

열화 무관 특징 임베딩 기반 블라인드 폭우 얼굴 영상 복원

김철환*, 손창환**

Blind Heavy Rain Face Image Restoration based on Degradation Agnostic Feature Embedding

Cheol-Hwan Kim*, Chang-Hwan Son**

본 성과물은 국립군산대학교의 연구비 지원에 의해 연구되었으며, 또한 2025년도 정부(과학기술정보통신부)의 재원으로 연구개발특구진흥재단-(군산) 2025년 강소특구 지역 특성화 육성 사업의 지원을 받아 수행된 연구임 (RS-2025-02312252)

요 약

이 논문에서는 폭우 환경에서의 저화질 얼굴 영상을 고품질로 복원하기 위한 블라인드 기반 생성형 얼굴 영상 복원 모델을 제안한다. 기존 얼굴 영상 복원 모델은 기상 환경으로 수반되는 빗줄기, 빗방울 축적, 해상도 저하 등 복합적인 열화 과정을 처리할 수 있는 전용 모듈이 부족해, 복원 영상에서 색상 및 텍스처 왜곡이나 주요 얼굴 구성 요소의 표현력 저하 문제가 나타난다. 이에 본 연구에서는 저화질 및 고품질 영상 간의 특징을 공유 임베딩 공간에서 정렬할 수 있는 열화 무관 특징 임베딩 학습 기법을 도입하고, 열화 유형과 관계 없이 저화질 영상에서 고품질의 특징을 추출할 수 있는 모듈을 설계하였다. 실험 결과, 제안한 모델은 기존 PSFRGAN 대비 PSNR이 평균 0.4544dB, SSIM이 0.0219, LPIPS는 0.0118, FID가 4.0268 향상되었으며, 최종 복원 영상에서 텍스처 및 색상 왜곡을 효과적으로 제거하고 세부 표현 능력을 향상시켰음을 확인하였다.

Abstract

This paper proposes a blind generative face restoration model to recover high-quality face images from low-quality inputs captured in heavy rain. Existing models lack dedicated modules to handle complex degradations like rain streaks, accumulation, and resolution loss, causing color and texture distortions and degraded facial details. To address this, we introduce a degradation-agnostic feature embedding learning technique that aligns features of low- and high-quality images in a shared embedding space and design a module extracting high-quality features regardless of degradation type. Experimental results show that our model improves PSNR by 0.4544 dB, SSIM by 0.0219, LPIPS by 0.0118, and FID by 4.0268 over PSFRGAN, effectively removing distortions and enhancing detailed representations in restored images.

Keywords

blind image restoration, degradation-agnostic, feature embedding, statistical distribution

* 국립군산대학교 소프트웨어학과 학사과정
- ORCID: <https://orcid.org/0009-0004-8591-5134>
** 국립군산대학교 소프트웨어학과 교수(교신저자)
- ORCID: <https://orcid.org/0000-0001-7077-3074>

· Received: Jul. 23, 2025, Revised: Aug. 20, 2025, Accepted: Aug. 23, 2025
· Corresponding Author: Chang-Hwan Son
Department of Software Science and Engineering, Kunsan National University, Republic of Korea
Tel.: +82-63-469-8915, Email: cson@kunsan.ac.kr

I. 서론

얼굴 영상 복원(Face image restoration)[1]은 열화된 얼굴 영상으로부터 고화질 얼굴 영상을 복원하는 기술로, 지능형 CCTV 시스템의 발전과 함께 실종자 탐색, 용의자 식별, 출입자 관리 등 다양한 분야에서 활용되고 있다. 기존 얼굴 영상 복원 기법들은 열화 조건(예: 잡음 수준, 다운샘플링 비율 등)이 사전에 알려졌다고 가정하지만, 이는 실제 환경과는 맞지 않다. 이러한 한계를 극복하기 위해, 열화 파라미터의 명시적 정보 없이도 작동 가능한 블라인드 얼굴 영상 복원 기법들이 최근 제안되고 있다.

생성적 적대 신경망(GAN, Generative Adversarial Networks)[2]은 얼굴 영상의 확률 밀도 분포를 적대적 학습(Adversarial learning)을 통해 효과적으로 모델링할 수 있기 때문에, 블라인드 얼굴 영상 복원의 핵심 프레임워크로 자리매김하고 있다. GAN 기반 블라인드 얼굴 영상 복원 기법은 기하학적 사전 정보(Geometric prior)[3][4], 참조 사전 정보(Reference prior)[5]-[7], 생성적 사전 정보(Generative prior)[8][9] 등 다양한 사전 정보를 활용하여 얼굴의 세부 표현을 향상하고 복원된 얼굴의 충실도를 높일 수 있다.

그러나 최근의 블라인드 얼굴 영상 복원 기법들은 여전히 몇 가지 한계를 내포하고 있다. 지능형 CCTV는 주로 실외 환경에 설치되므로, 폭우나 안개와 같은 악천후 조건에서 촬영된 영상은 빗줄기나 대비 저하와 같은 심각한 영상 결함을 포함할 수 있다. 기존 얼굴 영상 복원 기법들은 이러한 기상 환경에 기인한 열화를 효과적으로 처리할 수 있는 전용 복원 모듈을 포함하고 있지 않기 때문에, 복원 성능은 근본적인 한계를 지닐 수밖에 없다.

본 연구에서는 기상 환경으로 수반된 영상 결함을 효과적으로 제거할 수 있는 열화 무관 특징 임베딩(DAFE, Degradation-Agnostic Feature Embedding) 학습 기법을 제안한다. 특히, 제안한 DAFE 모듈은 두 종류의 영상 인코더를 활용하여 고화질 및 저화질 얼굴 영상의 임베딩 특징을 정렬함으로써, 추론 단계에서 저화질 얼굴 영상에서 고화질의 얼굴 특징을 추출할 수 있다. 이를 통해, 입력 영상의 열화 수준에 무관하게 고품질의 특징을 안정적으로 유도

할 수 있으며, 전반적인 얼굴 영상 복원 성능을 향상할 수 있다. 실험 결과를 통해, 제안한 DAFE 기법이 기상 조건에 의해 유발된 다양한 영상 결함을 효과적으로 제거하고, 얼굴의 형태 및 세부 표현을 보다 정확하고 복원할 수 있음을 확인하였다.

II. 관련 연구

최근 블라인드 방식의 얼굴 영상 복원 기법은 다양한 사전 정보를 모델링하는 데 주력해 왔다. 사전 정보는 크게 기하학적 사전 정보, 참조 사전 정보, 생성적 사전 정보로 구분될 수 있다. 첫째, 기하학적 사전 정보는 랜드마크[10][11], 파싱 맵[3][4], 히트맵[12]과 같이 얼굴의 구조적 특성을 표현할 수 있는 2차원 맵을 말한다. 이러한 정보는 GAN 기반의 생성형 모델에서 조건부 입력으로 활용될 수 있으며 얼굴 복원력을 강화하는 데 기여할 수 있다. 그러나 저화질 얼굴 영상에서 이러한 기하학적 사전 정보를 정확히 추정하기는 어렵다. 둘째, 참조 사전 정보는 동일 인물의 고화질 참조 영상[7]이나, 사전 학습을 통해 얼굴 전체를 표현할 수 있는 얼굴 사전(Face dictionaries)[5][6]을 말한다. 하지만 실제 환경에서는 고화질 참조 영상의 확보가 쉽지 않으며 얼굴 사전도 크기 용량의 제약으로 인해 얼굴 표현력의 다양성에 한계를 지닌다. 셋째, 생성적 사전 정보는 StyleGAN[13]과 같은 사전 학습된 생성형 모델에서 추출된 잠재 표현을 말한다. 이러한 생성형 모델은 얼굴의 구조, 색상, 스타일 등 다양한 잠재 정보를 함축하고 있기에 고품질의 얼굴 복원에 유용하게 활용될 수 있다.

상기 사전 정보 기반의 생성형 영상 복원 모델은 청명한 날씨에서 촬영된 저화질 영상 복원에는 효과적이거나, 기상 환경에 기인한 빗줄기, 잡음, 빗방울 축적(Rain accumulation)과 같은 다양한 영상 결함을 제거하기에는 부적합하다. 이는 폭우, 안개, 눈송이 등과 같은 다양한 기상 조건에서 발생하는 열화 양상에 특화된 전용 복원 모듈이 부재하기 때문이다. 따라서 기상 환경의 종류나 강도에 무관하게 고품질의 얼굴 특징을 안정적으로 추출할 수 있는 기상 강건성을 갖춘 학습 기법을 개발할 필요가 있다.

III. 제안한 열화 무관 특징 임베딩 기반 블라인드 폭우 얼굴 영상 복원 모델

본 연구에서는 기존 블라인드 얼굴 영상 복원 기법에 기상 강건성을 부여하기 위해 새로운 DAFE 학습 기법을 제안하고자 한다. DAFE 학습이란 저화질 영상과 고화질 영상의 특징을 임베딩 공간에서 정렬함으로써, 추론 단계에서 저화질 얼굴 영상에서 고화질 영상의 특징을 추출할 수 있는 기법을 말한다. 제안한 DAFE 모듈은 입력 영상의 열화 유형에 상관없이 안정적으로 고화질의 특징을 추출할 수 있어 다양한 기상 조건에서도 얼굴 복원 성능을 유지할 수 있을 것으로 기대된다. 제안한 DAFE 모듈을 기존의 GAN 기반의 생성형 모델인 PSFRGAN[4]에 적용하여 기상 강건성에 대한 효과를 실험적으로 검증하고자 한다.

3.1 DAFE 학습 과정

그림 1은 제안한 DAFE 학습 과정을 보여준다. 그림에서 보듯이, DAFE 모듈은 두 개의 영상 인코더가 필요하다. 첫 번째는 고화질 영상을 공유 임베딩 공간으로 사상하는 HQ(High-Quality) 인코더, 두 번째는 저화질 영상을 동일한 임베딩 공간으로 사상하는 LQ(Low-Quality) 인코더이다. 각 인코더에서 추출된 고화질 및 저화질 특징은 공유 임베딩 공간에서 식 (1)의 평균 제곱 오차(MSE, Mean Squared Error) 손실함수를 최소화함으로써 정렬된다.

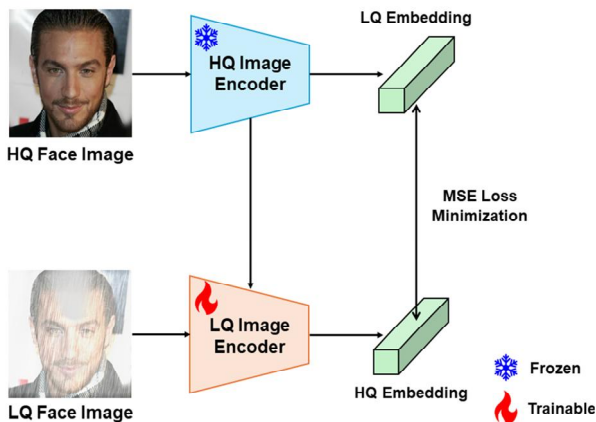


그림 1. DAFE 학습 과정
Fig. 1. DAFE training process

$$L_{MSE} = \| E_{HQ}(I_{HQ}) - E_{LQ}(I_{LQ}) \|_2^2 \quad (1)$$

여기서 L_{MSE} 는 DAFE의 손실함수이다. I_{HQ} 와 I_{LQ} 는 각각 고화질과 저화질 얼굴 영상을 나타내며, E_{HQ} 와 E_{LQ} 는 각각 HQ와 LQ 인코더를 의미한다. 그림 1에서 명시되었듯이, HQ 인코더의 파라미터는 고정된 상태에서 LQ 인코더의 파라미터만 갱신된다.

3.2 DAFE 기반 스타일 변환 파라미터 추정

최신 GAN 기반 생성형 얼굴 영상 복원 모델인 PSFRGAN[4]은 SOTA 모델 중 하나로써, 입력 난수 벡터로부터 스타일 변환(ST, Style Transformation)을 적용하여 최종 고화질 얼굴 영상을 복원한다. 여기서 스타일 변환이란 소스의 표준편차와 평균을 타겟의 표준편차와 평균으로 사상하는 기법을 말한다. PSFRGAN[4]에서는 소스는 입력 난수 벡터에서 생성된 특징을 말하며 타겟은 저화질 얼굴 영상과 파싱 영상에서 추출된 특징으로 정의된다.

제안한 DAFE 모듈을 기존 PSFRGAN[4] 모델과 통합하기 위해서는 DAFE 모듈에서 스타일 변환을 위한 표준편차와 평균 파라미터를 생성하는 과정이 필요하다. 그림 2는 DAFE 기반 스타일 변환 파라미터 추정 모듈(STPE, Style Transformation Parameter Estimation)으로써, DAFE를 활용하여 스타일 변환에 필요한 표준편차와 평균을 학습하는 과정을 보여준다.

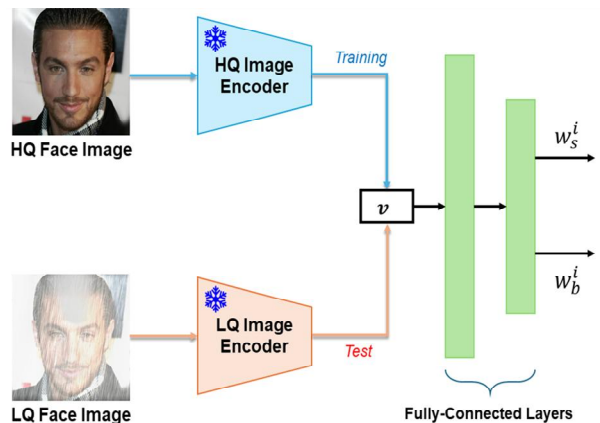


그림 2. DAFE 기반 스타일 변환 파라미터 추정
Fig. 2. DAFE-based style transformation parameter estimation

그림에서 보듯이, 스타일 변환 파라미터 추정은 훈련 단계(Training)와 추론(Inference) 단계가 다르게 동작한다. 즉, 훈련 단계에서는 HQ 인코더에서 임베딩 특징을 추출하고 추론 단계에서 LQ 인코더에서 임베딩 특징을 추출한다. 관련 스타일 변환 파라미터 추정은 식 (2)와 같이 표현된다.

$$\begin{cases} \text{Training} : v = E_{HQ}(I_{HQ}) \\ \text{Inference} : v = E_{LQ}(I_{LQ}) \end{cases} \quad (2)$$

여기서 v 는 공유 임베딩 공간으로 사상된 얼굴 특징을 의미하고, 훈련과 추론 단계에서 각각 HQ와 LQ 인코더로부터 임베딩 특징 벡터 v 를 추출한다.

스타일 변환을 위한 학습 파라미터는 식 (3)과 같이 공유 임베딩 특징 v 를 완전연결층(FC, Fully-Connected Layers)을 통과시켜 생성된다.

$$[w_s^i, w_b^i] = f_2^i(f_1^i(v)) \quad (3)$$

여기서 f_1 와 f_2 는 그림 2의 첫 번째와 두 번째 FC 계층을 의미하며, w_s 와 w_b 는 각각 표준편차와 평균에 해당하는 파라미터를 나타낸다. 인덱스 i 는 입력

난수 벡터에서 고화질 영상을 생성하기 위해 스타일 변환이 적용되는 스케일 정보에 해당한다. 즉, 표준편차와 평균 파라미터는 스케일별로 추정된다. 그리고 그림 2에서 두 개의 인코더는 모두 고정된 상태에서 FC 계층만 학습이 되고, 추후 스타일 변환 파라미터인 w_s 와 w_b 는 PSFRGAN[4]의 스타일 변환 보정을 위해 활용된다.

3.3 제안한 열화 무관 폭우 얼굴 영상 복원

그림 3은 제안한 DAFE 기반 얼굴 영상 복원 모델의 전체 아키텍처를 나타낸다. 제안한 모델은 스타일 변환 기반 생성형 모델인 PSFRGAN[4]을 변형한 것으로, 기존 PSFRGAN과의 주요 차이는 DAFE 모듈 추가에 있다. 제안한 모델은 난수 벡터를 입력으로 받아, 스타일 변환과 업샘플링 과정을 거쳐 최종 고화질 얼굴 영상을 생성한다. 이 과정에서 스타일 변환을 수행하기 위해, 저화질 얼굴 영상과 얼굴 파싱 영상을 추가로 입력받아 타겟의 표준편차와 평균을 계산한다. 전체 네트워크는 크게 스타일 변환 파라미터 추정(STPE) 모듈, 스타일 변환(ST) 모듈, 그리고 DAFE 기반 STPE 모듈로 구성된다.

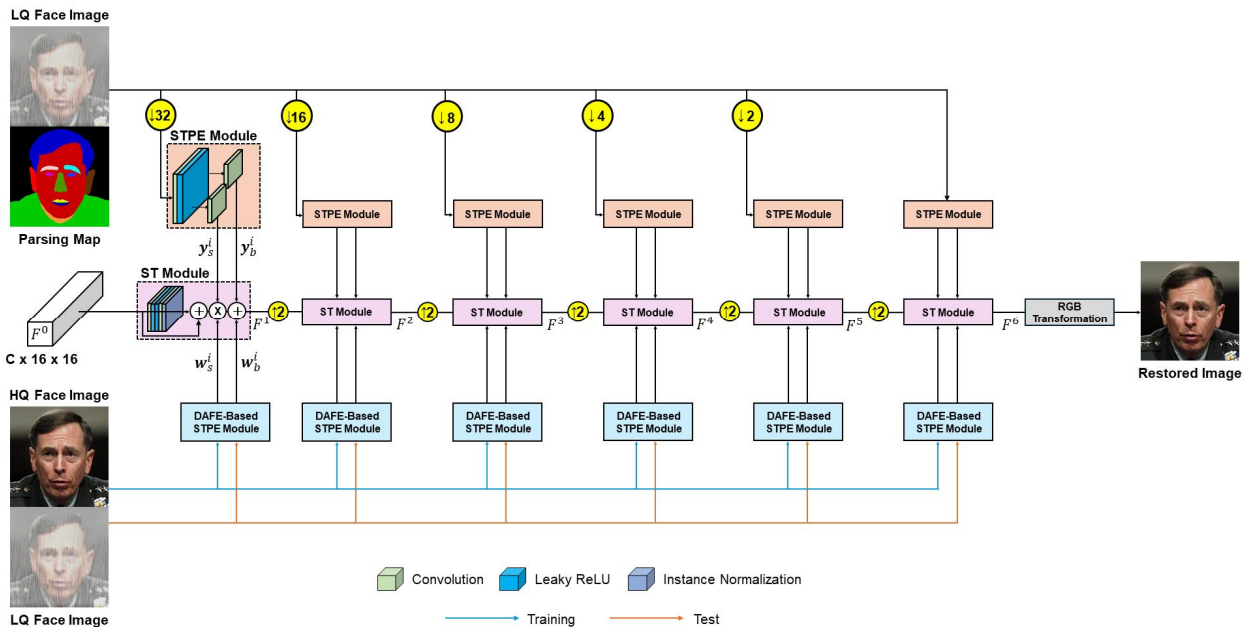


그림 3. 제안한 열화 무관 특징 임베딩 기반의 블라인드 폭우 얼굴 영상 복원 모델
Fig. 3. Proposed blind face restoration model under heavy rain based on degradation

먼저, STPE 모듈은 스타일 변환을 위해 타겟의 표준편차와 평균을 다음과 같이 계산한다.

$$[\mathbf{y}_s^i, \mathbf{y}_b^i] = \Psi(I_{LQ}^i, I_P^i) \quad (4)$$

여기서 Ψ 는 STPE 모듈을 나타내며 I_{LQ}^i 와 I_P^i 는 각각 i 번째 STPE 모듈의 입력인 다운샘플링된 저화질 얼굴 영상과 파싱 영상을 의미한다. 즉, STPE는 그림 3에서 보듯이 스케일별로 적용된다. 본 연구에서 실제 Ψ 는 합성곱(Convolution) 계층과 누설 정류선형 유닛(Leaky ReLU) 계층으로 구현되었다.

둘째, ST 모듈은 입력 소스의 표준편차와 평균을 타겟의 표준편차와 평균으로 변환한다.

$$F^i = (\mathbf{w}_s^i + \mathbf{y}_s^i) \frac{\Phi_{up}(F^{i-1}) - \mu(\Phi_{up}(F^{i-1}))}{\sigma(\Phi_{up}(F^{i-1}))} + (\mathbf{w}_b^i + \mathbf{y}_b^i) \quad (5)$$

식 (5)는 ST 모듈을 모델링한 것으로 F^{i-1} 는 i 번째 ST 모듈의 입력 소스를, Φ_{up} 는 업샘플링 계층을, μ 와 σ 는 각각 평균과 표준편차를 의미한다.

기존 PSFRGAN 모델에서는 스타일 변환을 위해 저화질 및 파싱 영상에서 추출된 y_s^i 와 y_b^i 만을 사용한다. 그러나 저화질 영상에서 타겟의 표준편차와 평균을 직접 추정하는 것은 정보 손실 및 왜곡으로 인해 정확도에 한계가 있을 수밖에 없다. 이를 보완하기 위해, 제안한 모델에서는 DAFE 기반 STPE 모듈에서 추출된 w_s^i 와 w_b^i 를 추가적으로 사용한다. DAFE 모듈은 저화질 영상과 고화질 영상 간의 특징을 정렬했기 때문에 추론 단계에서 고화질 영상의 특징을 추출할 수 있다. 따라서 w_s^i 와 w_b^i 의 추가적인 사용은 스타일 변환을 강화할 수 있고, 이는 궁극적으로 얼굴 영상 복원력 제고를 유도할 수 있다.

IV. 실험 및 결과

4.1 실험 환경

제안한 모델의 학습을 위해, 얼굴 복원 분야에서 널리 활용되는 CelebA-HQ[14] 데이터셋을 사용하였

다. 이 데이터셋은 전 세계 유명인의 고화질 얼굴 영상과 해당 얼굴 파싱 영상을 포함하고 있다. 총 30,000장의 영상 중 18,000장은 학습용으로, 5,000장은 테스트용으로 분할하여 사용하였다. 모든 실험은 RTX A6000 GPU 환경에서 수행되었으며, 배치 크기는 4, 학습 에폭 수는 100, 학습률은 0.0002로 설정하였다. 손실 함수는 기존 PSFRGAN[4] 모델에서 사용된 재구성 손실(Reconstruction loss), 적대 손실(Adversarial loss), 시맨틱 인식 스타일 손실(Semantic-aware style loss)을 동일하게 적용하였다.

4.2 정성적 화질 평가

제안한 모델의 성능을 평가하기 위해, 기존의 폭우 제거 기법(Heavy rain removal)[15]과 얼굴 복원 전용 모델인 HiFaceGAN[16], BFRfusion[17], PSFRGAN[4]과 비교 실험을 수행하였다. 이 중 폭우 제거 기법은 빗줄기 및 빗방울 추적 등으로 열화된 폭우 영상을 개선하여 고화질로 복원하는 데 중점을 둔 일반 영상 복원 기법이며, 나머지 세 모델은 얼굴 영상 복원에 특화된 방법들이다. 그림 4는 제안한 모델과 기존 모델을 적용한 복원 결과 영상을 보여준다. 첫 번째 열은 원본 고화질 얼굴 영상이며 두 번째 열은 블라인드 열화 모델[15]을 적용한 입력 폭우 영상이다. 열화 모델은 실세계 열화 과정을 최대한 반영하기 위해, 해상도, 빗줄기, 물방울 추적, 블러링 등에 대한 열화 파라미터를 랜덤하게 설정한다. 그림 4의 결과 영상에서, 기존 폭우 제거 모델은 색상 복원 능력이 미흡하여 결과 영상의 화질도 전반적으로 나빠진 것을 볼 수 있다. HiFaceGAN 모델은 비록 얼굴 영상 복원 모델이지만, 얼굴 결과 영상에서 텍스처 왜곡 현상이 많이 발생하는 것을 볼 수 있다. 반면, BFRfusion과 PSFRGAN 모델은 얼굴의 주요 윤곽과 세부 구조를 전체적으로 잘 복구한 것을 볼 수 있다. 하지만 BFRfusion 모델에 적용한 결과에서는 전반적으로 색상이 얼어지는 경향이 있으며, 네 번째 열에서는 귀결이의 모양이 원본과는 다르게 복원된 것을 볼 수 있다. 이는 BFRfusion 모델이 원본 얼굴을 충실하게 표현하기보다는 리얼리즘을 반영하여 귀결이 모양을 변형한 것으로 해석된다.

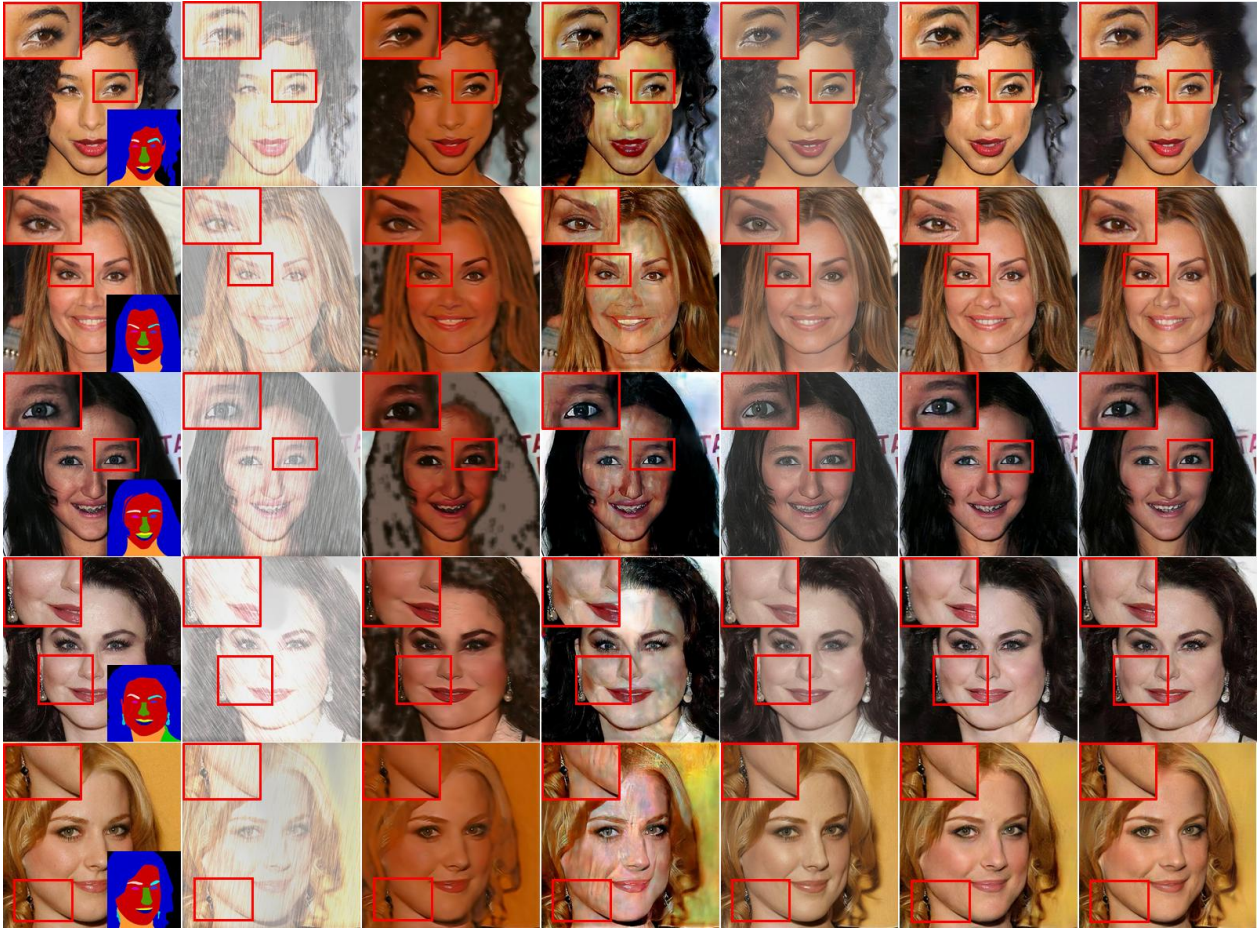


그림 4. 실험 결과 ; 원본 영상 (첫 번째 열), 폭우 얼굴 영상 (두 번째 열), 폭우 제거 기법[15] (세 번째 열), HiFaceGAN[16] (네 번째 열), BFRfusion[17] (다섯 번째 열), PSFRGAN[4] (여섯 번째 열), 제안한 모델 (마지막 열)
 Fig. 4. Experimental results ; Original images (first column), heavy rain face images (second column), heavy rain removal method[15] (third column), HiFaceGAN[16] (fourth column), BFRfusion[17] (fifth column), PSFRGAN[4] (sixth column), and the proposed model (last column)

PSFRGAN은 상대적으로 우수한 복원 능력을 보여주었지만, 제안한 모델에 비해 몇 가지 부족한 점이 확인되었다. 첫째, 육안으로 감지될 수 있는 텍스처 왜곡이 얼굴 구성 요소에서 전반적으로 확인되었다. 특히, 첫 번째 결과 영상에서 물결무늬와 같은 텍스처 결합이 존재하는 것을 확인할 수 있다. 둘째, 눈, 코, 입 등 얼굴 구성 요소의 선명도가 상대적으로 떨어진다. 특히, 네 번째 결과 영상의 귀걸이 모양이 흐릿하게 복원된 것을 볼 수 있다. 셋째, 세부 표현 능력에서도 상대적으로 미흡하다. 두 번째 결과 영상에서 눈썹의 끝부분이 일그러져 복원된 것을 관찰할 수 있다. 반면, 제안한 DAFE 기반 얼굴 영상 복원 모델은 입력 열화 유형에 무관하게 고화질의 특징을 효과적으로 추출함으로써, 보

다 정밀하고 선명한 복원 결과를 보여주었다. 특히, PSFRGAN에 DAFE 모듈을 결합함으로써, 얼굴 복원 성능이 전반적으로 향상되었다. 이는 DAFE 모듈이 기상 환경에 강인한 고화질 특징 추출에 효과적이며 얼굴 세부 구조와 텍스처 복원 능력을 강화할 수 있음을 입증한다.

4.3 정량적 화질 평가

정량적 화질 평가를 위해, 영상 복원 분야에서 널리 사용되는 네 가지 평가 지표인 PSNR(Peak Signal-to-Noise Ratio), SSIM(Structural Similarity Index Measure)[18], LPIPS(Learned Perceptual Image Patch Similarity)[19], FID(Frechet Inception

Distance)[20]를 적용하였다. 이 중 PSNR과 SSIM은 원본 영상과 복원 영상 간의 화질 유사도를 측정하는 참조 기반 지표이며, 값이 클수록 복원 품질이 우수함을 나타낸다. 반면, LPIPS와 FID는 원본 영상 없이 인간의 인지적 유사도를 기반으로 평가하는 비참조 기반 지표로써, 값이 작을수록 높은 인지적 화질을 의미한다.

표 1은 테스트 데이터에 적용한 각 모델의 정량적 화질 평가 결과이다. 표에서 보듯이, 제안한 모델이 PSNR, SSIM, LPIPS 지표에서 모두 가장 우수한 결과를 달성하였다. 이는 원본 영상과의 충실도 측면에서도 인지적 측면에서도 복원된 영상 품질이 제일 우수함을 의미한다. 비록 FID 지표에서는 BFRfusion 모델이 소폭 앞서는 결과를 보였으나, 그림 4의 시각적 평가와는 다소 차이가 있음을 보여준다. 이는 FID 지표가 영상의 통계적 분포에 기반한 평가 방식이기 때문에, 시각적으로 인지되는 화질과 반드시 일치하지 않을 수 있음을 보여준다. 표 1의 결과는 제안한 DAFE 모듈이 열화에 무관하게 고화질의 특징을 효과적으로 추출하고, 스타일 변환을 위한 파라미터의 정확도를 개선하여 최종 얼굴 복원 성능을 유의미하게 개선할 수 있음을 뒷받침해 준다.

4.4 DAFE 모듈의 효과 분석

DAFE 모듈의 특징 추출 효과를 확인하기 위해, 고차원 특징 벡터를 저차원 공간에 투영하여 시각화할 수 있는 t-SNE 기법[21]을 적용하였다.

그림 5는 2차원 공간으로 투영된 특징들을 시각화한 것으로, 고화질(HQ) 인코더를 통과한 고화질 영상의 특징과, 학습 전후의 저화질(LQ) 인코더를 통과한 저화질 영상의 특징을 각각 보여준다. 그림

5의 결과에서, DAFE 모듈을 적용한 후의 저화질 영상 특징이 원본 고화질 영상의 특징과 유사해진 것을 확인할 수 있다. 이는 DAFE 모듈이 저화질 영상으로부터 고화질 수준의 특징을 효과적으로 추출할 수 있음을 의미한다.

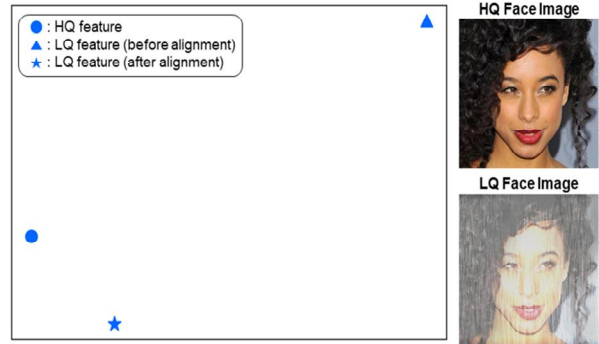


그림 5. HQ 인코더, 학습 전 LQ 인코더, 학습 후 LQ 인코더의 출력 특징 비교

Fig. 5. t-SNE visualization of features from the HQ encoder, pre-trained LQ encoder, and post-trained LQ encoder

V. 결 론

본 논문에서는 대표적인 기상 환경인 폭우로 인해 열화된 저화질 얼굴 영상에서 고화질의 얼굴 영상을 복구할 수 있는 새로운 얼굴 영상 복원 기법을 제안하였다. 제안한 모델은 폭우로 인해 수반된 빗줄기, 해상도 저하 및 물방울 축적과 같은 영상 결함을 효과적으로 제거하고, 고화질의 얼굴 영상을 복구하기 위해 열화에 강인한 특징 추출 기법인 DAFE 모듈을 도입하였다. 특히, LQ 및 HQ 인코더의 특징 결과를 정렬함으로써, 입력 열화에 무관하게 안정적으로 고화질의 특징을 추출할 수 있는 학습 구조를 설계하였다.

표 1. 정량적 화질 평가

Table 1. Quantitative quality assessment

	PSNR ↑	SSIM ↑	LPIPS ↓	FID ↓
Heavy rain removal[15]	14.5337	0.6087	0.3106	39.8703
HiFaceGAN[16]	20.3458	0.5925	0.2688	71.3781
BFRfusion[17]	18.5903	0.5800	0.2399	12.5220
PSFRGAN[4]	22.2651	0.6814	0.2317	18.6487
Proposed model	22.7195	0.7033	0.2199	14.6219

이 DAFE 모듈을 기존 GAN 기반 얼굴 영상 복원 모델인 PSFRGAN 모델에 적용하여 그 효과를 검증하였다. 실험 결과를 통해, 제안한 DAFE 기반 얼굴 영상 복원 모델이 결함 제거, 얼굴 세부 표현 복원 및 선명도를 개선하는 데 효과적임을 확인하였다. 향후 연구에서는 실제 폭우 환경에서 촬영된 실 환경 얼굴 영상 데이터를 활용하여, 제안한 모델의 성능을 테스트하고 고도화할 계획이다.

Acknowledgement

'2025년도 한국정보기술학회 하계종합학술대회에서 발표한 논문(열화-무관 특징 임베딩 기반 블라인드 폭우 얼굴 영상 복원)[22]을 확장한 것임'

References

- [1] X. Wang, Y. Li, H. Zhang, and Y. Shan, "Towards real-world blind face restoration with generative facial prior", Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition, Nashville, TN, USA, pp. 9168-9178, Jun. 2021. <https://doi.org/10.1109/CVPR46437.2021.00905>.
- [2] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets", Communications of the ACM, Vol. 63, No. 11, pp. 139-144, Oct. 2020. <https://doi.org/10.1145/3422622>.
- [3] Y. Chen, Y. Tai, X. Liu, C. Shen, and J. Yang, "FSRNet: End-to-end learning face super-resolution with facial priors", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, pp. 2492-2501, Jun. 2018. <https://doi.org/10.1109/CVPR.2018.00264>.
- [4] C. Chen, X. Li, L. Yang, X. Lin, L. Zhang, and K.-Y. K. Wong, "Progressive semantic-aware style transformation for blind face restoration", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, pp. 11896-11905, Jun. 2021. <https://doi.org/10.48550/arXiv.2009.08709>.
- [5] X. Li, C. Chen, S. Zhou, X. Lin, W. Zuo, and L. Zhang, "Blind face restoration via deep multi-scale component dictionaries", Proc. European Conference on Computer Vision, Montreal, Canada, pp. 399-415, Oct. 2021. <https://doi.org/10.48550/arXiv.2008.00418>.
- [6] X. Li, W. Li, D. Ren, H. Zhang, M. Wang, and W. Zuo, "Enhanced blind face restoration with multi-exemplar images and adaptive spatial feature fusion", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 2703-2712, Jun. 2020. <https://doi.org/10.1109/CVPR42600.2020.00278>.
- [7] B. Dogan, S. Gu, and R. Timofte, "Exemplar guided face image super-resolution without facial landmarks", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, pp. 1814-1823, Jun. 2019. <https://doi.org/10.48550/arXiv.1906.07078>.
- [8] J. Gu, Y. Shen, and B. Zhou, "Image processing using multi-code GAN prior", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 3012-3021, Jun. 2020. <https://doi.org/10.1109/CVPR42600.2020.00308>.
- [9] S. Menon, A. Damian, S. Hu, N. Ravi, and C. Rudin, "PULSE: Self-supervised photo upsampling via latent space exploration of generative models", Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, pp. 2437-2445, Jun. 2020. <https://doi.org/10.48550/arXiv.2003.03808>.
- [10] D. Kim, M. Kim, G. Kwon, and D. Kim, "Progressive face super-resolution via attention to facial landmark", Proc. British Machine Vision Conference, Cardiff, UK, No. 192, Sep. 2019. <https://doi.org/10.48550/arXiv.1908.08239>.
- [11] Y. Song, J. Zhang, L. Gong, S. He, L. Bao, J.

- Pan, Q. Yang, and M.-H. Yang, "Joint face hallucination and deblurring via structure generation and detail enhancement", *International Journal of Computer Vision*, Vol. 127, pp. 785-800, Jan. 2019. <https://doi.org/10.1007/s11263-019-01148-6>.
- [12] X. Yu, B. Fernando, B. Ghanem, F. Porikli, and R. Hartley, "Face super-resolution guided by facial component heatmaps", *Proc. European Conference on Computer Vision*, Munich, Germany, pp. 217-233, Sept. 2018. https://doi.org/10.1007/978-3-030-01240-3_14.
- [13] T. Karras, S. Laine, and T. Aila, "A style-based generator architecture for generative adversarial networks", *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 4401-4410, Jun. 2019. <https://doi.org/10.1109/CVPR.2019.00453>.
- [14] T. Karras, T. Aila, S. Laine, and J. Lehtinen, "Progressive growing of GANs for improved quality, stability, and variation", *Proc. International Conference on Learning Representations*, Vancouver, Canada, pp. 1-26, May 2018. <https://doi.org/10.48550/arXiv.1710.10196>.
- [15] R. Li, L.-F. Cheong, and R. T. Tan, "Heavy rain image restoration: Integrating physics model and conditional adversarial learning", *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA, USA, pp. 1633-1642, Jun. 2019. <https://doi.org/10.48550/arXiv.1904.05050>.
- [16] L. Yang, C. Liu, P. Wang, S. Wang, P. Ren, S. Ma, and W. Gao, "Hifacegan: Face renovation via collaborative suppression and replenishment", *Proc. ACM International Conference on Multimedia*, NY, United States, pp. 1551-1560, Oct. 2020. <https://doi.org/10.1145/3394171.3413965>.
- [17] X. Chen, J. Tan, T. Wang, K. Zhang, W. Luo, and X. Cao, "Toward real-world blind face restoration with generative diffusion prior", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 34, No. 9, pp. 8494-8508, Sep. 2024. <https://doi.org/10.48550/arXiv.2312.15736>.
- [18] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structure similarity", *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp. 600-612, Apr. 2004. <https://doi.org/10.1109/TIP.2003.819861>.
- [19] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The unreasonable effectiveness of deep features as a perceptual metric", *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 586-595, Jun. 2018. <https://doi.org/10.1109/CVPR.2018.00068>.
- [20] M. Heusel, H. Ramsauer, T. Unterthiner, B. Nessler, and S. Hochreiter, "GANs trained by a two time-scale update rule converge to a local Nash equilibrium", *Proc. Advances in Neural Information Processing Systems*, Long Beach, CA, USA, pp. 6626-6637, Dec. 2017. <https://doi.org/10.48550/arXiv.1706.08500>.
- [21] L. van der Maaten and G. Hinton, "Visualizing data using t-SNE", *Journal of Machine Learning Research*, Vol. 9, No. 11, pp. 2579-2605, Nov. 2008.
- [22] C. H. Kim, J. H. Shin and C. H. Son, "Blind Face Restoration in Heavy Rain with Degradation-Agnostic Feature Embedding", *Proc. Korea Information Technology Conference*, Jeju, Korea, pp. 571-575, Jun. 2025.

저자소개

김 철 환 (Cheol-Hwan Kim)



2020년 3월 ~ 현재 :
국립군산대학교
소프트웨어학과 학사과정
관심분야 : 컴퓨터 비전, 영상처리,
기계학습, 딥 러닝

손 창 환 (Chang-Hwan Son)



2002년 2월 : 경북대학교
전자전기공학부(공학사)
2004년 2월 : 경북대학교
전자공학과(공학석사)
2008년 8월 : 경북대학교
전자공학과(공학박사)
2017년 4월 ~ 현재 :

국립군산대학교 소프트웨어학과 교수
관심분야 : 컴퓨터 비전, 영상처리, 기계학습, 딥 러닝