

## StyleGAN을 이용한 미래 2세대 얼굴 예측 웹 서비스

김 황<sup>0</sup>, 김민정\*, 이지현\*, 정진아\*, 김동욱\*, 곽호영\*\*

<sup>0</sup>제주대학교 대학원 컴퓨터공학과,

\*제주대학교 소프트웨어학부 컴퓨터공학전공,

\*\*제주대학교 컴퓨터공학과

e-mail: kimhwang@sk.com<sup>0</sup>, {k1950035, bomnae1005, SpaceDust, bradl0809}@stu.jejunu.ac.kr\*,  
kwak@jejunu.ac.kr\*\*

## Future 2<sup>nd</sup> generation face prediction web service using StyleGAN

Hwang Kim<sup>0</sup>, Min-Jeong Kim\*, Ji-Hyeon Lee\*, Jin-Ah Jung\*, Dong-Uk Kim\*, Ho-Young Kwak\*\*

<sup>0</sup>Dept. of Engineering, Graduate School, Jeju Nat'l University,

\*Dept. of Computer Engineering, Jeju Nat'l University,

\*\*Dept. of Computer Engineering, Jeju Nat'l University

### ● 요약 ●

최근 생성형 AI에 대한 수요가 상승하고 있으며, MZ세대의 자기에 성향으로 자신의 얼굴을 활용한 미디어 콘텐츠에 대한 호기심이 높아지고 있다. 이에 따라 본 논문에서는 MZ세대의 창의성과 미디어 소비를 고취시키기 위해, StyleGAN 기술을 중심으로 자신과 닮은 2세의 가상 모습을 생성하는 웹 서비스를 설계하고 구현하였다.

**키워드:** styleGAN, 2세대(2<sup>nd</sup> generation), perceptual loss, 얼굴 합성(face prediction)

## I. Introduction

최근 생성형 AI를 활용한 합성 미디어 콘텐츠에 대한 수요가 급증하고 있다. 이러한 흐름은 특히 MZ세대의 자기에 성향과 함께 더욱 두드러지게 나타난다. MZ세대는 자신의 모습을 기반으로 한 창의적이고 독특한 디지털 콘텐츠에 대한 특별한 관심을 보이고 있다.

본 논문에서는 이런 MZ세대의 미디어 소비 성향을 고려한 미디어 콘텐츠로서, 'StyleGAN'을 이용한 2세대 얼굴 합성 서비스를 제안하였다. 이 서비스를 통하여 사용자의 얼굴을 기반으로 가상의 2세의 모습을 생성할 수 있는 기술적 플랫폼을 제공함으로써, 자신의 모습을 바탕으로 창의적인 미디어 콘텐츠를 만들어내는 즐거움을 경험할 수 있도록 하였다.

## II. Preliminaries

### 1. Related works

#### 1.1 StyleGAN과 StyleGAN Encoder

컴퓨터 시스템이 인공지능 기술을 이용하여 새로운 콘텐츠(텍스트, 이미지, 음악 등)를 만들어내는 생성형 AI 연구에 있어서, 생성자와 평가자가 서로 경쟁하면서 새로운 데이터를 생성하는 네트워크인 GAN의 경우 다양한 분야에서 혁신적인 성과를 보여주고 있다.

StyleGAN은 NVIDIA에서 개발한 GAN 기반의 이미지 생성 모델로 고품질 및 고해상도의 가상 이미지의 생성에 사용되고 있으며, StyleGAN Encoder는 기존 이미지를 StyleGAN의 잠재 공간으로 매핑하는 역할을 하는 도구로 StyleGAN이 해석 가능한 형식으로 변환한 후, 해당 이미지의 특정 스타일이나 속성을 수정할 수 있게 한다[1][2].

### III. The Proposed Scheme

#### 1. 시스템 아키텍처 설계

본 연구에서는 StyleGAN과 StyleGAN Encoder를 활용하여 사용자의 얼굴 이미지로부터 2세대 합성 이미지를 생성하기 위한 시스템을 Fig. 1과 같이 설계하였다.

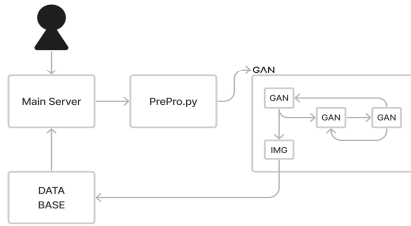


Fig. 1. System Architecture

#### 2. 제안된 모델의 구현

##### 2.1 데이터 세트 및 전처리

본 논문에서는 StyleGAN 모델의 구현을 위해, tensorflow 1.5 버전과 CUDA Toolkit 11.4 버전을 사용하였다. 또한 특정 알고리즘과의 호환성을 위해 CUDA Toolkit 10 버전의 내장 알고리즘 동적 링크 라이브러리 파일을 추가하였다. StyleGAN 모델은 'karras2019stylegan-ffhq-1024x1024.pkl'을 사용하여 구현하였으며, 이는 NVIDIA의 FFHQ 데이터 세트로 사전 훈련시켰다[3].

FFHQ는 이미지의 정확한 전처리와 얼굴 특징의 정렬이 가능하므로 효과적인 합성 이미지를 생성하기 위해 'shape\_predictor\_68\_face\_landmarks.dat.bz2' 파일을 사용하였다.

##### 2.2 모델 구현 및 학습

StyleGAN 모델의 학습 과정에서 더 정교한 특징의 추출 및 perceptual loss의 계산을 위해 사전 학습된 네트워크 가중치를 활용하였다[4]. 고차원 특징들은 모델이 학습하는 동안 이미지의 본질적인 양상을 효과적으로 포착하도록 하였다.

또한 perceptual loss는 생성된 이미지와 실제 이미지 간의 인식적 차이를 측정하는 지표이기 때문에 이미지의 시각적 품질을 개선하는데 적용하였다.

StyleGAN Encoder는 심층 신경망을 통해 구현되도록 하였으며, 고차원 특징과 얼굴의 미세한 구조를 포착하기 위한 다양한 계층을 포함시켰다. 사용자 이미지를 입력으로 받아, 해당 이미지의 잠재적 특성을 인코딩한 벡터를 출력한다[5].

Fig. 2는 생성된 2세의 얼굴 이미지를 보이고 있으며, 이미지의 품질 평가 및 최적화를 위해 VGG16 모델의 가중치를 이용한 perceptual matrix를 기반으로 생성된 이미지의 품질 평가를 진행하였다.

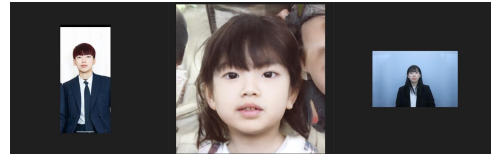


Fig. 2. The results of 2<sup>nd</sup> generation facial image generation

또한 품질 최적화 기법으로는 손실 함수를 세밀하게 조정하여, perceptual loss를 최소화하고 생성된 이미지의 시각적 품질이 향상되도록 하였다.

### IV. Conclusions

본 논문에서는 StyleGAN 기반의 자신과 닮은 2세의 가상 모습을 생성하는 웹 서비스를 개발하였다. FFHQ-data set을 통해 다양하고 현실적인 얼굴 이미지를 획득하였고, 나이와 성별을 나타내는 latent의 적용을 통하여 생성 이미지에서 명확한 2세 얼굴의 특징을 강조하고자 함에 성공하였다.

향후 연구에서는 한국인 얼굴 데이터 세트의 확보 및 효과적인 활용과 다양한 사용자 특성을 고려한 모델의 개선 연구가 필요하다.

### ACKNOWLEDGEMENT

"본 연구는 2023 제주산학융합원 Project Lab 지원사업의 연구결과로 수행되었음".

### REFERENCES

- [1] T. H. Cho, "Deep-learning for everyone," [KIIP] Seoul: Gilbut, pp. 1-308, 2017.
- [2] Tero Karras, and SamuliLaine, and Timo Aila, "A style-Based Generator Architecture for Generative Adversarial Networks," arxiv:1812.04948
- [3] NVlabs, "Flickr-Faces-HQ Dataset(FFHQ)," <https://github.com/NVlabs/ffhq-dataset>
- [4] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen and T. Aila, "Analyzing and Improving the Image Quality of StyleGAN", IEEE/CVF&CVPR, pp. 8107- 8116, June 2020. <https://doi.org/10.1109/CVPR42600.2020.00813>
- [5] Richardson, Elad, et al. "Encoding in style: a style gan encoder for image-to-image translation," In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. p. 2287- 2296, 2021.