

# 신경망 기술 기반 이미지 부호화 성능 개선

## 요약

정보통신기술과 하드웨어의 급격한 발전으로 연구소 및 기업 외에도 일반 사용자들이 대용량 이미지 데이터를 다룰 수 있게 되었다. 이렇게 초고화질 이미지에 대한 수요가 증가함에 따라 이미지 부호화 기술은 중요한 위치를 차지하고 있다. 또한, 최근 다양한 분야에 딥러닝 기반 기술이 적용되면서 이미지 부호화에서도 이를 활용한 기술이 제안되고 있다. 따라서 본 연구에서는 신경망 기술 기반의 영상 부호화 성능을 개선하는 방법을 제안하고자 한다.

먼저, 신경망 기반 이미지 부호화(Neural Network Image Compression, NNIC) 모델과 다운스케일링(Downscaling) 기법, 업스케일링(Upscaling) 기법을 통해 낮은 비트율에서의 성능 향상을 목표로 한다. 그 후, 입력 받은 이미지를 블록 기반으로 처리하는 방법을 도입하여 높은 비트율에서도 좋은 성능이 나오도록 한다. 성능 평가는 비트율(bitrate)과 PSNR을 통해 진행한다.

## 1. 서론

### 1.1 연구배경

오늘날 사람들은 하루동안 수많은 고화질 이미지를 활용하고 있다. 정보통신기술과 하드웨어의 급격한 발전으로 연구소 및 기업 외에도 일반 사용자들이 대용량 이미지 데이터를 다룰 수 있게 되었다. 이렇게 초고화질 이미지에 대한 수요가 증가함에 따라 이미지 부호화 기술은 중요한 위치를 차지하고 있다. 또한, 최근 다양한 분야에 딥러닝 기반 기술이 적용되면서 이미지 부호화에서도 이를 활용한 기술이 제안되고 있다. 따라서 본 연구에서는 신경망 기술 기반의 영상 부호화 성능을 개선하는 방법을 제안하고자 한다.

### 1.2 연구목표

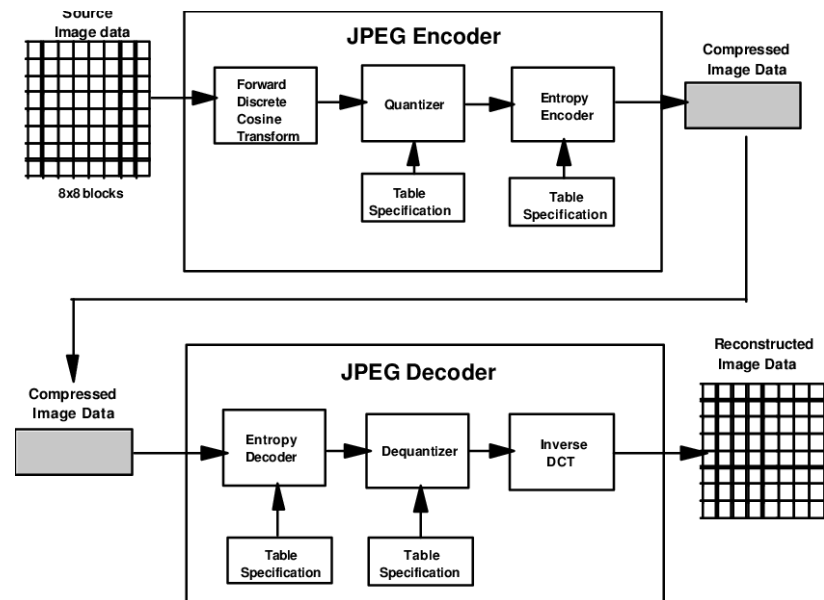
본 연구는 신경망 기반 이미지 부호화 모델을 만들고, 다운스케일링 기법, 업스케일링 기법과 블록 기반 처리 방법을 통해 기존 신경망 기반 이미지 부호화 모델보다 개선된

성능을 보이하고자 한다. 성능 평가는 비트율과 PSNR을 비교한 RD-Curve를 통해 진행한다.

## 2. 관련연구

### 2.1 JPEG

1992년 정지 화상을 위해서 만들어진 손실 압축 방법 표준이다. ISO와 ITU-T에서 제정하였다. JPEG/JFIF는 웹 상에서 사진 등의 화상을 보관하고 전송하는 데 가장 널리 사용되는 파일 형식이다. 하지만 이 압축 방법은 문자, 선, 세밀한 격자 등 고주파 성분이 많은 이미지의 변환에서는 GIF나 PNG에 비해 불리하며, 나쁜 품질을 보이는 경우가 많다.



### 2.2 JPEG 2000

JPEG 2000은 웨이블릿 변환에 기초한 래스터 이미지 압축 기술이다. ISO와 ITU의 공동 조직인 Joint Photographic Experts Group에서 JPEG의 후속으로 규격화되었다. JPEG보다 높은 압축률과 이미지 품질을 얻기 위해 개발되었다. 비손실 압축을 지원하고 한 이미지에 여러 해상도의 이미지를 포함할 수 있다.

JPEG 2000의 일부가 ISO 표준(ISO/IEC 15444-1)으로 발표되었으나, 2020년 기준 현재도 대부분의 웹 브라우저에서 지원되지 않는다.

## 2.3 NN-based Image Compression

NN-based Image Compression의 대표적인 기술로 변형가능한 오토 인코더를 기반의 end-to-end 모델이 연구되었다. 합성곱 오토 인코더의 은닉 벡터를 부호화하며, 비트스트림을 반복적으로 생성하는 경우 RNN Residual 오토 인코더가 활용되었다. 심층 신경망의 학습은 복원 이미지의 화질은 높이고 발생하는 비트량은 낮추는 방향으로 이뤄졌다. 기존의 JPEG, WebP 등과 비교하면 저 비트율에서의 복원 이미지 화질이 크게 개선되었다.[1][2]

## 3. 프로젝트 내용

Image를 input 라고 하자. NNIC는 앞선 논문에서 제안된 Mean Scale Hyperprior 모델을 기본으로 수정된 모델이다[3][4].

Input  $X$ 는 이미지라고 하자. 이미지를 입력으로 받는 NNIC의 블록도는 다음과 같다.

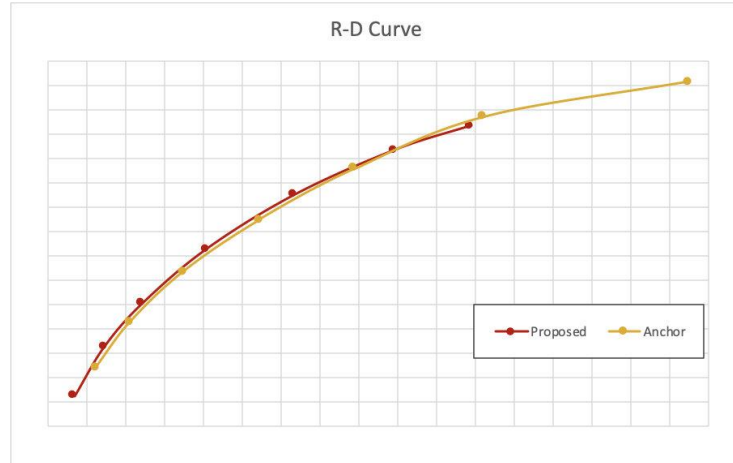


Image  $X$ 를 NNIC의 입력으로 넣는다. NNIC 과정을 통해 압축 후 복원된 이미지  $\hat{X}$ 를 얻는다.  $X$ 와  $\hat{X}$ 를 통해 PSNR을 구한 후, PSNR을 통해 성능평가를 진행한다. 이를 통해 얻은 결과는 Anchor가 되어 이후 진행할 연구의 기초 비교대상이 된다.

스케일링을 중간 과정에 포함시킨다. 블록도는 다음과 같다.

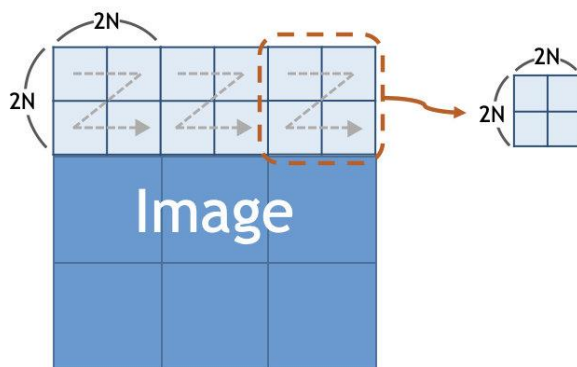


다own스케일링 과정을 통해 해상도를 낮추고, Super Resolution 또는 Bicubic Interpolation과 같은 기법을 통해 Upscaling 하여 결과를 얻는다. 낮은 bpp에선 Anchor보다 PSNR 성능이 좋게 나오지만, 높은 bpp에선 Anchor보다 더 낮은 성능을 보일 것이다. 즉, 저해상도에선 더 좋은 성능을 보이지만 고해상도에선 더 낮은 성능이 예상된다. R-D Curve를 확인해보면 아래와 같을 것으로 예상된다.

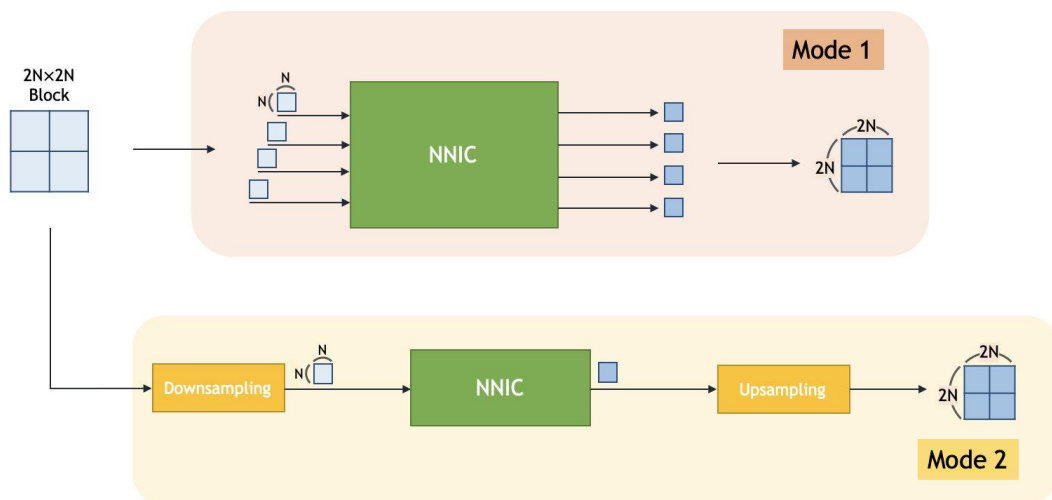


Scaling을 하게 되면 복잡도로 인해 하드웨어 처리에 불리해지고 bitrate에 따른 성능 변화가 생기므로 블록을 기반으로 처리하는 방안을 제안한다.

먼저 Image를 여러개의  $2N \times 2N$  블록으로 나눈다.

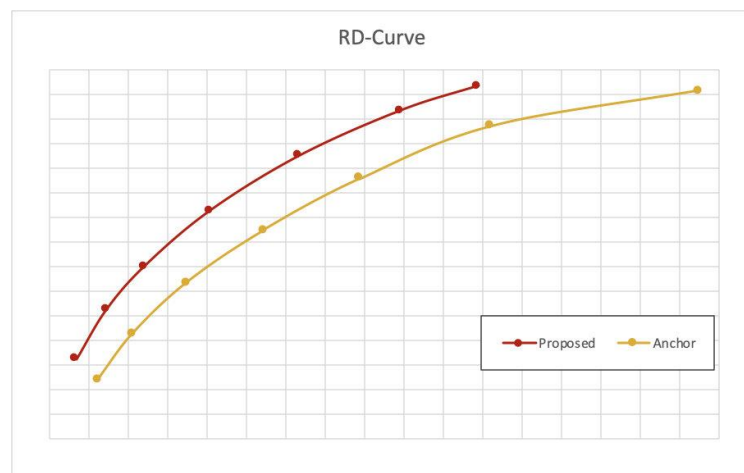


Mode 1과 Mode 2, 두가지 모드로 실험을 진행한다.



먼저 Mode 1에서는  $2N \times 2N$  블록을  $N \times N$  블록 4개로 나누어  $N \times N$  블록을 입력으

로 받는 NNIC에 통과시켜 복원된 블록을 얻는다. 이를 다시  $2N \times 2N$  블록으로 복원하여, 복원된  $2N \times 2N$  블록(하늘색)과 오리지널  $2N \times 2N$  블록(파란색)으로 성능을 측정한다. Mode 2에서는  $2N \times 2N$  블록을 Downsampling 하여  $N \times N$  블록으로 만든다. 이 블록을  $N \times N$  블록을 입력으로 받는 NNIC에 통과시켜 복원된 블록을 얻고, 이를  $2N \times 2N$  블록으로 Upsampling 한다. 복원된  $2N \times 2N$  블록(하늘색)과 오리지널  $2N \times 2N$  블록(파란색)으로 성능을 측정한다. 각각 Mode에서 입력이 이미지일 때와 블록일 때의 성능을 비교하고, mode끼리의 성능 또한 비교한다. 인코더에선 모드 정보를 저장하고, 디코더에선 모드 정보를 받아서 복원한다. 예상되는 결과는 다음과 같다.



#### 4. 결론 및 기대효과

이 방안은 기존 신경망 기술 기반의 영상 부호화 성능을 개선할 수 있을 것으로 보인다.

## 참고문헌

- [1] 조승현, 김연희, 임웅, 김휘용, 최진수, "딥 러닝 기반의 이미지와 비디오 압축 기술 분석", 방송공학회논문지 제23권 제3호, 2018.
- [2] Johannes Ballé, David Minnen, Saurabh Singh, Sung Jin Hwang, Nick Johnston, "Variational image compression with a scale hyperprior", ICLR, 2018.
- [3] Ballé, Johannes, et al. "Variational image compression with a scale hyperprior." arXiv preprint arXiv:1802.01436 (2018).
- [4] Minnen, David, Johannes Ballé, and George D. Toderici. "Joint autoregressive and hierarchical priors for learned image compression." Advances in neural information processing systems 31 (2018).
- [5]