

j.inst.korean.electr.electron.eng.

[Technical Paper]

- Protocol Design and Controller Implementation of Automotive LED Matrix Headlamp Control ● Changmin Lee, Wonchae Kim, Seonghyun Yang, and Seongsoo Lee (368)
- Deep Reinforcement Learning-Based C-V2X Distributed Congestion Control for Real-Time Vehicle Density Response ● Byeong Cheol Jeon, Woo Yoel Yang, Han-Shin Jo (379)
- Quasi-nonvolatile Memory Characteristics of Silicon Nanosheet Feedback Field-effect Transistors ● Seungho Ryou, Hyojoo Heob, Kyoungah Chob, Sangsig Kim (386)
- Regulation control of a dc motor by ϵ -PID controller based on the frequency response analysis ● Alhassan Dodo Adamou Soudeize, Ji-Sun Park, Ho-Lim Choi (391)
- Improvement of Current Path by Using Ferroelectric Material in 3D NAND Flash Memory ● Jihwan Lee, Jaewoo Lee, Myounggon Kang (399)
- Design of Dual-Band Patch Antenna Using Asymmetric Inset Feeding ● Dong-Kook Park (405)
- Performance Analysis of a UAV Energy Harvesting Relay Network in the Terahertz Band ● Yeongi Cho, Saifur Rahman Sabuj, Han-Shin Jo (411)
- Effects of Annealing on Electrical Characteristics of Double-Gated Silicon Nanosheet Feedback Field-Effect Transistors ● Hyojoo Heoa, Yunwoo Shina, Jaemin Sona, Seungho Ryub, Kyoungah Choa, Sangsig Kim (418)
- A study on scheme for activating active elements in RIS aided wireless communication system ● Jinsoo Bae, Seung-Geun Yoo, Hyoung-Kyu Song (425)
- Improved Estimation Method for the Capacitor Voltage in Modular Multilevel Converters Using Distributed Neural Network Observer ● Mehdi Syed Musadiq, Dong-Myung Lee (430)
- Indoor Environment Drone Detection through DBSCAN and Deep Learning ● Ha Tran Thi, Hien Pham The, Yun-Seok Mun, Ic-Pyo Hong (439)
- A High-efficiency Buck-boost Half-bridge Inverter for Single-phase Photovoltaic Generation ● Hyung-Min Ryu (450)
- Land Cover Classification of Satellite Image using SSResUnet Model ● Joohyung Kang, Minsung Kim, Seongjin Kim, Sooyeong Kwak (456)
- A study on age distortion reduction in facial expression image generation using StyleGAN Encoder ● JHee-Yeol Lee, Seung-Ho Lee (464)
- Fundamental Frequency Estimation of Voiced Speech Signals Based on the Inflection Point Detection ● Byeongwan Iem (472)
- Analysis of the Effect of Alternating Current Ripple on Electrical State of Health Degradation of 21700 Lithium-ion Battery ● Bongwoo Kwak (477)
- Automatic hand gesture area extraction and recognition technique using FMCW radar based point cloud and LSTM ● Seung-Tak Ra, Seung-Ho Lee (486)
- Analysis of the Image Processing Speed by Line-Memory Type ● Si-Yeon Han, Semin Jung, Bongsoon Kang (494)
- Design of SAE J2716 SENT Full Modes Controller ● Joonho Chung, Jaehyuk Cho, and Seongsoo Lee (501)
- A study on vertical alignment liquid crystal devices for electrically polarization controlled camera ● Na-Kyung Lee, Hyeon-Sik Ahn, Sung-Min Kim, Min-Sang Kim, Seungseo Park, Yoonseuk Choi (512)
- Diagnosis of Inter Turn Short Circuit in 3-Phase Induction Motors Using Applied Clarke Transformation ● Yeong-Jin Goh and Kyoung-Min Kim (518)
- Design and Implementation of Automotive Intrusion Detection System Using Ultra-Lightweight Convolutional Neural Network ● Myeongjin Lee, Hyungchul Im, Minseok Choi, Minjae Cha, and Seongsoo Lee (524)
- Data Preprocessing Method for Lightweight Automotive Intrusion Detection System ● Sangmin Park, Hyungchul Im, and Seongsoo Lee (531)
- Recent advances in few-shot learning for image domain: a survey ● Ho-Sik Seok (537)
- A Study on IAM-Based Personal Data Protection Techniques in BaaS ● Mi-Hui Kim, Myung-Joe Kang (548)
- The Impact of N-Non Implantation on Deep-Level Defects and Carrier Lifetime in 4H-SiC SBDs ● Myeong-cheol Shin, Geon-Hee Lee, Ye-Hwan Kang, Jong-Min Oh, Weon Ho Shin, San-Mo Koo (556)
- A Study on Improvement of Image Classification Accuracy Using Image-Text Pairs ● Mi-Hui Kim, Ju-Hyeok Lee (561)
- Breakdown Voltage and On-resistance Analysis of Partial-isolation LDMOS ● Sin-Wook Kim, Myoung-jin Lee (567)
- Additional Thermometer Code Locking Technique for Minimizing Quantization Error in Low Area Digital Controlled Oscillators ● Byeongseok Kang, Young-Sik Kim, Shinwoong Kim (573)
- Research on LTE D2D implementation and operation plan for PPDR communication ● Seong-Gyoon Park, Soong-Hwan Ro, Jongtae Ihm (579)
- A study on an optimal design of the high frequency transformer in LLC DC to DC resonant converter ● Jong-Hae Kim (587)
- A Study on fault diagnosis of DC transmission line using FPGA ● Tae-Hun Kim, Jun-Soo Che, Seung-Yun Lee, Byeong-Hyeon An, Jae-Deok Park, Tae-Sik Park (601)
- Implementation of ROS-Based Intelligent Unmanned Delivery Robot System ● Seong-Jin Kong, Won-Chang Lee (610)
- A Study on Service Quality Diagnosis Techniques for LTE/5G Network Backhaul ● Ji-Hyun Yoo (617)
- A Study on Design and Implementation of Scalable Angle Estimator Based on ESPRIT Algorithm ● Dohyun Lee, Byunghyun Kim, Jongwha Chong, Sungjin Lee, Kyeongyuk Min (624)
- Electrical characterization of 4H-SiC MOSFET with aluminum gate according to design parameters ● Seung-Hwan Baek, Jeong-Min Lee, U-yeol Seo, Yong-Seo Koo (630)
- Adaptive Fast Calibration Method for Active Phased Array Antennas using PPO Algorithm ● Sunge Lee, Kisk Byun, Hong-Jib, Yoon (636)
- Analysis of Passing Word Line Induced Leakage of BCAT Structure in DRAM ● Su Yeon Kim, Dong Yeong Kim, Je Won Park, Shin Wook Kim, Chae Hyuk Lim So won Kim, Hyeona Seo, Hye Rin Lee, Ju Won Kim, Jeong Hyeon Yun, Young-Woo Lee, Hyoung-Jin Joe, Myoung Jin Lee (644)
- Channel Modeling and RF Performance Verification in mmWave Bands Based on NS-3 ● Seung-Min Lee, Jun-Seok Seo, Hong-Je Jang, Myung-Ryul Choi (650)
- Re-estimation of PV hosting capacity by improving parameters for voltage controls of the smart inverter ● Juhyeon Kim, Gihwan Yoon, Yoondong Sung, Hak-Geun Jeong, Jongbok Baek, Moses Kang (657)
- FlexRay: Protocol, Time Hierarchy, Message Frame, Communication Controller, and Conformance Test ● Seokjun Hahn, Sua Shin, Naeun Park, Chan Park, Daegi Lee, and Seongsoo Lee (668)

[Short Papers]

- Clustering Methods for Cluster Uniformity in Wireless Sensor Networkse ● Joong-Ho Lee (679)
- Design of a Delayed Dual-Core Lock-Step Processor with Automatic Recovery in Soft Errors ● Juho Kim, Seonghyun Yang, and Seongsoo Lee (683)



전기전자학회 논문지

j.inst.korean.electr.electron.eng.

제27권 제4호 2023년 12월

[논문]

- 차량용 LED 매트릭스 헤드램프 제어를 위한 LED 제어 프로토콜 설계 및 제어기 구현 ● 이창민, 김원채, 양성현, 이성수 (368)
- 실시간 차량 밀도에 대응하는 심층강화학습 기반 C-V2X 분산혼잡제어 ● 전병철, 양우열, 조한신 (379)
- 실리콘 나노시트 피드백 전계효과 트랜지스터의 준비회발성 메모리 특성 연구 ● 류승호, 허효주, 조경아, 김상식 (386)
- Regulation control of a dc motor by ϵ -PID controller based on the frequency response analysis ● Alhassan Dodo Adamou Soudeize, Ji-Sun Park, Ho-Lim Cho (391)
- 3D NAND Flash Memory에 Ferroelectric Material을 사용한 Current Path 개선 ● 이지환, 이재우, 강명곤 (399)
- 비대칭 인셋 급전을 사용한 이중 대역 패치 안테나 설계 ● 박동국 (405)
- 테라헤르츠 대역 무인비행체 에너지 수확 릴레이 네트워크 성능분석 ● 조연기, Saifur Rahman Sabuj, 조한신 (411)
- 더블게이트 실리콘 나노시트 피드백 전계효과 트랜지스터의 전기적 특성에 미치는 열처리 효과 ● 허효주, 신언우, 손재민, 류승호, 조경아, 김상식 (418)
- RIS를 활용한 무선 통신 시스템에서 능동 반사 소자를 활성화하는 기법에 관한 연구 ● 배진수, 유승근, 송형규 (425)
- Improved Estimation Method for the Capacitor Voltage in Modular Multilevel Converters Using Distributed Neural Network Observer ● Mehdi Syed Musadiq, Dong-Myung Lee (430)
- Indoor Environment Drone Detection through DBSCAN and Deep Learning ● Ha Tran Thi, Hien Pham The, Yun-Seok Mun, Ic-Pyo Hong (439)
- 단상 태양광 발전용 고효율 벡부스트 하프브리지 인버터 ● 류형민 (450)
- SSResUnet 모델을 이용한 위성 영상 토지피복분류 ● 강주형, 김민성, 김성진, 곽수영 (456)
- StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡 감소에 대한 연구 ● 이희열, 이승호 (464)
- 변곡점 검출에 기반한 음성인식의 기본 주파수 추정 ● 임병관 (472)
- 교류 리플이 21700 리튬 이온 배터리의 전기적 건강 상태 열화에 미치는 영향 분석 ● 광봉우 (477)
- FMCW 레이더 기반의 포인트 클라우드로 LSTM을 이용한 자동 핸드 제스처 영역 추출 및 인식 기법 ● 라승탁, 이승호 (486)
- 라인메모리 유형에 따른 이미지 처리 속도의 분석 ● 한시연, 정세민, 강봉순 (494)
- SAE J2716 풀 모드 SENT 컨트롤러의 설계 ● 정준호, 조재혁, 이성수 (501)
- 전기적 편광 조절형 카메라를 위한 수직 배향형 액정 소자 연구 ● 이나경, 안현식, 김성민, 김민상, 박승서, 최윤석 (512)
- Clarke 변환을 응용한 3상 유도전동기의 Inter Turn Short Circuit 진단 ● 고영진, 김경민 (518)
- 초경량 Convolutional Neural Network를 이용한 차량용 Intrusion Detection System의 설계 및 구현 ● 이명진, 임형철, 최민석, 차민재, 이성수 (524)
- 차량용 경량화 침입 탐지 시스템을 위한 데이터 전처리 기법 ● 박상민, 임형철, 이성수 (531)
- 이미지 분석을 위한 퓨사 학습의 최신 연구동향 ● 석호식 (537)
- BaaS에서 IAM을 이용한 개인정보 보호 기법에 관한 연구 ● 김미희, 강명조 (548)
- N-이온주입이 4H-SiC SBDs의 깊은 준위 결함 및 소수 캐리어 수명에 미치는 영향 ● 신명철, 이건희, 강예환, 오종민, 신원호, 구상모 (556)
- 이미지-텍스트 쌍을 활용한 이미지 분류 정확도 향상에 관한 연구 ● 김미희, 이주혁 (561)
- Partial-isolation LDMOS의 항복전압과 온저항 분석 ● 김신욱, 이명진 (567)
- 저면적 디지털 제어 발전기의 양자화 에러 최소화를 위한 추가 서모미터 코드 잠금 기법 ● 강병석, 김영식, 김신웅 (573)
- 재난통신을 위한 LTE D2D 구현 및 운용 방안 연구 ● 박성균, 노승환, 임종태 (579)
- LLC DC to DC 공진 컨버터의 고주파 변압기 최적화 설계에 관한 연구 ● 김중해 (587)
- FPGA를 활용한 DC계통 고장진단에 관한 연구 ● 김태훈, 채준수, 이승윤, 안병현, 박재덕, 박태식 (601)
- ROS 기반 지능형 무인 배송 로봇 시스템의 구현 ● 공성진, 이원창 (610)
- LTE/5G 네트워크 백홀(Backhaul)의 서비스 품질진단 기법에 관한 연구 ● 유지현 (617)
- ESPRIT 알고리즘 기반 재구성 가능한 각도 추정기 설계에 관한 연구 ● 이도현, 김병현, 정경화, 이성진, 민경욱 (624)
- Aluminium Gate를 적용한 4H-SiC MOSFET의 Design parameter에 따른 전기적 특성 분석 ● 백승환, 이정민, 서우열, 구용서 (630)
- PPO 알고리즘을 이용한 능동위상배열안테나 적용형 고속 보정 방법 ● 이성계, 변기식, 윤홍집 (636)
- BCAT구조 DRAM의 패시워드 라인 유도 누설전류 분석 ● 김수연, 김동영, 박재원, 김신욱, 임재혁, 김소원, 서현아, 이주원, 이혜란, 윤정현, 이영우, 조형진, 이명진 (644)
- NS-3 기반의 mmWave 대역 채널 모델링 및 RF 성능 검증 ● 이승민, 서준석, 장홍제, 최명철 (650)
- 스마트인버터 전압제어의 파라미터 개선을 통한 PV hosting capacity 재추정 방법 ● 김주현, 윤기환, 성윤동, 정학근, 백종복, 강모세 (657)
- FlexRay: 프로토콜, 시간 계층, 메시지 프레임, 커뮤니케이션 컨트롤러, 적합성 시험 ● 한석준, 신수아, 박나은, 박찬, 이대기, 이성수 (668)

[단편]

- 무선센서 네트워크에서 클러스터 균일화를 위한 클러스터링 방법 ● 이종호 (679)
- 소프트웨어 발생 시 자동 복구하는 이중 코어 지연 락스텝 프로세서의 설계발 ● 김주호, 양성현, 이성수 (683)

전기전자학회 논문지

제27권 제4호 2023년 12월

사단법인 한국전기전자학회



StyleGAN Encoder를 활용한 표정 이미지

생성에서의 연령 왜곡 감소에 대한 연구

A study on age distortion reduction in facial expression

image generation using StyleGAN Encoder

이 희 열*, 이승 호*

Hee-Yeol Lee*, Seung-Ho Lee*

Abstract

In this paper, we propose a method to reduce age distortion in facial expression image generation using StyleGAN Encoder. The facial expression image generation process first creates a face image using StyleGAN Encoder, and changes the expression by applying the learned boundary to the latent vector using SVM. However, when learning the boundary of a smiling expression, age distortion occurs due to changes in facial expression. The smile boundary created in SVM learning for smiling expressions includes wrinkles caused by changes in facial expressions as learning elements, and it is determined that age characteristics were also learned. To solve this problem, the proposed method calculates the correlation coefficient between the smile boundary and the age boundary and uses this to introduce a method of adjusting the age boundary at the smile boundary in proportion to the correlation coefficient. To confirm the effectiveness of the proposed method, the results of an experiment using the FFHQ dataset, a publicly available standard face dataset, and measuring the FID score are as follows. In the smile image, compared to the existing method, the FID score of the smile image generated by the ground truth and the proposed method was improved by about 0.46. In addition, compared to the existing method in the smile image, the FID score of the image generated by StyleGAN Encoder and the smile image generated by the proposed method improved by about 1.031. In non-smile images, compared to the existing method, the FID score of the non-smile image generated by the ground truth and the method proposed in this paper was improved by about 2.25. In addition, compared to the existing method in non-smile images, it was confirmed that the FID score of the image generated by StyleGAN Encoder and the non-smile image generated by the proposed method improved by about 1.908. Meanwhile, as a result of estimating the age of each generated facial expression image and measuring the estimated age and MSE of the image generated with StyleGAN Encoder, compared to the existing method, the proposed method has an average age of about 1.5 in smile images and about 1.63 in non-smile images. Performance was improved, proving the effectiveness of the proposed method.

* Dept. Electronic Engineering, Hanbat National University

★ Corresponding author

E-mail : shlee@cad.hanbat.ac.kr, Tel : +82-42-821-1137

※ Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT)(No. NRF-2022R1F1A1066371)

This results was supported by "Regional Innovation Strategy (RIS)" through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(MOE)(2021RIS-004)

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ICAN (ICT Challenge and Advanced Network of HRD) program (IITP-2022-RS-2022-00156212) supervised by the IITP (Institute of Information & Communications Technology Planning & Evaluation).

Manuscript received Nov. 24, 2023; revised Dec. 11, 2023; accepted Dec. 15, 2023.

This is an Open-Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License(<http://creativecommons.org/licenses/by-nc/3.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

요 약

본 논문에서는 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡을 감소시키는 방법을 제안한다. 표정 이미지 생성 과정은 StyleGAN Encoder를 사용하여 얼굴 이미지를 생성하고, SVM을 이용하여 학습된 boundary를 잠재 벡터에 적용하여 표정을 변화시킨다. 그러나 웃는 표정의 boundary를 학습할 때 표정 변화에 따른 연령 왜곡이 발생한다. 웃는 표정에 대한 SVM 학습에서 생성된 smile boundary는 표정 변화로 인해 생긴 주름이 학습 요소로 포함되어 있으며 연령에 대한 특성도 함께 학습된 것으로 판단한다. 이를 해결하기 위해, 제안된 방법에서는 smile boundary와 age boundary의 상관계수를 계산하고, 이를 이용하여 smile boundary에서 age boundary를 상관계수에 비례하여 조절하는 방식을 도입한다. 제안된 방법의 효과를 확인하기 위해 공개된 표준 얼굴 데이터셋인 FFHQ 데이터셋을 사용하고 FID score를 측정하여 실험한 결과는 다음과 같다. Smile 이미지에서는 기존 방법에 비하여, Ground Truth와 제안된 방법으로 생성된 smile 이미지의 FID score가 약 0.46 향상되었다. 또한, Smile 이미지에서 기존 방법에 비하여, StyleGAN Encoder로 생성된 이미지와 제안된 방법으로 생성된 smile 이미지의 FID score가 약 1.031 향상되었다. Non-smile 이미지에서는 기존 방법에 비하여, Ground Truth와 본 논문에서 제안된 방법으로 생성된 non-smile 이미지의 FID score가 약 2.25 향상되었다. 또한, Non-smile 이미지에서 기존 방법에 비하여, StyleGAN Encoder로 생성된 이미지와 제안된 방법으로 생성된 non-smile 이미지의 FID score가 약 1.908 향상됨을 확인하였다. 한편, 각 생성된 표정 이미지의 연령을 추정하여 StyleGAN Encoder로 생성된 이미지의 추정된 연령과 MSE를 측정한 결과, 기존방법 대비 제안하는 방법이 smile 이미지에서 약 1.5, non-smile 이미지에서 약 1.63의 성능 향상되어 제안한 방법에 대한 성능의 효율성이 입증되었다.

Key words : DeepLearning, Generative Adversarial Networks, Encoder, Image generation, Correlation coefficient

I. 서론

최근 생성형 딥러닝 모델 기술의 발전과 함께 이미지 생성 및 특성 편집에 대한 연구는 더욱 중요해지고 있다. 이미지 생성과 특성 편집은 컴퓨터 비전 및 인공지능 분야에서 중요한 주제로써 얼굴 이미지의 표정 및 연령 등의 수정 등이 가능해졌다. 이러한 분야의 작업은 특히 영화산업 및 영상매체 산업에서 각광받고 있다. 해당 산업은 과거 영상 편집자의 수작업으로 영상을 제작하여 왔으나 GAN(Generative Adversarial Networks)의 발전을 시작으로 현재는 생성형 딥러닝 모델을 통해 영상 제작 속도 및 품질을 크게 향상시켜왔다. 특히 StyleGAN을 이용한 이미지 생성 기술은 사진과 유사한 고품질의 이미지를 생성하며 해당 분야의 큰 영향을 주었다. 이후 이미지의 잠재 벡터를 조작하여 특정 인물의 이미지에 표정 및 연령 조작이 가능해져 고연령 배우의 연기 영상을 젊은 시절의 영상으로 변환하는 등의 작업이 가능해졌다. 그러나 표정 변화에서 웃는 표정의 이미지에 주름이 더해짐으로써 학습 결과에서 웃는 표정일 때 더 많은 나이의 이미지가 생성되는 현상을 확인하였다. 이러한 현상은 생성된 이미지에서 표정과 나이 간의 상호작용에 대한 측면을 제시한다.

본 논문에서는 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡을 확인하고 표정 변화와 연령 변화의 연관성을 분석한다. 또한, 상관계수를 적용하여 표정 이미지 생성에서의 연령 왜곡 개선 방안을 제시

하고 이를 검증하기 위한 실험을 수행하여 제안한 방법의 성능을 입증한다.

II. 본론

1. 전체 개요도

본 논문에서 제안하는 표정 이미지 생성에서의 연령 왜곡 개선 방법의 전체 개요도는 그림 1과 같다.

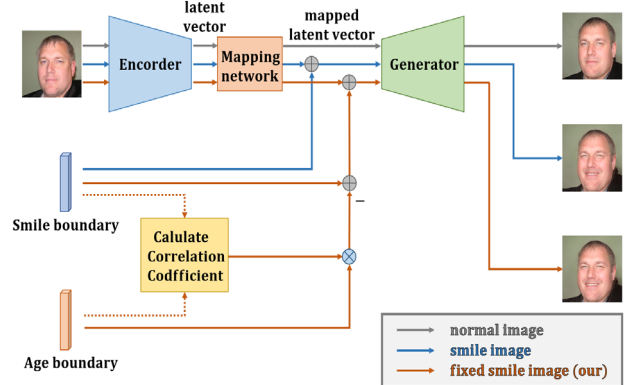


Fig. 1. A complete overview of the proposed method. 그림 1. 제안하는 방법의 전체 개요도

2. StyleGAN Encoder를 이용한 얼굴 이미지 생성

StyleGAN은 잠재 벡터를 입력받아 이미지를 생성하는 GAN이다[1]. 여기서 잠재 벡터는 Mapping Network를 통해 학습된 스타일 정보로 인코딩되어 Generator에 입력되고, 입력한 잠재 벡터에 해당하는 이미지가 생성된

다[2]. 이렇게 생성된 이미지는 다양한 용도에 활용되며, 대표적으로 데이터 증강에 효과적으로 사용된다[3]. 그림 2는 StyleGAN의 구성도이다.

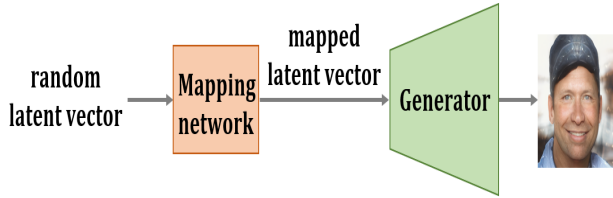


Fig. 2. Structure of StyleGAN.
그림 2. StyleGAN의 구성도

Mapping Network는 잠재 벡터의 각 차원을 다양한 스타일 특성으로 매핑한다. 이러한 스타일 특성은 얼굴 이미지의 다양한 시각적 특징을 제어한다. 예를 들어 머리카락 스타일, 미소의 정도, 눈의 크기 등이 여기에 해당한다. Generator의 각 레이어에서 스타일이 조절되며, 이러한 과정을 통해 고품질의 얼굴 이미지를 생성된다. StyleGAN의 Generator는 다계층 구조를 가지고 있으며, Mapping Network를 통해 맵핑된 잠재 벡터를 입력으로 받아 이미지를 생성한다. 이때 이미지의 다양성을 높이기 위해 각 레이어에서는 노이즈가 입력되며 이는 스타일의 변화를 증가시키고 디테일한 특성을 부여한다. 높은 차원의 스타일 특성은 각 레이어에서 일정 부분에 전파되어 전체적인 스타일을 형성한다. 또한 AdaIN이라는 정규화 방법이 각 레이어에서 사용된다. 이 방법은 입력된 스타일 특성을 이미지에 적응적으로 적용하여 스타일의 다양한 변화를 가능하게 한다. 일반적인 StyleGAN은 난수로부터 생성된 잠재 벡터를 사용하여 이미지를 생성하므로 무작위하게 학습된 형태의 이미지가 생성된다. 특정한 이미지를 생성하기 위해서는 Encoder를 사용하여 한다. 그림 3은 StyleGAN Encoder의 구성도이다.

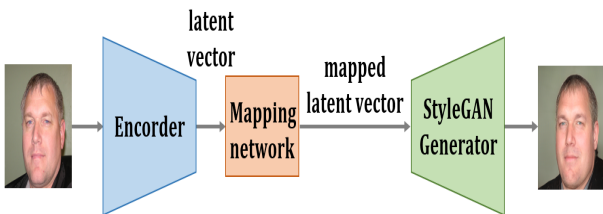


Fig. 3. Structure of StyleGAN Encoder.
그림 3. StyleGAN Encoder의 구성도

StyleGAN Encoder는 난수로부터 생성된 잠재 벡터를 사용하지 않고, 입력받은 이미지로부터 잠재 벡터를 추정한다[4]. 이 과정은 Generator의 기능을 반대로 수

행하는 것과 같다. 이렇게 추정된 잠재 벡터는 원하는 특성을 강조하거나 억제하기 위해 수정 또는 조작될 수 있다. 추출된 잠재 벡터는 Generator에 입력되고, Generator는 초기에 입력받은 이미지와 일치되는 이미지를 출력한다.

3. 특성 boundary를 이용한 얼굴 표정 이미지를 생성

StyleGAN Encoder를 통해 추출한 잠재 벡터는 원하는 특성을 강조하거나 억제하기 위해 InterFaceGAN을 사용한 추가적인 학습이 필요하다[5]. 먼저, 추출한 잠재 벡터와 특성값을 데이터셋으로 구축한다. 이후, 구축된 데이터셋은 특성값에 따라 Positive 클래스와 Negative 클래스로 나뉘어진다. 여기서 Positive 클래스는 특성값이 있는 데이터로 구성되고, Negative 클래스는 특성값이 없는 데이터로 구성된다. Positive와 Negative로 나뉘어진 데이터는 SVM(Support Vector Machine)의 이진 분류 학습을 통해 두 클래스로 나뉘어진다. SVM은 분류 및 회귀 문제에 활용되는 지도 학습 알고리즘으로, 데이터를 고차원 공간에 매핑하여 서로 다른 클래스를 나누는 초평면을 계산한다. 이 초평면은 각 클래스 간의 최대 마진을 갖도록 학습되며, 분류 문제에서 클래스 간의 경계를 형성한다. 분류 학습이 완료되면 학습된 특성 boundary 데이터가 생성된다. 이후, Generator에 입력되는 잠재 벡터에 해당 boundary 데이터를 추가함으로써, 기본 이미지에서 특성값이 증가한 이미지가 출력된다. 이 과정을 통해 이미지의 원하는 특성을 더욱 강조하거나 조절할 수 있다. 그림 4는 smile boundary를 적용 smile 이미지 생성 구성도이다.

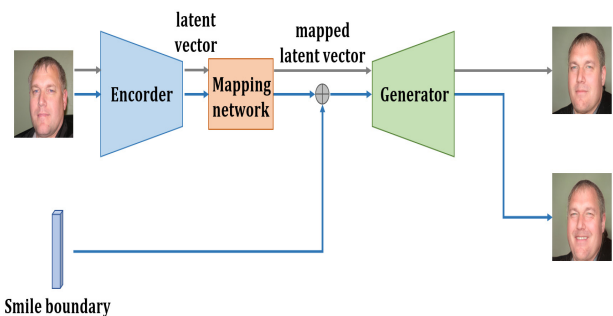


Fig. 4. Structure of creating a smile image by applying a smile boundary.

그림 4. smile boundary를 적용 smile 이미지 생성 구성도

4. 상관계수를 적용하여 표정 이미지 생성에서의 연령 왜곡 개선

그림 5는 smile boundary를 적용하여 추출한 이미지이다.

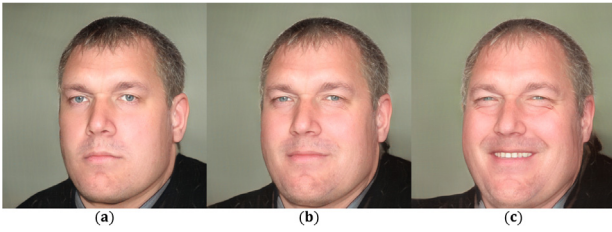


Fig. 5. Image extracted by applying smile boundary.
그림 5. smile boundary를 적용하여 추출한 이미지

(b)는 기본 출력 이미지이고, (a)는 boundary 음의 값을 주어 smile 특성을 줄였고, (c)는 양의 값을 주어 smile 특성을 높였다. 해당 이미지를 관찰하였을 때 상대적으로 (a)는 (b)보다 나이가 적게, (c)는 (b)보다 나이가 많게 출력되었다. 이러한 현상은 웃음 때 발생하는 얼굴의 주름으로 인해 age 특성이 smile 특성의 학습에 영향을 주었다고 판단된다. 따라서, 상관계수를 적용하여 표정 이미지 생성에서의 연령 왜곡을 개선하기 위하여 smile boundary와 age boundary의 상관계수를 이용하여 값을 조정한다. 상관계수는 두 변수 간 상관관계의 정도를 나타내는 수치로, n 차원의 두 변수 x , y 가 주어질 때 상관계수 r 은 식(1)과 같다.

$$r = \frac{S_{xy}}{\sqrt{S_{xx}} \cdot \sqrt{S_{yy}}} \quad (1)$$

단, \bar{x} , \bar{y} 는 각각 x 와 y 의 평균을 나타내고, S_{xx} , S_{yy} 는 각각 x 와 y 의 분산, S_{xy} 는 x 와 y 의 공분산이다. 식은 다음과 같다.

$$\bar{x} = \frac{1}{n} \sum (x_i) \quad (2)$$

$$\bar{y} = \frac{1}{n} \sum (y_i) \quad (3)$$

$$S_{xx} = \sum (x_i - \bar{x})^2 \quad (4)$$

$$S_{yy} = \sum (y_i - \bar{y})^2 \quad (5)$$

$$S_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) \quad (6)$$

상관계수 r 은 -1과 1 사이의 값이다. -1은 완벽한 음의 상관관계를, 1은 완벽한 양의 상관관계를 나타낸다. 0은 두 변수가 완벽하게 독립(수직)인 것을 의미한다. 높은 상관관계는 두 변수 사이의 종속성이 높아 한 변수의 변화가 다른 변수에 영향을 미친다는 것을 나타낸다.

상관계수 r 에 비례한 종속성을 활용하여 두 boundary의 상관계수를 계산하고, 식(7)을 사용하여 잠재 벡터의 특성을 조작한다.

$$\overrightarrow{latent'} = \overrightarrow{latent} + \alpha (\overrightarrow{BDRY}_{smile} - r \cdot \overrightarrow{BDRY}_{age}) \quad (7)$$

\overrightarrow{latent} 는 입력 이미지에서 추출된 잠재 벡터이며, $\overrightarrow{BDRY}_{smile}$ 와 $\overrightarrow{BDRY}_{age}$ 는 각각 smile boundary와 age boundary를 나타내고 α 는 boundary를 적용할 스케일 크기를 나타낸다. 위 수식을 통해 smile boundary로 인한 age 증가 현상을 완화 할 수 있다. 그림 6은 상관계수를 적용한 개선된 smile 이미지 생성 구성도이다.

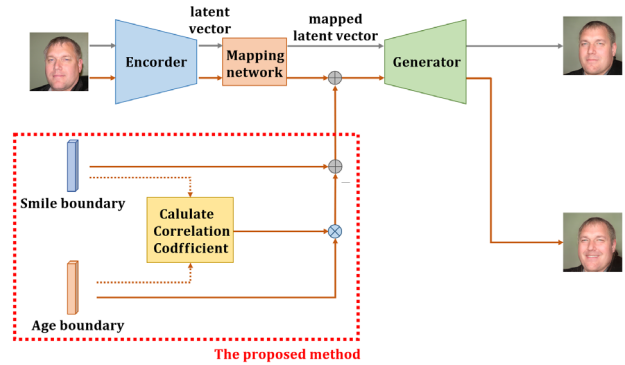


Fig. 6. structure of improved smile image generation using correlation coefficient.

그림 6. 상관계수를 적용한 개선된 smile 이미지 생성 구성도

그림 7은 상관계수를 적용한 개선된 smile 이미지도이다. (a)는 기존의 smile boundary를 적용하여 생성한 이미지이고, (b)는 제안하는 방법을 적용하여 개선된 이미지이다. (c)는 두 이미지의 차이를 강조하기 위해 차 연산을 수행 후 노출을 높인 이미지이다. (a)와 (b)를 자세히 비교하면, (a)가 (b)에 비해 노화에 의한 흰머리와 더 깊은 주름 및 더 거친 피부결 등을 확인할 수 있다. 이 차이는 (c)를 통해 명확하게 확인할 수 있다. 특히, 눈가의 주름이 크게 줄었음을 (c)를 통해 확인할 수 있다.

4. 실험

가. 실험 환경

실험에 사용된 운영체제 및 하드웨어는 Ubuntu 20.04.2 LTS 운영체제를 기반으로 Intel(R) Xeon(R) Silver 4210 CPU @ 2.20GHz, RAM 128GB, NVIDIA RTX A5000 (VRAM 24GB) GPU로 구성되어 있다. 개발도구는 Visual Studio Code와 Pytorch 1.8.0, CUDA 11.1, cuDNN 8.0.5 라이브러리를 사용하였다. 본 논문에서 제안하는 방법의 검증에 사용된 데이터셋으로 FFHQ(Flickr-Face-HQ) 데이터셋을 사용하였다.



Fig. 7. Improved smile image using correlation coefficient.
그림 7. 상관계수를 적용한 개선된 smile 이미지

FFHQ 데이터셋은 GAN의 벤치마크로 만들어진 고품질 얼굴이미지 데이터셋이다. 해당 데이터셋은 1024×1024

해상도의 PNG 이미지 70000개로 구성되어 있으며 그림 8과 같이 연령, 포즈, 표정, 민족, 배경 등의 다양한 환경이 포함되어 있다. 또한 안경, 모자 등의 액세서리를 포함한 이미지도 상당량 분포되어있다.



Fig. 8. FFHQ dataset.
그림 8. FFHQ 데이터셋

실험에 사용한 StyleGAN Encoder로 ReStyle-pSp를 사용하였으며 내부의 generator 모델로 stylegan3를 사용하였다[6][7]. FFHQ 데이터셋의 pretrain되어 있는 모델을 로드하여 사용하였다[6]. 또한 SVM은 scikit-learn 라이브러리를 이용하여 boundary 학습에 사용되었다.

나. 실험 결과

본 실험에 사용된 지표는 FID(Fr chet Inception Distance) score로 생성된 영상의 품질비교에 사용되며 원본 데이터셋과 생성된 데이터셋과의 분포 거리를 나타낸다[8]. 원본 데이터와 유사할수록 거리가 가까워 낮은 점수가 출력된다. 해당 점수는 Inception 모델을 사용하여 중간 layer에서 feature를 추출한다. 추출한 feature에서 평균과 공분산을 이용하여 거리를 계산한다. 두 데이터셋 X, Y 간의 FID 수식은 식(8)과 같다.

$$FID = \frac{1}{2} \left(\|\mu_X - \mu_Y\|_2^2 + \text{Tr}(\Sigma_X + \Sigma_Y - 2\sqrt{(\Sigma_X \Sigma_Y)}) \right) \quad (8)$$

FID score는 두 데이터셋간의 거리를 나타내므로 제

안하는 방법을 적용한 데이터셋이 원본과 비교했을 때 연령의 왜곡을 줄이는 효과를 입증하기 위해 채택하였다. 실험 방법은 그림 9와 같이 Ground truth와 InterfaceGAN으로 생성된 smile 이미지의 FID score를 계산하고, Ground truth와 제안하는 방법으로 생성된 smile 이미지의 FID score를 계산하여 비교한다. 또한, StyleGAN Encoder로 생성된 결과와 InterfaceGAN으로 생성된 smile 이미지의 FID score를 계산하고, StyleGAN Encoder로 생성된 결과와 제안하는 방법으로 생성된 smile 이미지의 FID score를 계산하여 비교한다. smile 특성을 음의 값으로 입력하여 생성된 non-smile 이미지에 대해서도 위의 방법으로 실험한다. 또한, 연령 왜곡의 개선을 확인하기 위해 연령 추정 모델을 사용해 StyleGAN Encoder로 생성된 결과와 InterfaceGAN으로 생성된 smile 이미지와 non-smile 이미지, 제안하는 방법으로 생성된 smile 이미지와 non-smile 이미지의 연령 오차를 MSE로 측정하였다. 연령 추정은 Deep EXpectation 네트워크의 pretrained 모델을 로드해 사용하였다[9].

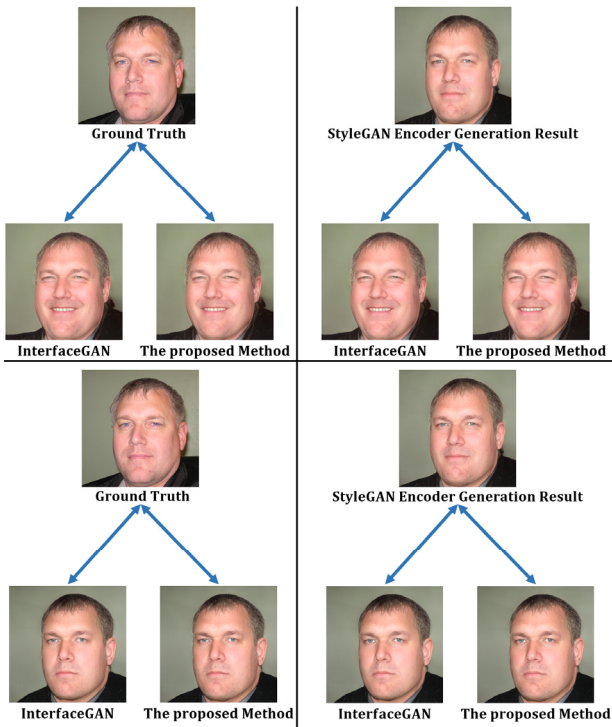


Fig. 9. 실험 방법.
그림 9. Experimental method

실험 결과, 표 1과 같이 Ground Truth에 대한 smile 이미지들의 FID score는 InterfaceGAN 방법이 약 36.475로 측정되었고, 제안하는 방법이 약 36.015로 측정되었다. 따라서, 본 논문에서 제안하는 방법이 기존

InterfaceGAN의 방법과 비교하여 FID score가 약 0.46 향상되었음을 확인하였다.

Table 1. FID scores of smile images for Ground Truth.
표 1. Ground Truth에 대한 smile 이미지들의 FID score

smile	FID score ↓
	Ground Truth
InterfaceGAN[5]	36.475
The proposed method	36.015

표 2에서는 StyleGAN Encoder로 생성된 결과에 대한 smile 이미지들의 FID score는 InterfaceGAN 방법이 약 5.187로 측정되었고, 제안하는 방법이 약 4.156로 측정되었다. 따라서, 본 논문에서 제안하는 방법이 기존 InterfaceGAN의 방법과 비교하여 FID score가 약 1.031 향상되었음을 확인하였다.

Table 2. FID scores of smile images for results generated with StyleGAN Encoder.
표 2. StyleGAN Encoder로 생성된 결과에 대한 smile 이미지들의 FID score

smile	FID score ↓
	StyleGAN Encoder Results
InterfaceGAN[5]	5.187
The proposed method	4.156

표 3에서는 Ground Truth에 대한 non-smile 이미지들의 FID score는 InterfaceGAN 방법이 약 36.269로 측정되었고, 제안하는 방법이 약 34.018로 측정되었다. 따라서, 본 논문에서 제안하는 방법이 기존 InterfaceGAN의 방법과 비교하여 FID score가 약 2.25 향상되었음을 확인하였다.

Table 3. FID scores of non-smile images for Ground Truth.
표 3. Ground Truth에 대한 non-smile 이미지들의 FID score

non-smile	FID score ↓
	Ground Truth
InterfaceGAN[5]	36.269
The proposed method	34.018

표 4에서는 StyleGAN Encoder로 생성된 결과에 대한 non-smile 이미지들의 FID score는 InterfaceGAN

방법이 약 5.968로 측정되었고, 제안하는 방법이 약 4.06로 측정되었다. 따라서, 본 논문에서 제안하는 방법이 기존 InterfaceGAN의 방법과 비교하여 FID score가 약 1.908 향상되었음을 확인하였다.

Table 4. FID scores of non-smile images for results generated with StyleGAN Encoder.

표 4. StyleGAN Encoder로 생성된 결과에 대한 non-smile 이미지들의 FID score

non-smile	FID score ↓
	StyleGAN Encoder Results
InterfaceGAN[5]	5.968
The proposed method	4.060

표 5는 연령 추정 모델을 사용해 StyleGAN Encoder로 생성된 결과와 각 방법으로 생성된 이미지의 연령 MSE 측정 결과이다. InterfaceGAN으로 생성된 smile 이미지는 약 3.529, InterfaceGAN으로 생성된 non-smile 이미지는 약 3.549, 제안하는 방법으로 생성된 smile 이미지는 약 2.029, 제안하는 방법으로 생성된 non-smile 이미지는 약 1.918의 MSE 측정 결과를 보였다. 따라서, 본 논문에서 제안하는 방법이 기존 InterfaceGAN의 방법과 비교하여 smile 이미지에서 약 1.5, non-smile 이미지에서 약 1.63의 성능 향상을 확인하였다.

Table 5. Age estimation MSE for each generated image.

표 5. 생성된 이미지들의 연령 추정에 대한 MSE

age	MSE ↓	
	smile	non-smile
InterfaceGAN[5]	3.529	3.549
The proposed method	2.029	1.918

FID score는 이미지의 분포 및 품질이 원본과 얼마나 유사한지를 나타내므로, 실험 결과에서 연령 왜곡을 원본 이미지의 특징을 유지하면서 복원을 되었는지를 확인할 수 있다. 연령 추정에 대한 MSE는 연령을 직접적으로 추정하여 왜곡이 얼마나 발생했는지 FID보다 더 명확하게 파악이 가능하며 또한 연령 특성에 대해 얼마나 복원을 했는지 알수있는 지표이다. 따라서, 실험 결과를 통해 제안하는 방법이 연령 왜곡을 감소시키는데 효과적임을 입증하였다.

III. 결론

본 논문에서는 StyleGAN Encoder를 활용한 표정 이미지 생성에서의 연령 왜곡을 감소시키는 방법을 제안하였다. 표정 이미지 생성 과정은 StyleGAN Encoder를 사용하여 얼굴 이미지를 생성하고, SVM을 이용하여 학습된 boundary를 잠재 벡터에 적용하여 표정을 변화시킨다. 그러나 웃는 표정의 boundary를 학습할 때 표정 변화에 따른 연령 왜곡이 발생하였다. 웃는 표정에 대한 SVM 학습에서 생성된 smile boundary는 표정 변화로 인해 생긴 주름이 학습 요소로 포함되어 있으며 연령에 대한 특성도 함께 학습된 것으로 판단하였다. 이를 해결하기 위해, 제안된 방법에서는 smile boundary와 age boundary의 상관계수를 계산하고, 이를 이용하여 smile boundary에서 age boundary를 상관계수에 비례하여 조절하는 방식을 도입하였다. 제안된 방법의 효과를 확인하기 위해 공개된 표준 얼굴 데이터셋인 FFHQ 데이터셋을 사용하고 FID score를 측정하여 실험한 결과는 다음과 같다. Smile 이미지에서는 기존 방법에 비하여, Ground Truth와 제안된 방법으로 생성된 smile 이미지의 FID score가 약 0.46 향상되었다. 또한, Smile 이미지에서 기존 방법에 비하여, StyleGAN Encoder로 생성된 이미지와 제안된 방법으로 생성된 smile 이미지의 FID score가 약 1.031 향상되었다. Non-smile 이미지에서는 기존 방법에 비하여, Ground Truth와 본 논문에서 제안된 방법으로 생성된 non-smile 이미지의 FID score가 약 2.25 향상되었다. 또한, Non-smile 이미지에서 기존 방법에 비하여, StyleGAN Encoder로 생성된 이미지와 제안된 방법으로 생성된 non-smile 이미지의 FID score가 약 1.908 향상됨을 확인하였다. 한편, 각 생성된 표정 이미지의 연령을 추정하여 StyleGAN Encoder로 생성된 이미지의 추정된 연령과 MSE를 측정한 결과, 기존방법 대비 제안하는 방법이 smile 이미지에서 약 1.5, non-smile 이미지에서 약 1.63의 성능 향상을 이루어 제안한 방법에 대한 성능의 효율성이 입증되었다. 향후 연구 방향으로 다양한 특징 이미지 생성을 위한 다중 특성의 관계성에 대한 연구가 필요하다고 사료된다.

References

[1] Karras, Tero, Samuli Laine, and Timo Aila, "A style-based generator architecture for generative

adversarial networks,” *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.

DOI: 10.48550/arXiv.1812.04948

[2] J. Y. Kim, S. H. Lee, “3D Point Cloud Reconstruction Technique from 2D Image Using Efficient Feature Map Extraction Network,” *j.inst.Korean.electr.electron.eng*, vol.26, no.3, pp.408-415, 2022.

[3] J. H. Shim, J. E. Lee, E. J. Hwang, “A Scheme for Preventing Data Augmentation Leaks in GAN-based Models Using Auxiliary Classifier,” *j.inst.Korean.electr.electron.eng*, vol.26, no.2, pp.176-185, 2022.

[4] Richardson, Elad, et al. “Encoding in style: a stylegan encoder for image-to-image translation,” *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. pp.2287-2296, 2021. DOI: 10.1109/CVPR46437.2021.00232

[5] Shen, Yujun, et al. “Interpreting the latent space of gans for semantic face editing,” *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2020.

DOI: 10.1109/CVPR42600.2020.00926

[6] Alaluf, Yuval, Or Patashnik, and Daniel Cohen-Or, “Restyle: A residual-based stylegan encoder via iterative refinement,” *Proceedings of the IEEE/CVF International Conference on Computer Vision*. 2021. DOI: 10.48550/arXiv.2104.02699

[7] KARRAS, Tero, et al. “Alias-free generative adversarial networks,” *Advances in Neural Information Processing Systems*, 34, pp.852-863. 2021.

DOI: 10.48550/arXiv.2106.12423

[8] Heusel, Martin, et al., “Gans trained by a two time-scale update rule converge to a local nash equilibrium,” *Advances in neural information processing systems 30* (2017).

DOI: 10.48550/arXiv.1706.08500

[9] Rothe, Rasmus, Radu Timofte, and Luc Van Goo, “Dex: Deep expectation of apparent age from a single image,” *Proceedings of the IEEE international conference on computer vision workshops*. 2015. DOI: 10.1109/ICCVW.2015.41

BIOGRAPHY

Hee-Yeol Lee (Student Member)



2016 : BS degree in Electronic Engineering, Hanbat National University
National University
2018 : MS degree in Electronic Engineering, Hanbat National University

2018~current : Ph. D degree course of Electronic Engineering, Hanbat National University

Seung-Ho Lee (Member)



1986 : BS degree in Electronic Engineering, Hanyang University
1989 : MS degree in Electronic Engineering, Hanyang University
1994 : Ph. D degree in Electronic Engineering, Hanyang University

1994~current : Professor, Department of Electronic Engineering, Hanbat National University