

얼굴 인식에 적합한 AI 업스케일링 알고리즘에 관한 연구

곽두일¹, 박광영¹
¹승실대학교 AI테크노융합학과
 ai@soongsil.ac.kr, 1004pky@ssu.ac.kr

A study on AI upscaling algorithms suitable for facial recognition

Doo-il Kwak¹, Kwang-Young Park¹
¹Dept. of AI Techno Convergence, Soongsil University

요 약

CCTV가 범죄 예방 및 수사에 사용되는데, 수사를 위해 저화질 CCTV 영상에서 특정인의 얼굴 인식엔 어려움을 겪어 CCTV 본연의 역할의 희석된다. 따라서 본 논문은 저화질 영상을 고화질로 변환하여 얼굴 인식의 정확성을 높일 수 있는 알고리즘을 연구하는 것을 목적으로 한다. 기존에 연구된 인공지능 기반의 업스케일링 알고리즘을 분석하여 K-FACE 데이터셋에 적절한 모델을 제안한다. 이를 위해 2020년 이전과 이후의 AI 업스케일링 관련 연구를 비교 분석한다. 향후 제시된 모델을 대상으로 동일한 환경내에서 K-FACE 데이터셋을 학습시켜 통일된 기준의 지표 산출이 필요하다.

1. 서론

범죄자 색출에 결정적인 단서의 제공 및 시설 안정과 화재 예방, 교통 단속 등의 목적으로 폐쇄회로 텔레비전(이하 CCTV)가 급속도로 설치 및 운영되고 있다. 통계청 자료에 따르면 국내 CCTV 설치 및 운영 대수는 2021년 145만 8,465대로 2008년 대비 827.8% 증가했으며, 한국인터넷진흥원에 따르면 CCTV의 대수는 1,600만대로 추정된다. 하지만, 범죄 수사를 위해 수사 대상자의 신원 확인 시 CCTV를 통해 확보한 대부분의 저화질 화면에서는 얼굴 인식 성능 저하로 인해 신원 확인에 어려움을 겪게 된다.

따라서, 본 논문은 CCTV에서 출력된 저화질 화면에서 특정 사건 발생 시 특정인을 특정하여 고화질 화면으로 출력할 수 있는 알고리즘 개발에 필요한 업스케일링(upscaling) 알고리즘 기술을 분석하여 가장 우수한 알고리즘을 선택하여 제안한다. 기존 연구된 보간법(interpolation Algorithm)은 구현이 간단하고 쉽지만, 이미지 스케일링만 가능하고 품질을 개선해주지는 않는다. 그렇기에 인공지능을 활용한 'AI 업스케일링' 모델을 분석하고, 얼굴 인식에 한정하여 제안된 최신 알고리즘들을 비교 분석하여 얼굴

에 가장 적합한 알고리즘을 제안하는 것을 목표로 한다.



(그림 1) 얼굴에 특화된 업스케일링 모델 시연 예시

2. 2020년 이전의 AI 업스케일링 관련 연구

하단 도표는 인공지능 업스케일링에 사용된 모델에 대한 비교 분석이다.

PSNR은 Peak Signal-to-Noise Ratio를 의미하며, SSIM은 Structural Similarity Index를 의미한다. 두 측정 지표 모두 높을수록 원래 이미지와 가까운 고해상도 이미지를 생성하는 능력이 높음을 나타낸다.

- SRCNN(Super-Resolution Convolutional Neural Network)[2]: 3개의 합성곱 레이어를 사용하여 입력 이미지를 고해상도 이미지로 변환한다. 속도가 빠르고, 더 많은 데이터와 깊은 모델 사용으로 성능을 높일 수 있다.

<표 1> 2014~2018년간 AI 업스케일링 모델 비교

모델	연도	최적화 알고리즘	추가적인 기술	PSNR(dB)	SSIM ↑	데이터셋
SRCNN	2014	SGD	없음	32.33	0.9090	Set14
EDSR	2017	Adam	Residual Learning, Dense Connection	32.82	0.9290	Set14
SRGAN	2017	Adam	GAN, VGG loss, Feature loss	30.28	0.8472	Set14
ESRGAN	2018	Adam	GAN, Relativistic Discriminator, Feature loss	30.50	0.9155	Set14

- EDSR(Enhanced Deep Super-Resolution)^[3]: Batch Normalization을 제거하여 GPU의 사용량을 획기적으로 줄이고 유연성을 증가시켰다. 줄어든 메모리로 더 깊고 넓은 신경망 구조를 만들 수 있다.
- SRGAN(Super-Resolution Generative Adversarial Network)^[4]: GAN을 사용하여 고해상도 이미지를 생성한다. 시각 지표인 MOS 테스트에서 획기적인 성능 향상을 이루었다. 네트워크 구조가 깊을수록 성능이 좋지만 high-frequency artifact 이슈가 있어 학습이 어렵다.
- ESRGAN(Enhanced Super-Resolution Generative Adversarial Network)^[5]: SRGAN에서 Batch Normalization을 제거하고 Relativistic Discriminator를 도입하여 더 선명한 이미지 생성을 가능하게 한다.

3. 2020년 이후의 AI 업스케일링 관련 연구

2020년에 출판된 'PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models'^[6]은 자기 감독 학습 방식을 이용한 고해상도 이미지 생성 방법이다. 이 논문에서는 이미지의 잠재 공간을 탐색하고, 이를 기반으로 고해상도 이미지를 생성한다. 이 방법은 기존의 고해상도 이미지 생성 방법들과 달리, 대규모 데이터셋이 필요하지 않아도 높은 성능을 보인다.

2021년에 나온 'High Fidelity Image Generation Using Diffusion Models'^[7]는 최근에 인기를 끌고 있는 확산 모델을 이용한 고해상도 이미지 생성 방법이다. 이미지를 확산하는 과정을 이용하여 고해상

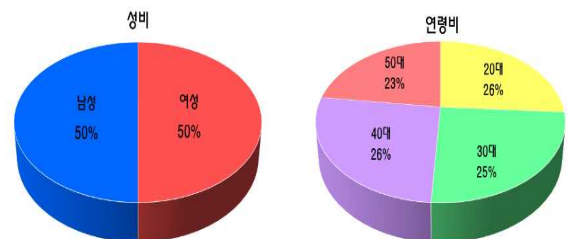
<표 2> 기존 모델 일부와 최신 알고리즘 비교

모델	PSNR	SSIM	학습 시간	하드웨어	데이터셋
PULSE	22.21	0.55	1일 (100K 반복)	NVIDIA V100 GPU 1개	CelebA HQ
Diffusion Model	N/A	N/A	3일 (300K 반복)	NVIDIA V100 GPU 8개 + TPU v3-32 Pod 1개	CelebA-HQ + ImageNet

도 이미지를 생성하는데, 기존의 고해상도 이미지 생성 방법들과 달리, 높은 해상도를 지원하면서도 매우 고품질의 이미지를 생성할 수 있다. 이 방법은 여러 이미지 데이터셋을 이용하여 실험을 진행했으며, 실험 결과는 매우 높은 성능을 보여주었지만 데이터셋과 측정지표가 달라 비교가 어려웠다.

4. K-FACE 데이터베이스

AI 업스케일링 알고리즘의 성능은 학습 데이터의 양과 질, 모델의 구조에 따라 달라진다. 따라서 한국에서 수집되는 CCTV 영상 속 특정인의 신원 확인을 위한 기술은 학습 데이터 단계에서부터 고려되어야 한다. K-FACE 데이터베이스는 한국인 얼굴의 통계적 특징을 반영하는 한국인 얼굴 데이터베이스 구축을 목표로 하여 만들어진 데이터 세트로서 성별과 연령을 고려한 1,000명의 피실험자에 대해 100만장 이상의 HR 이미지를 구축했다. 포즈, 조명 등의 외부 환경은 물론이고, 20대부터 50대까지의 연령대도 고려하여 구축되었다^[8]. 이는 본 저자가 구현하고자 하는 'CCTV에서 추출된 저화질 이미지의 업스케일링으로 목표하는 사람의 얼굴을 정확하게 식별할 수 있는 고화질 이미지로의 업스케일 모델 구현'에 적합한 데이터베이스라 할 수 있다.



(그림 2) K-FACE 데이터셋의 인구통계학적 분포

5. K-FACE 데이터셋에 적절한 알고리즘

K-FACE 데이터셋은 한국인의 특성을 고려하여 구축된 고화질 얼굴 이미지 데이터셋으로, 저화질 CCTV 영상에서 추출된 얼굴 이미지의 업스케일링에 적합한 데이터셋이다.

PULSE와 Diffusion Models와 같이 자기 감독 학습 방식을 이용한 고해상도 이미지 생성 알고리즘은 대규모 데이터셋이 필요하지 않아도 높은 성능을 보이는 장점이 있다. 따라서 K-FACE 데이터셋과 같이 작은 규모의 데이터셋을 이용하여도 충분히 좋은 결과를 얻을 수 있다는 장점이 있다.

또한, K-FACE 데이터셋은 연령, 성별 등 다양한 속성을 가진 얼굴 이미지를 포함하고 있으므로, 이를 이용하여 속성별 업스케일링 알고리즘을 개발하는 것도 가능하다. 예를 들어, 연령대별로 다른 업스케일링 알고리즘을 적용하여 더욱 정확한 결과를 얻을 수 있다.

6. 결론

본 논문은 저해상도 CCTV 영상을 인공지능 업스케일링을 통해 고해상도 영상으로 업스케일링하여 특정 개인을 식별할 수 있는 알고리즘을 제안하는 것을 목표로 한다. 이를 위해 업스케일링 알고리즘을 분석하고, 얼굴 인식을 위해 제안된 최신 알고리즘을 소개했다. 자기 감독 학습 방식을 이용한 PULSE와 Diffusion Models와 같은 알고리즘은 한국인의 특성을 고려한 K-FACE 데이터셋과 함께 사용한다면 더욱 정확하고 고품질의 이미지 생성이 가능하다. 따라서 K-FACE 데이터셋과 최근 발전되는 딥러닝 알고리즘을 활용하여, 보다 정확하고 고품질의 AI 업스케일링이 가능해보인다.

본 논문은 범죄 수사 및 공공 안전에 필수적인 저해상도 CCTV 영상에서 개인을 식별하는 문제에 대한 해결책을 제시하기 위해 학습 모델에서 사용하는 데이터셋을 먼저 선정하고 이에 맞는 알고리즘을 찾는 시도를 하였다. 하지만 제시된 알고리즘 간 동일한 데이터셋이나 하드웨어를 통한 비교 분석이 되지 않아 정확한 모델 간 비교가 되지 않았다는 한계가 있다. 향후 K-FACE 데이터베이스를 사용하여 본 논문에서 제시된 알고리즘을 적용한 실험을 진행하고 측정 지표를 산출하여 최상의 적절한 알고리즘을 선정하는 후속 연구가 필요하다.

ACKNOWLEDGMENT

“본 연구는 과학기술정보통신부 및 정보통신기획평가원의 지역지능화혁신인재양성사업의 연구결과로 수행되었음” (IITP-2023-RS-2022-00156360)

참고문헌

- [1] 지표서비스 | E-나라지표. “공공기관 CCTV 설치 및 운영”. Retrieved from https://www.index.go.kr/unity/potal/main/EachDtIPageDetail.do?idx_cd=2855. 2022.
- [2] Dong, C., Loy, C. C., He, K., & Tang, X. “Image Super-Resolution Using Deep Convolutional Networks.” 2014. <https://arxiv.org/abs/1501.00092>
- [3] Lim, B., Son, S., Kim, H., Nah, S., & Lee, K. M. “Enhanced Deep Residual Networks for Single Image Super-Resolution.” 2017. <https://arxiv.org/abs/1707.02921>
- [4] Ledig, C., Theis, L., Huszar, F., Caballero, J., Cunningham, A., Acosta, A., Aitken, A., Tejani, A., Totz, J., Wang, Z., & Shi, W. “Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network.” 2016. <https://arxiv.org/abs/1609.04802>
- [5] Wang, X., Yu, K., Wu, S., Gu, J., Liu, Y., Dong, C., Loy, C. C., Qiao, Y., & Tang, X. “ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks.” 2018. <https://arxiv.org/abs/1809.00219>
- [6] Menon, S., Damian, A., Hu, S., Ravi, N., & Rudin, C. “PULSE: Self-Supervised Photo Upsampling via Latent Space Exploration of Generative Models.” 2020. <https://arxiv.org/abs/2003.03808>
- [7] Ho, J., Saharia, C., Chan, W., Fleet, D. J., Norouzi, M., & Salimans, T. “Cascaded Diffusion Models for High Fidelity Image Generation.” 2021. <https://arxiv.org/abs/2106.15282>
- [8] 한국지능정보사회진흥원. “한국인 안면 이미지”. 2020, https://github.com/k-face/k-face_2019