

웹툰 남주가 토르라면

곽도윤, 김건우, 윤형진, 이유주

25-1 DSL Modeling, Team Generative Model

March 28th, 2025

Abstract

본 보고서는 네이버 웹툰『국세청 망나니』의 주인공 ‘신재현’의 얼굴을 사용자가 선택한 특정 얼굴로 대체하는 생성모델 기반의 이미지 변환 파이프라인을 제안한다. 파이프라인은 웹툰 이미지 수집 및 전처리, Advanced Classifier(VGG19)를 통한 특정 인물의 얼굴 분류, JoJoGAN을 이용한 이미지 스타일 전이, facial landmark 기반 Deepfake 기술을 활용한 얼굴 교체, U-Net 및 PatchGAN 기반 이미지 복원, OCR 기반 이름 교체까지의 과정을 포함한다.

1 Introduction

최근 생성모델 분야의 발전은 이미지를 현실적이고 자연스럽게 변형하는 다양한 응용을 가능하게 했다. 그 중에서도 특히 얼굴을 특정 인물로 교체하는 기술은 개인화된 콘텐츠 제작과 사용자 맞춤형 미디어에 높은 활용 가능성을 보인다. 본 프로젝트에서는 네이버 웹툰『국세청 망나니』의 주인공 ‘신재현’의 얼굴을 크리스 햄스워스와 같은 유명인의 얼굴로 자연스럽게 교체하는 것을 목표로 하며, 전체 프로세스를 자동화할 수 있는 생성모델 기반의 통합 파이프라인을 제안한다.

2 Pipeline

본 프로젝트의 전체 파이프라인은 웹툰 이미지 크롤링, 데이터셋 구축, 얼굴 분류(Classification), 이미지 스타일 전이(Style Transfer), 얼굴 교체(Face Swap), 얼굴 복원(Face Inpainting), 이름 변경(OCR) 등의 단계로 구성된다.

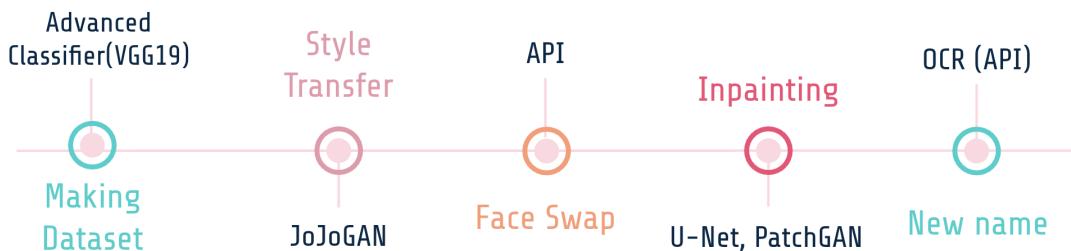


Figure 1: 웹툰 얼굴 변환 파이프라인 전체 흐름

Data Collection Preprocessing: 웹툰『국세청 망나니』에서 총 3,523장의 이미지를 크롤링하여, Face++ API를 활용해 얼굴과 머리카락까지 포함한 영역을 bbox로 크롭하였다. 크롭된 이미지의

좌표값은 JSON 형태로 저장되었으며, (크기 7.21KB 미만, 해상도 100x100 pixel 미만) 저화질 이미지를 필터링하여 최종적으로 고화질 얼굴 데이터 2,571장을 확보하였다. 이 과정에서 Face++ API가 일부 얼굴을 인식하지 못해 수작업 보완이 필요했다.

Face Classification: 정확한 얼굴 분류를 위해 CLIP과 Logistic Regression을 결합한 Naive Classifier를 먼저 적용하여 신재현 후보 713장을 추출하였다. 이후 수작업 검토를 거쳐 최종 VGG19 학습용 데이터를 구축(신재현 283장, 신재현이 아닌 인물 547장)하였으며, VGG19 모델의 마지막 6개 layer를 Fine-tuning하여 최종적으로 867장의 데이터를 확보하였다. Fine-tuning 시 VGG19의 기본 학습률의 1/10을 적용하여 학습하였다.



Figure 2: VGG19로 분류된 얼굴 이미지 예시. (a)는 신재현으로 분류된 이미지들이며, (b)는 기타 등장인물로 분류된 이미지들이다.

	precision	recall	f1-score	support
0	1.00	0.99	1.00	110
1	0.98	1.00	0.99	58
accuracy			0.99	168
macro avg	0.99	1.00	0.99	168
weighted avg	0.99	0.99	0.99	168

Figure 3: VGG19 기반 얼굴 분류기의 검증 결과. 두 클래스에서 모두 높은 정밀도와 재현율을 기록 하였으며, 주인공 식별에 있어 우수한 성능을 보였다.

Style Transfer Face Swap: Style Transfer를 위해 Prompt 입력 없이 작동하며 네이버 웹툰 사례에서 검증된 JoJoGAN을 선택하였다. JoJoGAN 적용 시 Identity Loss를 추가하여 극단적 스타일 변화에도 원본 얼굴 정체성을 보존할 수 있었다. Style Mixing 레이어 조정을 통해 (예: layer 7-17은 포즈 및 헤어스타일 보존, layer 7-9는 색감 정보 보존) 최적의 결과를 도출하였다.

$$w_i = (1 - \alpha)M \cdot w + \alpha(1 - M) \cdot FC(z_i) \quad (1)$$

여기서 M 은 style mixing mask, $FC(z_i)$ 는 latent vector z_i 에 대한 fully connected transformation을 의미한다.

얼굴 교체는 Deepfake에서 활용되는 facial landmark 기반 Face Swap 방식을 사용하여 표정, 시선, 얼굴 방향의 불일치 문제를 효과적으로 해결하였으며, Style Transfer 이후에 적용하는 것이 최적임을 확인하였다.

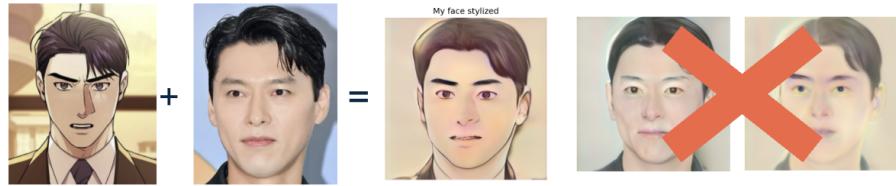


Figure 4: Identity 보존 여부에 따른 스타일 전이 결과 예시. 왼쪽은 원본 얼굴의 주요 특징이 유지된 경우이며, 오른쪽은 정체성이 유지되지 않은 실패 사례이다.

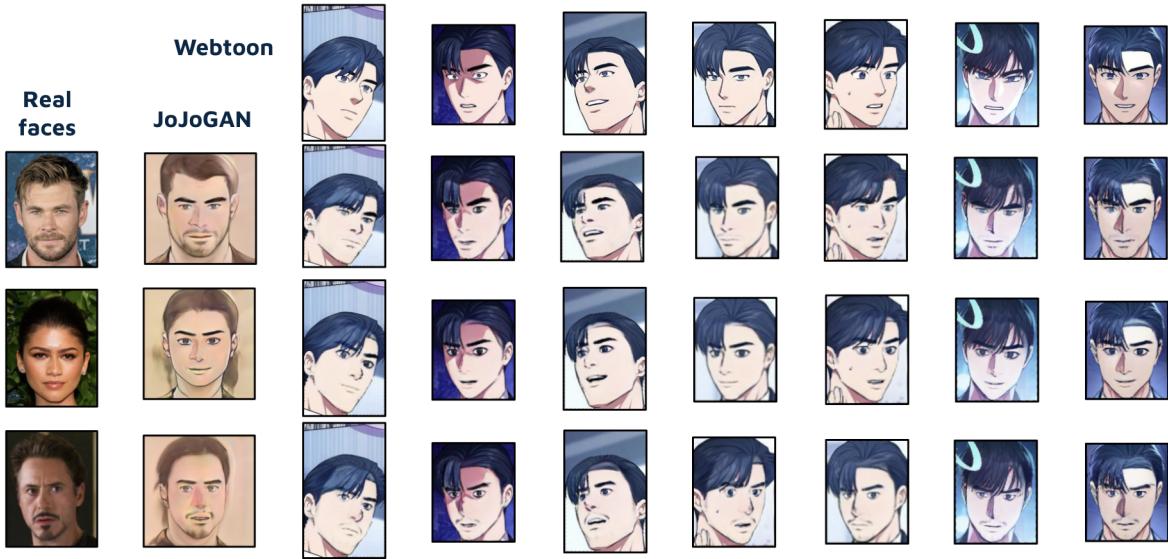


Figure 5: JoJoGAN을 이용한 스타일 전이 결과. 좌측 두 열은 실제 얼굴과 해당 스타일 전이 결과이며, 이후 열은 웹툰 스타일의 얼굴 변환 결과이다. 다양한 표정과 밝기에서도 원본 인물의 정체성이 유지됨을 확인할 수 있다.

Face Inpainting: 얼굴 교체 후 발생한 이미지 손상 및 부자연스러운 경계를 복원하기 위해 Face Inpainting을 적용하였다. 초기 AutoEncoder 방식의 시도는 성능이 낮았으므로, 경계 특성 보존에 강한 U-Net과 PatchGAN을 선택하였다. 데이터 분포 분석을 통해 Random Masking 크기를 128x128 픽셀로 결정하였으며, L1 Loss, Adversarial Loss, Total Variation Loss, Boundary Loss를 조합하여 학습하였다. Batch Size는 GPU 자원 제약으로 2로 설정하였고, Epoch 10에서 성능 정체를 확인하여 추가 학습은 중단하였다.

OCR-based Name Replacement: 웹툰 이미지에서 주인공 이름을 자동으로 추출 및 교체하기 위해 OCR API를 활용하였다. 추출된 텍스트 중 주인공 이름이 포함된 부분만 선택적으로 변경하여 최종 결과물의 자연스러움을 확보하였다.



Figure 6: OCR 기반 이름 교체 과정 예시 (Episode 112, Cut 64). (a)는 원본 이미지이며, (b)는 텍스트 제거를 위한 inpainting 결과, (c)는 ‘신재현’을 ‘크리스’로 자연스럽게 교체한 최종 결과이다.

3 Model

본 파이프라인에서는 여러 최신 생성모델 및 이미지 처리 기법이 통합적으로 사용되었다. 얼굴 분류 단계에서는 VGG19 모델을 활용하여 Fine-tuning을 통해 높은 정확도를 달성하였고, Style Transfer 단계에서는 JoJoGAN을 사용하여 Identity Loss 추가와 Hyperparameter 최적화를 통해 원본 인물의 특성을 유지하면서 웹툰 스타일로 변환하였다. 또한, 얼굴 교체는 Deepfake 기술에서 활용되는 facial landmark 기반 Face Swap 방식을 도입하였고, 이미지 복원에는 경계 보존에 강점을 가진 U-Net과 PatchGAN 모델을 적용하여 높은 복원 성능을 보였다.

4 Performance and Results

얼굴 분류 결과, VGG19 Classifier는 평균 정밀도 0.99, 재현율 1.00을 기록하였으며, 총 867장의 고정밀 데이터를 확보하였다. JoJoGAN 기반의 Style Transfer는 Identity Loss와 세부 레이어 조정을 통해 원본 인물의 특성을 효과적으로 유지하면서 웹툰 스타일로 변환하는 데 성공하였다. Face Swap 기술을 통해 얼굴의 방향, 표정, 시선의 불일치 문제를 자연스럽게 보정하였으며, Face Inpainting은 L1, Adversarial, Total Variation, Boundary Loss의 조합으로 부자연스러운 경계를 효과적으로 복원하였다. 최종적으로 OCR 기반 이름 변경을 통해 결과물 내 텍스트도 성공적으로 교체되었다.

최종 결과 예시는 Figure 7에 제시된 바와 같이, 112화의 70번째 컷에 적용되었다. 얼굴 교체와 텍스트 교체가 모두 적용되었으며, 자연스러운 연결성과 캐릭터 정체성 유지 측면에서 우수한 품질을 보였다.

5 Limitation & Future Work

Face++ API의 얼굴 탐지 과정에서 일부 실패 사례가 발생하여 수작업 보완이 필요하였고, JoJoGAN은 얼굴 형상이나 표정의 급격한 차이가 있는 경우 만족스러운 결과를 얻기 어려웠다. 또한, Face Swap 기술은 Deepfake 윤리 문제로 인한 데이터 확보의 한계가 존재하며, 복잡한 이미지 복원 시 채도 변화 등의 문제가 발생하였다. 향후 파이프라인 전용 추가 학습을 통해 복원 성능을 더욱 향상시키고, Invisible Watermark 기술 도입을 통해 저작권 보호 및 위/변조 방지 기능을 강화하는 연구가 필요하다.

6 References

- Chong, M. J., & Forsyth, D. A. (2022). JoJoGAN: One shot face stylization.
- Kim, S., & Baek, J. (2023). WebtoonMe 개발기: NAVER DEVVIEW 2023.
- Albisua et al. (2012). The quest for optimal class distribution. Progress in Artificial Intelligence, 2(1), 45–63.



(a) Original



(b) Final Result

Figure 7: Episode 112, Cut 70에서의 최종 결과 예시. (a)는 원본 이미지이며, (b)는 얼굴과 이름이 모두 변경된 결과로, 주인공 ‘신재현’의 얼굴은 크리스 햄스워스로 교체되었고, 이름 역시 OCR 기반으로 자연스럽게 수정되었다.