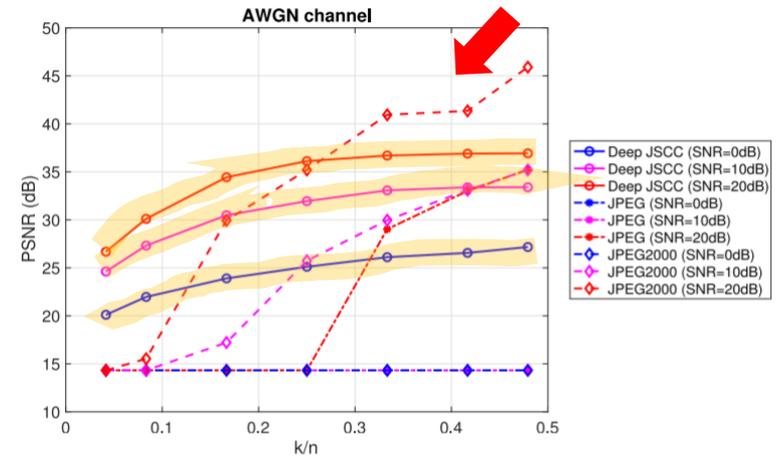
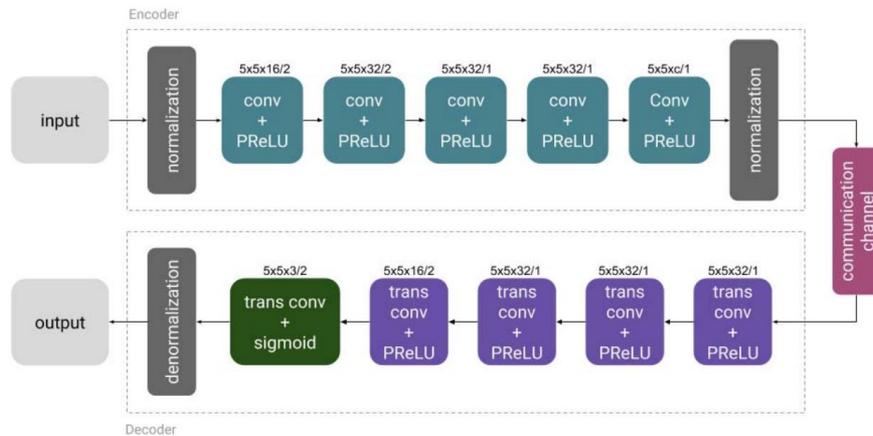
- **소스 부호화**
압축 알고리즘을 통해 이미지 픽셀들 간의 상관관계를 줄이고 전송되는 정보량을 줄임
ex) JPEG, JPEG2000
- **채널 부호화**
오류정정부호를 통해 무선 채널 환경에서 발생하는 오류를 극복함
ex) LDPC, Turbo, Polar code

<소스-채널 부호화의 결합>

- 소스 압축률과 채널 부호율을 가변적으로 적용하여 대역폭 사용의 효율을 높임

사전 연구 1

■ Deep JSCC (Deep Joint Source-Channel Coding)

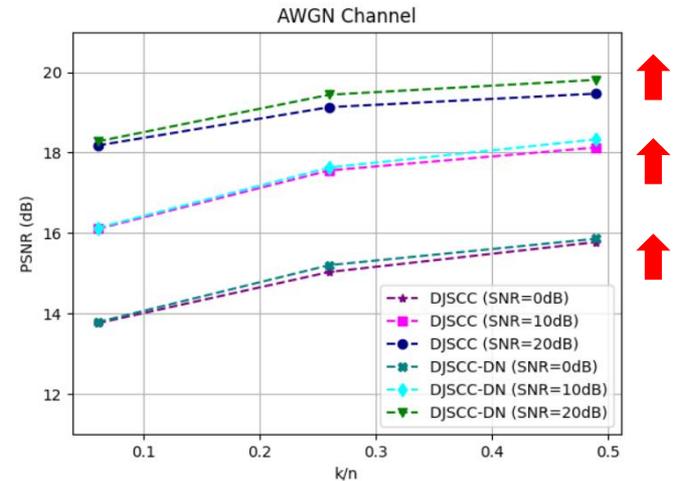
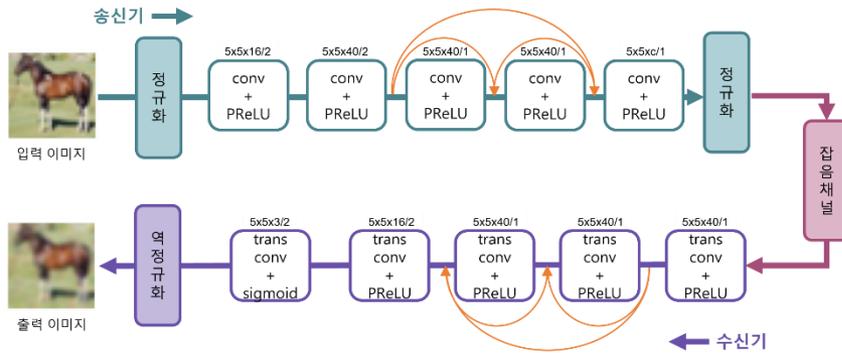


딥러닝 기반의 결합 소스-채널 부호화 (DJSCC)는 대역폭 압축률(k/n)의 변화에도 급격한 성능 저하를 겪지 않는 장점이 있지만, 일부 채널 환경(SNR)에서는 여전히 기존의 분리된 기법보다 낮은 성능을 보임

E. Boursoulatz, D. Burth 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, Sep. 2019.

사전 연구 2

■ DJSCC-DN (Deep Joint Source-Channel Coding using DenseNet)



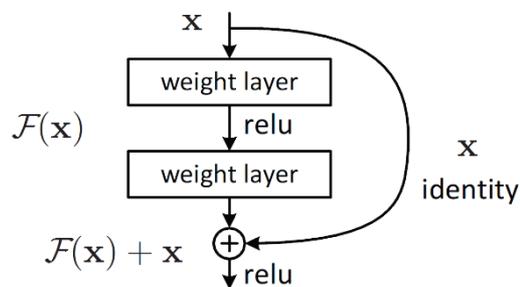
DenseNet 기법을 활용하여 DJSCC보다 더 나은 성능을 보이도록 제안된 DJSCC-DN 구조는 기존 모델보다 **향상된 정도가 매우 미약하여** 모델을 더욱 개선시킬 필요가 있음

임수빈, 김정현, 송홍엽. "무선 이미지 전송을 위한 딥러닝 기반의 효율적인 결합 소스-채널 부호화," 제2회 한국 인공지능 학술대회, Sep. 2021.

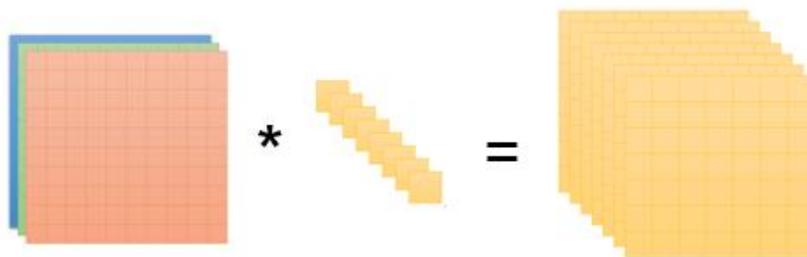
연구 목표

■ 결합 소스-채널 부호화의 성능 개선

1. 이미지 특징 추출 기능을 향상시키기 위해 층 간 연결을 하는 ResNet 구조 활용



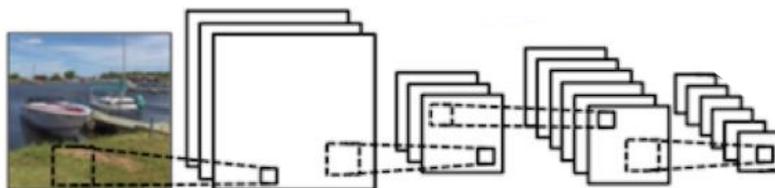
2. 연산량을 감소시키기 위해 1X1 크기의 필터 사용



개선 방안

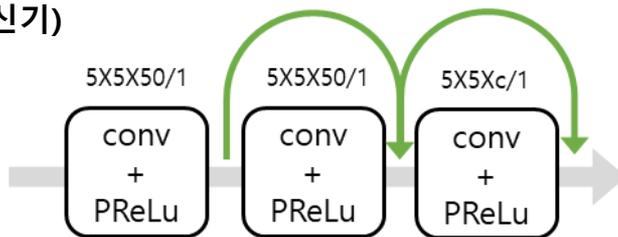
1. 이미지 특징 추출 기능을 향상시키기 위한 ResNet 기법

- 이미지에 대하여 필터와의 합성곱 연산 또는 풀링을 통해 **피쳐맵을 생성**
- 층이 점점 깊어지게 되면, **이전의 정보가 사라지는 문제 발생**

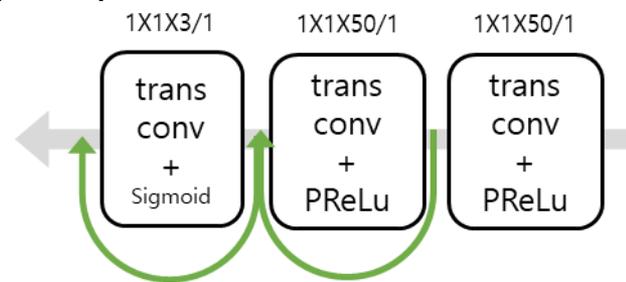


- **ResNet 기법을 활용하여 각 층에서 생성된 피쳐맵을 이전 층에서 생성된 피쳐맵과 연결**
- 정보가 손실되는 것을 방지

(송신기)



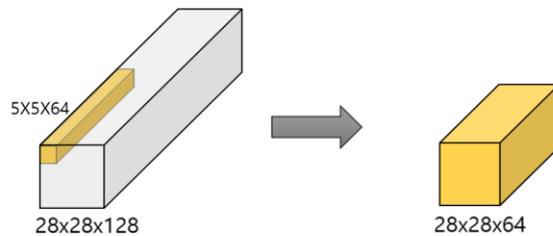
(수신기)



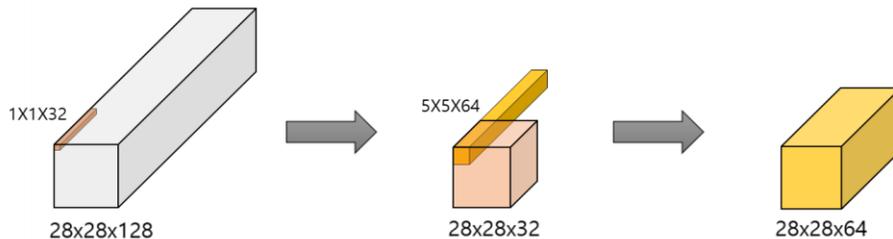
개선 방안

2. 연산량 감소를 위해 사용한 1X1 크기의 필터

- 심층 CNN 구조(GoogLeNet, Xception, MobileNets 등)에서 연산량 감소를 위해 1X1 크기의 필터를 적극적으로 사용
- 제안모델의 수신기의 컨볼루션 층에 적용함
- 지역적 특징을 활용하지 못하는 대신에 필터 수를 증가시켜 특징 추출에 효율적이도록 함



$$\#params = 28 \times 28 \times 64 \times 5 \times 5 \times 128 = 160M$$



$$\#params = 28 \times 28 \times 32 \times 128 \times 1 \times 1 = 4.8M$$

$$\#params = 28 \times 28 \times 64 \times 5 \times 5 \times 32 = 40M$$

$$\#total = 44.8M$$

대역폭 압축률이 0.06일 경우, 파라미터 수

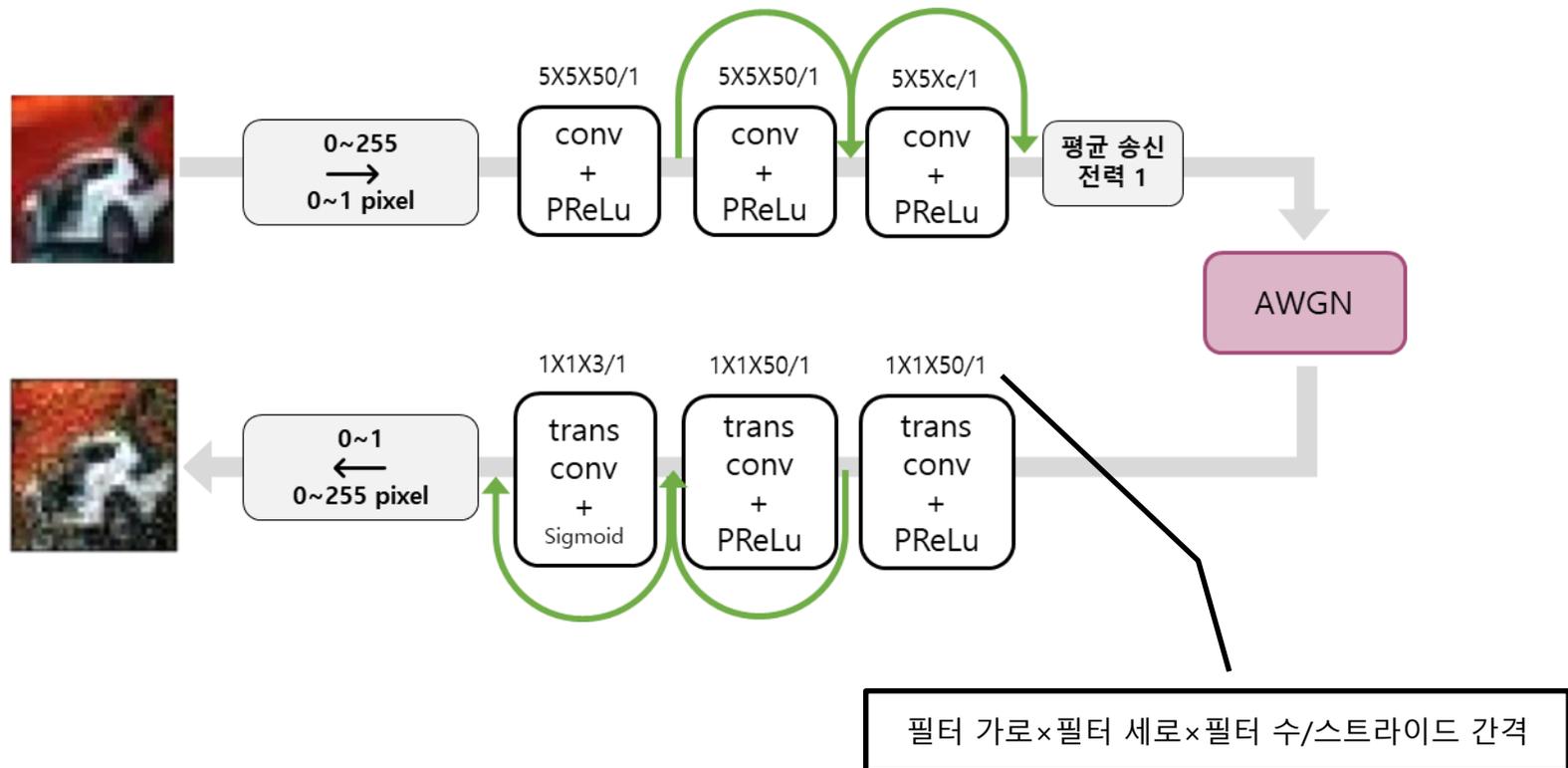
DJSCC-DN
369,699개



DJSCC-Advanced
290,180개

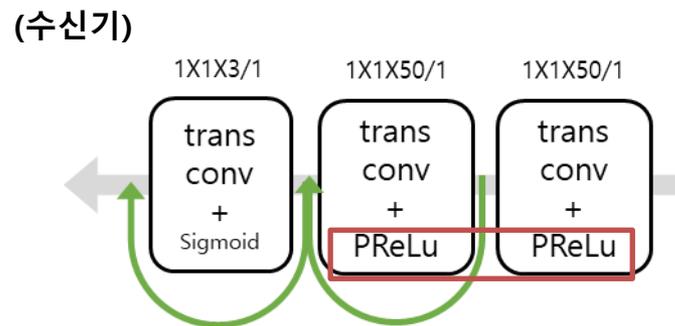
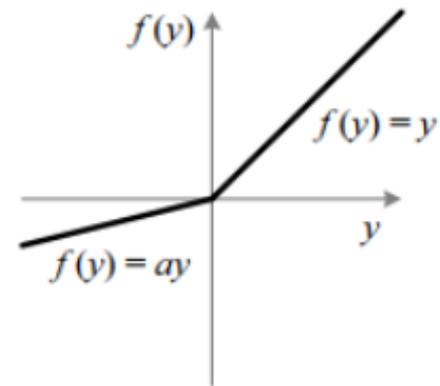
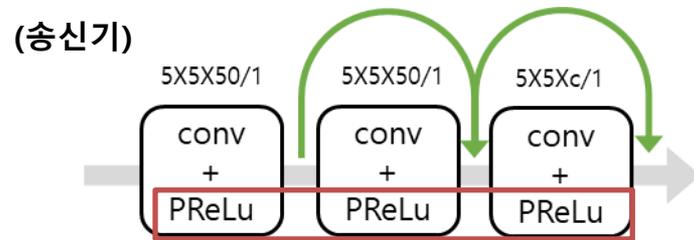
제안 모델

■ 전체 구조



제안 모델

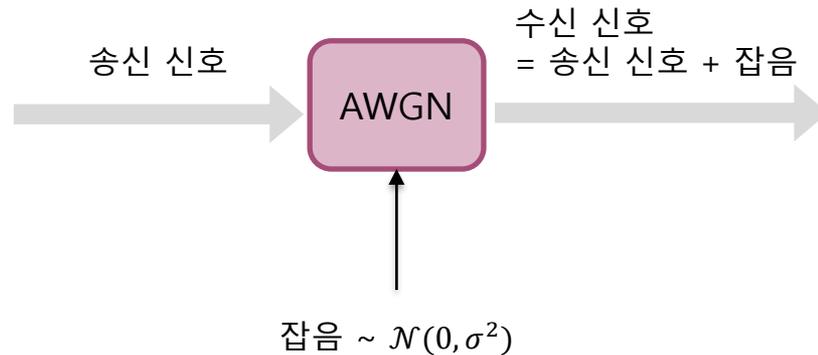
■ 활성화 함수 PReLU



- 은닉층에서 가장 많이 사용되는 ReLU 계열 함수
- y 가 음수인 영역에서도 학습 가능
- 각 층마다 알맞은 파라미터 a 를 학습

실험 환경

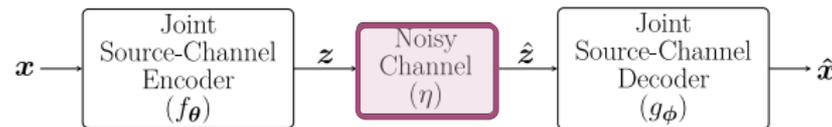
- 잡음채널 AWGN
(Additive White Gaussian Noise)



- SNR
(Signal-to-Noise Ratio)

$$(SNR)_{dB} = 10 \log_{10} \frac{\text{signal power}}{\text{noise power}}$$

- SNR이 클수록 채널 환경이 좋음
- SNR이 0, 10, 20dB인 경우에서 실험 진행



실험 환경

■ 대역폭 압축률 k/n

(송신기)



대역폭
 $n=32 \times 32$
(실수)

소스 부호화

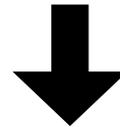
소스 압축률

채널 부호화

채널 부호율

대역폭
 $k=410$
(복소수)

잡음
채널



(송신기)



결합 소스-채널 부호화

대역폭 압축률
 $k/n = 0.4$

잡음
채널

k/n 이 0.06, 0.26, 0.4인 경우에서
실험 진행

데이터 및 파라미터 설정

▪ CIFAR-10 데이터셋



- 10개 클래스
- 32x32 픽셀의 컬러 이미지
- 학습 이미지 50,000장
- 테스트 이미지 10,000장

▪ 파라미터 설정 값

손실 함수	평균제곱오차(MSE)
최적화 방법	Adam, 학습률=0.001
배치 크기	100
반복 수	20

성능 평가 지표

■ PSNR (Peak Signal-to-Noise Ratio)

신호가 가질 수 있는 최대 전력에 대한 잡음 비율
주로 이미지 압축에서 화질 손실 정보를 평가할 때 사용

$$PSNR = 10 \log_{10} \frac{MAX^2}{MSE} (dB)$$

- MAX: 이미지의 한 픽셀이 가질 수 있는 최대값
- MSE: 송신 이미지와 수신 이미지의 픽셀 간의 평균제곱오차



Original image



PSNR 25.8699dB



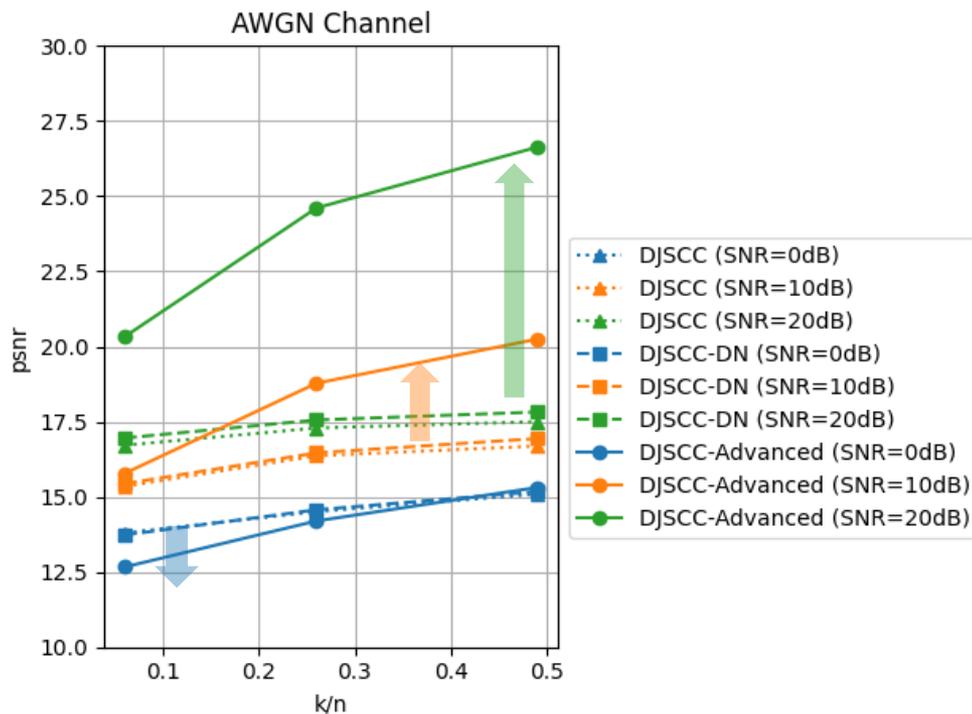
PSNR 30.9394dB



PSNR 34.8227dB

성능 평가 그래프

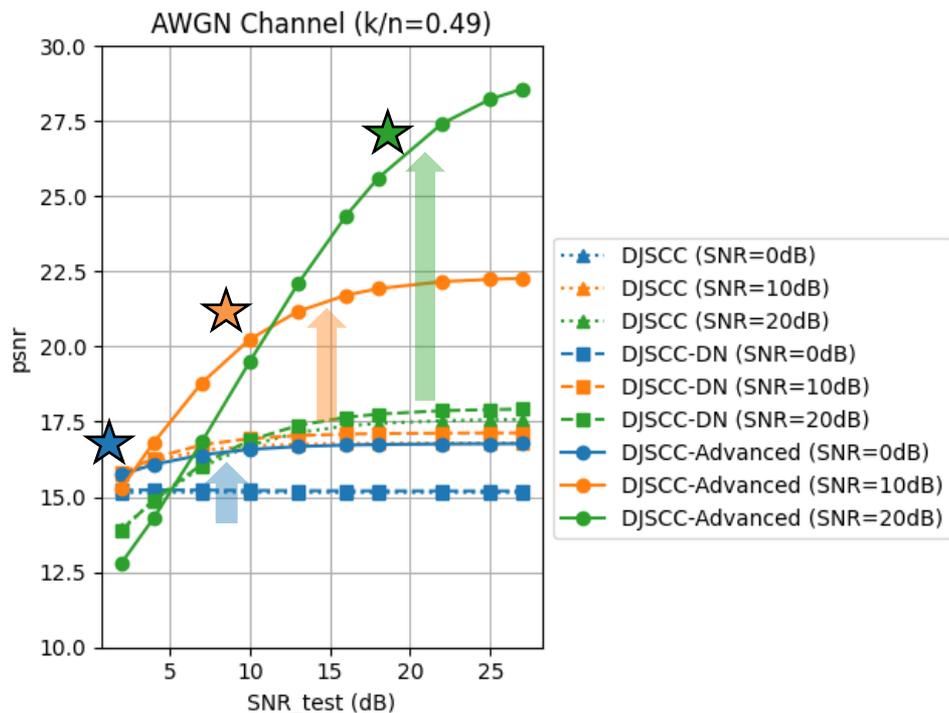
■ 대역폭 압축률(k/n)에 따른 성능



- SNR이 0dB인 경우를 제외하고, 10,20dB인 경우에는 모든 대역폭 압축률(k/n)에 대해 제안하는 기법이 기존 기법 대비 같거나 더 나은 성능을 보임

성능 평가 그래프

- 다양한 테스트 환경(SNR_test)에 따른 성능



- 테스트 환경(SNR_test)이 타겟 학습 환경(SNR_train=0, 10, 20dB)과 유사한 경우, 제안기법이 기존기법 대비 더 나은 성능을 보임
- 테스트 채널 환경이 좋을수록 격차는 크게 증가

향후 연구

- 새로운 신경망 구조 활용하여 추가 성능 개선
- JPEG, JPEG2000 등 기존의 분리된 소스-채널 부호화의 성능 비교
- 다양한 채널 환경에서 실험을 통한 성능 검증
- 다양한 크기의 이미지 데이터에 대한 성능 검증
- 성능 평가지표로 PSNR 뿐만 아니라 **SSIM**도 동시에 적용

경청해주셔서 감사합니다.
