

저작권양도서

(Copyright Transfer Form)

소속 : 인하대학교 정보통신공학과

성명 : 김지민, 양유진

학번 : 12205102, 12211796

논문제목 : 다양한 영상 열화를 개선하는 딥러닝 파운데이션 모델 개발

본인은 상기 논문을 2024학년도 2학기 정보통신프로젝트 최종 보고서 겸 결과 논문으로 제출하고자 합니다. 본 논문의 내용은 저자가 직접 연구한 결과인 것과 이전에 출판된 적이 없음을 확인합니다. 또한 공저자와 더불어 인하대학교 정보통신공학과에서 발간하는 논문집에 본 논문을 수록하는 것을 허락하며 제반 저작권을 정보통신공학과에 양도합니다.

2024년 1월 06일

주저자 : 김지민 (김지민)

: 양유진 (양유진)

정보통신공학과장 귀하

다양한 영상 열화를 개선하는 딥러닝 파운데이션 모델 개발

Deep learning Foundation model for Enhancing Various Image Degradations

지도교수 확인함

김지민, 양유진
(Jimin Kim and Yujin Yang)



Abstract: Denoising, Deblurring, Anti-aliasing, Inpainting을 모두 수행할 수 있는 파운데이션 모델을 개발한다. 이를 위해 트랜스포머 기반의 디블러링 및 디노이징 모델을 확장하여 모델 구조를 개선하고, 다양한 열화 요소에 대응할 수 있도록 설계한다. 최적의 성능을 확보하기 위해 각 열화 요소에 적합한 데이터셋을 활용하여 4가지 테스크에 대해 개별적으로 파인튜닝을 수행한 후, 성능을 비교·분석한다. 본 연구는 단일 모델을 활용한 다중 열화 복원 가능성을 탐색하며, 향후 영상 복원 및 품질 향상 분야에서의 실용성을 높이는 데 기여할 것으로 기대된다.

Keywords: Image Restoration, Foundation model, Transformer, Deep Learning, Multi-task Learning, Deblurring, Denoising, Anti-aliasing, Inpainting

I. 서론

현대 사회에서 이미지와 영상 데이터는 의료 영상, 자율주행, AR/VR, 감시 시스템 등 다양한 산업에서 핵심적인 역할을 하고 있다. 그러나 현실 세계에서 수집되는 데이터는 다양한 열화 요소 (Denoising, Deblurring, Anti-aliasing, Inpainting)로 인해 품질 저하가 발생하며, 이는 데이터의 분석 및 활용도를 크게 저하시킨다. 기존 연구들은 주로 특정 열화 문제(예: 노이즈 제거, 블러 복원)만을 해결하는 데 중점을 두었으며, 한 가지 문제에 최적화된 모델이 다른 문제에는 적합하지 않다는 한계가 있었다. 또한, 여러 문제를 다룬다고 하여도 좁은 범위의 열화(예: 노이즈, 블러)를 다루었고 넓은 범위의 열화(예: 노이즈와 오클루전)를 다루는 모델을 찾기 어렵다.

현재 주목받고 있는 ‘파운데이션 모델’은 다양한 테스크를 하나의 모델로 통합적으로 해결할 수 있는 가능성을 제시하지만, 영상 복원 분야에서는 이러한 접근이 아직 초기 단계에 머물러 있다. 이에 따라, 본 연구는 Transformer 기반의 Restormer 모델을 확장하여 다양한 열화 요소를 동시에 해결할 수 있는 다목적 파운데이션 모델을 개발하는 것을 목표로 한다. 이러한 모델은 의료 영상의 진단 정확도 향상, 자율주행 시스템의 안전성 강화, 콘텐츠 제작의 품질 개선 등 폭넓은 사회적, 경제적 기여할 수 있다.

본 논문에서는 먼저 각 열화 이미지와 관련된 데이터셋 전처리 방식을 소개한다. 이후, 학습의 기반이 된 Restormer 모델의 구조와 특징을 설명한 뒤, 인페인팅 학습을 위해 적용된 모델 구조 수정 방법을 논의한다. 마지막으로, 학습된 모델의 성능을 평가하고, 파인튜닝 이후 성능과 비교함으로써 본 연구가 열화 요소 제거를 통해 다양한 분야에서 가져올 잠재적 기여와 의의를 논의한다.

II. 본론

1. 개발 환경 및 도구

본 연구에서는 딥러닝 모델 학습 및 실험을 수행하기 위해 NVIDIA TITAN RTX GPU와 CUDA를 사용했다. 딥러닝 프레임워크로는 PyTorch를 사용하며, 데이터 처리와 모델 학습 지원을 위해 Anaconda 및 torchvision과 같은 오픈소스 라이브러리를 활용했다. 개발 도구로는 VSCode와 Python을 사용하여 코드 작성 및 디버깅을 진행했다.

2. 시스템 흐름도 및 개요

본 연구는 다양한 열화 유형(Denoising, Deblurring, Anti-aliasing, Inpainting)을 복원하기 위한 딥러닝 기반 파운데이션 모델 설계를 목표로 한다. 이를 위해 열화 유형별로 데이터를 분리하여 학습하고, 각 유형에 특화된 Fine-tuning 과정을 통해 최적의 복원 성능을 확보한다.

입력 데이터는 열화 유형에 따라 Denoising, Deblurring, Anti-aliasing, Inpainting으로 구분되며 각각 별도의 디렉토리에 저장된다. 이러한 데이터 구분은 열화 유형별로 처리 방식이 상이하기 때문에 필수적이다.

Inpainting의 경우 랜덤 스크래치 마스크를 활용하여 손상된 이미지를 생성하고, 이를 네트워크 입력과 함께 사용하여 손상된 영역의 복원에 집중할 수 있도록 설계한다.

Restormer를 기반으로 설계된 파운데이션 모델은 여러 열화 유형 데이터를 학습하여 범용적인 복원 능력을 갖춘다. 이후, 열화 유형별로 Fine-tuning을 수행하여 특정 열화에 특화된 성능을 확보한다. Fine-tuning 과정은 각 열화 유형의 데이터셋을 활용하며, 실험적으로 학습 데이터셋 크기 및 Iteration 수에 따른 성능 변화를 평가한다. 이를 통해, 열화 유형별로 차별화된 학습 전략이 모델 성능에 미치는 영향을 분석한다.

다양한 영상 열화를 개선하는 딥러닝 파운데이션 모델 개발

Deep learning Foundation model for Enhancing Various Image Degradations

김지민, 양유진
(Jimin Kim and Yujin Yang)

Abstract: Denoising, Deblurring, Anti-aliasing, Inpainting을 모두 수행할 수 있는 파운데이션 모델을 개발한다. 이를 위해 트랜스포머 기반의 디블러링 및 디노이징 모델을 확장하여 모델 구조를 개선하고, 다양한 열화 요소에 대응할 수 있도록 설계한다. 최적의 성능을 확보하기 위해 각 열화 요소에 적합한 데이터셋을 활용하여 4가지 태스크에 대해 개별적으로 파인튜닝을 수행한 후, 성능을 비교·분석한다. 본 연구는 단일 모델을 활용한 다중 열화 복원 가능성을 탐색하며, 향후 영상 복원 및 품질 향상 분야에서의 실용성을 높이는 데 기여할 것으로 기대된다.

Keywords: Image Restoration, Foundation model, Transformer, Deep Learning, Multi-task Learning, Deblurring, Denoising, Anti-aliasing, Inpainting

I. 서론

현대 사회에서 이미지와 영상 데이터는 의료 영상, 자율주행, AR/VR, 감시 시스템 등 다양한 산업에서 핵심적인 역할을 하고 있다. 그러나 현실 세계에서 수집되는 데이터는 다양한 열화 요소 (Denoising, Deblurring, Anti-aliasing, Inpainting)로 인해 품질 저하가 발생하며, 이는 데이터의 분석 및 활용도를 크게 저하시킨다. 기존 연구들은 주로 특정 열화 문제(예: 노이즈 제거, 블러 복원)만을 해결하는 데 중점을 두었으며, 한 가지 문제에 최적화된 모델이 다른 문제에는 적합하지 않다는 한계가 있었다. 또한, 여러 문제를 다룬다고 하여도 좁은 범위의 열화(예: 노이즈, 블러)를 다루었고 넓은 범위의 열화(예: 노이즈와 오클루전)를 다루는 모델을 찾기 어렵다.

현재 주목받고 있는 ‘파운데이션 모델’은 다양한 태스크를 하나의 모델로 통합적으로 해결할 수 있는 가능성을 제시하지만, 영상 복원 분야에서는 이러한 접근이 아직 초기 단계에 머물러 있다. 이에 따라, 본 연구는 Transformer 기반의 Restormer 모델을 확장하여 다양한 열화 요소를 동시에 해결할 수 있는 다목적 파운데이션 모델을 개발하는 것을 목표로 한다. 이러한 모델은 의료 영상의 진단 정확도 향상, 자율주행 시스템의 안전성 강화, 콘텐츠 제작의 품질 개선 등 폭넓은 사회적, 경제적 기여할 수 있다.

본 논문에서는 먼저 각 열화 이미지와 관련된 데이터셋 전처리 방식을 소개한다. 이후, 학습의 기반이 된 Restormer 모델의 구조와 특징을 설명한 뒤, 인페인팅 학습을 위해 적용된 모델 구조 수정 방법을 논의한다. 마지막으로, 학습된 모델의 성능을 평가하고, 파인튜닝 이후 성능과 비교함으로써 본 연구가 열화 요소 제거를 통해 다양한 분야에서 가져올 잠재적 기여와 의의를 논의한다.

II. 본론

1. 개발 환경 및 도구

본 연구에서는 딥러닝 모델 학습 및 실험을 수행하기 위해 NVIDIA TITAN RTX GPU와 CUDA를 사용했다. 딥러닝 프레임워크로는 PyTorch를 사용하며, 데이터 처리와 모델 학습 지원을 위해 Anaconda 및 torchvision과 같은 오픈소스 라이브러리를 활용했다. 개발 도구로는 VSCode와 Python을 사용하여 코드 작성 및 디버깅을 진행했다.

2. 시스템 흐름도 및 개요

본 연구는 다양한 열화 유형(Denoising, Deblurring, Anti-aliasing, Inpainting)을 복원하기 위한 딥러닝 기반 파운데이션 모델 설계를 목표로 한다. 이를 위해 열화 유형별로 데이터를 분리하여 학습하고, 각 유형에 특화된 Fine-tuning 과정을 통해 최적의 복원 성능을 확보한다.

입력 데이터는 열화 유형에 따라 Denoising, Deblurring, Anti-aliasing, Inpainting으로 구분되며 각각 별도의 디렉토리에 저장된다. 이러한 데이터 구분은 열화 유형별로 처리 방식이 상이하기 때문에 필수적이다.

Inpainting의 경우 랜덤 스크래치 마스크를 활용하여 손상된 이미지를 생성하고, 이를 네트워크 입력과 함께 사용하여 손상된 영역의 복원에 집중할 수 있도록 설계한다.

Restormer를 기반으로 설계된 파운데이션 모델은 여러 열화 유형 데이터를 학습하며 범용적인 복원 능력을 갖춘다. 이후, 열화 유형별로 Fine-tuning을 수행하여 특정 열화에 특화된 성능을 확보한다. Fine-tuning 과정은 각 열화 유형의 데이터셋을 활용하여, 실험적으로 학습 데이터셋 크기 및 Iteration 수에 따른 성능 변화를 평가한다. 이를 통해, 열화 유형별로 차별화된 학습 전략이 모델 성능에 미치는 영향을 분석한다.

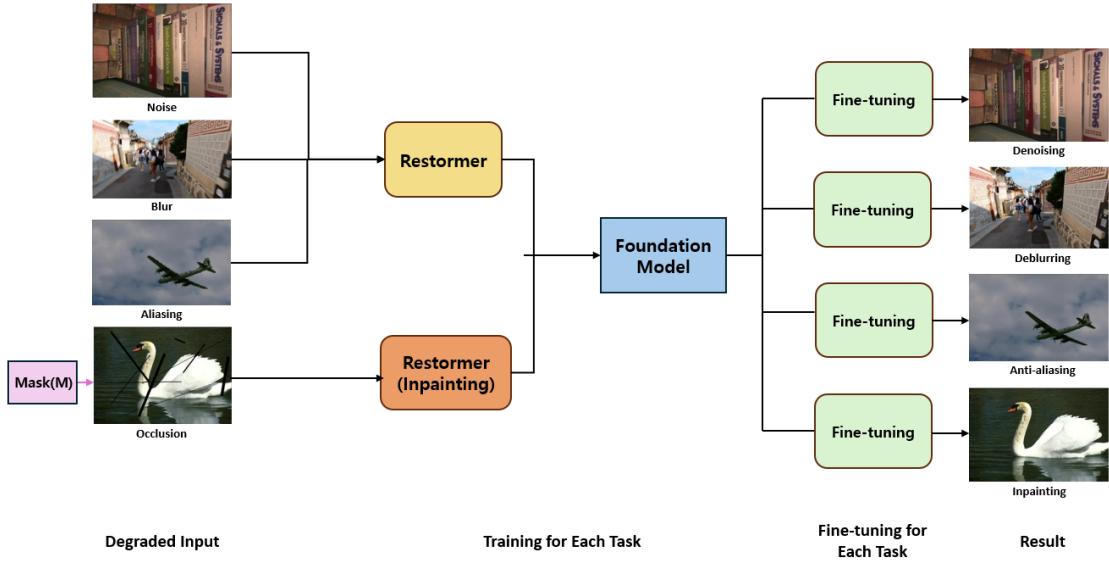


그림 1. 본 연구의 흐름도

3. 파운데이션 모델 및 Fine-tuning 데이터셋 구축

다양한 열화 유형에 대응할 수 있는 파운데이션 모델과 이를 각 열화 유형에 맞게 특화하기 위한 파인튜닝 데이터셋을 구축한다. 데이터셋 구축은 Denoising, Deblurring, Anti-aliasing, Inpainting과 같은 주요 열화 복원 작업에 필요한 데이터의 다양성과 복잡성을 확보하는 데 초점을 맞춘다.

Denoising 작업을 위해 SIDD 데이터셋을 활용한다. 이 데이터셋은 현실 세계에서 발생하는 다양한 노이즈 패턴을 포함하고 있어 노이즈 제거 성능 평가에 적합하다. 추가적으로, CelebA, BSD500, ImageNet 데이터셋에 Poisson Noise, Gaussian Noise, Salt and Pepper Noise를 직접 입력 열화된 이미지를 생성한다.

Deblurring 작업에서는 GOPRO와 REDS 데이터셋을 사용한다. 두 데이터셋 모두 Motion Blur와 같은 현실적인 블러 이미지를 포함하고 있어 모델이 실제 상황에서 발생하는 블러를 효과적으로 복원할 수 있도록 학습할 수 있다. 또한, CelebA, BSD500, ImageNet 데이터셋에 Gaussian Blur와 Motion Blur를 직접 적용하여 추가적인 블러 데이터를 생성한다.

Anti-aliasing 작업을 위해 CelebA, BSD500, ImageNet 데이터셋에 열화를 직접 생성하여 사용한다. Aliasing 열화는 이미지의 해상도를 원래 크기의 30%로 축소한 뒤 다시 원래 해상도로 확대하는 과정을 통해 생성한다. 이 과정을 통해 해상도 변화로 인해 발생하는 특징적인 열화 패턴을 포함한 데이터셋을 구축한다.

Inpainting 작업을 위한 데이터셋은 CelebA, BSD500, ImageNet 데이터셋을 기반으로 한다. 열화 이미지는 기존 데이터셋에 랜덤하게 2~6px 두께의 스크래치를 적용하여 생성한다. 이러한 스크래치는 다양한 크기와 형태를 가지도록 설계되었으며, 모델이 손상된 영역의 복원을 효과적으로 학습할 수 있도록 마스크와 함께 입력으로 제공된다.

파운데이션 모델의 학습에는 총 3000개의 Train 데이터셋과 1700개의 Validation 데이터셋을 사용한다. Train 데이터셋은 Gaussian Noise(400개),

Salt and Pepper Noise(400개), Poisson Noise(400개), Motion Blur(200개), Gaussian Blur(200개), Aliasing(900개), Inpainting(500개)로 구성되며, Validation 데이터셋은 Gaussian Noise(300개), Salt and Pepper Noise(300개), Poisson Noise(300개), Motion Blur(200개), Gaussian Blur(100개), Aliasing(300개), Inpainting(200개)로 구성되었다. 이러한 데이터셋 구성은 모델이 다양한 열화 환경에서도 강건한 복원 성능을 발휘할 수 있도록 하는 중요한 기초 작업이다.

4. Restormer 기반 파운데이션 모델 설계

다양한 열화 유형을 효과적으로 처리할 수 있는 범용적인 딥러닝 기반 파운데이션 모델로 Restormer를 선택하고 이를 수정하여 설계한다. Restormer는 Transformer 기반의 Encoder-Decoder 구조를 갖추고 있으며, Attention Mechanism과 Feed-Forward Network를 통해 이미지 내 전역적 및 국소적 특징을 학습한다. Restormer를 기반 모델로 선택한 이유는 다음과 같다. 첫째, Transformer 구조는 Attention Mechanism을 통해 이미지 전역에서의 장거리 의존성과 복잡한 패턴을 학습할 수 있어 전역적 관계 처리에 강점을 가진다. 둘째, 네트워크 확장성 측면에서 다른 구조(예: UNet, GAN, Diffusion) 대비 Fine-tuning 과정에서 새로운 데이터나 복원 작업에 더 쉽게 적응할 수 있다. 셋째, 입력 데이터의 유연성이 뛰어나 다양한 해상도 및 입력 크기를 지원하며, 이는 Fine-tuning 과정에서 안정적인 학습을 가능하게 한다.

본 연구에서는 Inpainting 작업의 특성을 반영하여 기존의 Attention 모듈을 수정한다. Inpainting의 경우, 손상된 영역에서 Attention Map이 불필요한 정보를 학습하지 않도록 Mask-aware Attention Mechanism을 도입한다. 수정된 Attention 클래스는 마스크 정보를 입력받아 Mask-aware Attention Map을 생성하며, 손상된 영역을 무시하고 유효한 영역에 집중할 수 있도록 설계한다. 반면, Denoising, Deblurring, Anti-aliasing의 경우 기존의 Attention Mechanism을 사용하여 Query, Key, Value만으로 기존의

Attention Map을 생성한다.

Mask-aware Transformer Block은 Attention 모듈과 Feed-Forward Network를 결합하여 특징을 학습하는 구조를 가진다. 이 블록은 마스크 유무에 따라 학습 방식을 달리한다. 마스크가 있는 경우, 손상된 영역의 특징 복원에 집중할 수 있도록 마스크 정보를 Attention 모듈과 Feed-Forward Network에 함께 전달한다. 마스크가 없는 경우에는 전체 이미지의 전역적 특징을 학습하는 기준 구조를 유지한다. Feed-Forward Network는 Attention 모듈에서 학습된 특징을 기반으로 복원 과정을 강화하며, 이는 Restormer 내에서 중요한 역할을 담당한다.

파운데이션 모델 학습 과정에서 발견된 주요 문제점으로는 특정 태스크에 과적합될 가능성과 태스크 간 학습 간섭이다. 이를 해결하기 위해 멀티태스크 학습의 교차 학습 전략을 도입한다. 각 태스크별 학습이 독립적이면서도 균형 있게 진행되도록, 250 iteration 단위로 학습 태스크를 교체하며 총 50,000 iteration까지 학습을 진행한다. 이러한 교차 학습 전략은 열화 유형 간 간섭을 최소화하면서도 모든 태스크에서 균형 잡힌 성능을 확보할 수 있도록 한다.

5. Fine-tuning 설계 및 성능 평가

학습된 파운데이션 모델을 기반으로 열화 유형별 Fine-tuning을 수행하여 각 열화 유형에 특화된 성능을 확보하고자 한다. 파인튜닝 과정은 네 가지 열화 유형(Denoising, Deblurring, Anti-aliasing, Inpainting)에 대해 개별적으로 진행되었으며, 이를 통해 열화 유형별 복원 성능 최적화를 목표로 한다. Fine-tuning 실험은 Train 데이터셋 크기와 Iteration 횟수, 두 가지 기준에 따라 진행한다. Train 데이터셋 크기가 복원 성능에 미치는 영향을 분석하기 위해, 데이터셋 크기를 점진적으로 증가시키며 진행한다. 모든 실험은 동일한 Iteration 수(200회)로 진행되었으며, 데이터셋 크기에 따른 Train/Validation 분리 구성은 아래와 같다:

- Train 10개 / Validation 2개
- Train 20개 / Validation 4개
- Train 40개 / Validation 8개
- Train 80개 / Validation 16개

Iteration 횟수가 성능에 미치는 영향을 평가하기 위해, Train/Validation 데이터 크기를 고정(Train 40개 / Validation 8개)하고, Iteration 횟수를 달리하여 실험을 진행한다. Iteration 설정은 다음과 같다:

- Iteration 100회
- Iteration 200회
- Iteration 300회
- Iteration 400회
- Iteration 500회
- Iteration 600회
- Iteration 700회
- Iteration 800회

III. 연구 결과

1. 정량적 성과

총 3000개의 Train 데이터셋과 1700개의 validation 데이터셋을 사용하여 학습한 파운데이션 모델의 성능은 다음과 같다.

표 1. 파운데이션 모델 성능

	PSNR(dB)	LPIPS
Denoising	32.10	0.1605
Deblurring	26.60	0.2651
Anti-aliasing	24.63	0.2849
Inpainting	33.75	0.0920

파운데이션 모델을 바탕으로 파인튜닝을 진행한 결과는 다음과 같다.

1) Train 데이터셋 크기에 따른 결과

Train 데이터셋의 크기를 10, 20, 40, 80으로 설정하고 동일한 Iteration 수(200회)로 학습을 진행한 결과, 각 열화 유형별로 특정 데이터셋 크기에서 최적의 성능을 보였다.

표 2. Train 데이터셋 크기별 PSNR(dB) 수치

*	Baseline	Train 10	Train 20	Train 40	Train 80
Denoising	32.10	36.69	37.23	38.02	37.99
Deblurring	26.60	<u>27.76</u>	27.75	27.89	27.75
Anti-Aliasing	24.63	24.60	25.04	25.42	<u>25.24</u>
Inpainting	33.75	35.00	35.17	<u>35.08</u>	35.07

표 3. Train 데이터셋 크기별 LPIPS 수치

*	Baseline	Train 10	Train 20	Train 40	Train 80
Denoising	0.1605	0.1177	0.1164	<u>0.0999</u>	0.0898
Deblurring	0.2651	0.2549	<u>0.2547</u>	0.2460	0.2547
Anti-Aliasing	0.2849	0.2642	0.2581	0.2541	<u>0.2562</u>
Inpainting	0.0920	0.0407	0.0382	0.0412	0.0393

Denoising의 경우, Train 데이터셋 크기가 40일 때 평균 PSNR(38.02 dB)과 LPIPS(0.0999)가 가장 우수하였다. 이는 충분한 데이터 크기가 네트워크 학습에 유리하게 작용함을 보여준다. 그러나 데이터셋 크기가 80으로 증가했을 때는 성능 개선이 거의 나타나지 않아, 데이터셋 크기가 일정 수준을 초과하면 학습 효율이 감소할 수 있음을 시사한다.

Deblurring에서도 Train 데이터셋 크기가 40일 때 PSNR(27.89 dB)과 LPIPS(0.2460)로 최적의 성능을 보였다. 데이터셋 크기를 80으로 확장했을 때 성능이 소폭 감소하는 경향이 관찰되었으며, 이는 Fine-tuning 관점에서 적정한 데이터 크기를 유지하는 것이 과적합 방지와 효율적 학습에 중요하다는 것을 의미한다.

Anti-aliasing 태스크에서는 Train 데이터셋 크기 40에서 PSNR(25.42 dB)과 LPIPS(0.2541)로 최적 성능을 나타냈다. 데이터 크기를 계속 확장할 경우 추가적인 성능 개선이 거의 없었으며, 이는 Fine-tuning에서 불필요한 데이터 증가가 모델 효

* 볼드체로 표시된 값은 해당 지표에서의 최고 성능을 나타내며, 밑줄로 표시된 값은 두 번째로 높은 성능을 나타낸다.

율성을 저하할 수 있음을 보여준다.

Inpainting 작업은 Train 데이터셋 크기가 20일 때 PSNR(35.17 dB)과 LPIPS(0.0382)로 가장 높은 성능을 기록하였다. 이는 Inpainting 태스크가 다른 열화 유형에 비해 적은 데이터셋으로도 최적 성능을 달성할 수 있음을 보여준다. 마스크 정보가 네트워크 학습에 효과적으로 작용하여 적은 데이터셋으로도 강력한 복원 능력을 가능하게 했기 때문으로 분석된다.

2) Iteration 횟수에 따른 파인튜닝 결과

파인튜닝 결과, 4가지 태스크 모두에서 Iteration이 증가함에 따라 PSNR과 LPIPS 값은 초기에는 지속적으로 개선되는 경향을 보였으나, 일정 Iteration 이후에는 성능 개선이 둔화되거나 수렴하는 양상을 확인할 수 있었다.

표 4. Iteration 횟수별 PSNR(dB) 수치

*	Denoising	Deblurring	Anti-Aliasing	Inpainting
Baseline	32.10	26.60	24.63	33.75
100	37.39	27.55	25.27	34.75
200	37.57	27.89	<u>25.42</u>	<u>35.08</u>
300	37.67	27.85	24.97	34.79
400	37.92	27.83	25.08	34.74
500	38.05	27.81	24.99	35.28
600	<u>38.19</u>	<u>27.88</u>	25.60	35.00
700	38.25	27.74	24.89	34.98
800	38.00	27.78	25.01	34.92

표 5. Iteration 횟수별 LPIPS 수치

*	Denoising	Deblurring	Anti-Aliasing	Inpainting
Baseline	0.1605	0.2651	0.2849	0.0920
100	0.1045	0.2590	0.2571	0.0451
200	0.0979	0.2460	0.2541	0.0418
300	<u>0.0970</u>	0.2471	0.2562	0.0451
400	0.1023	0.2415	<u>0.2512</u>	0.0451
500	0.0973	0.2462	0.2545	0.0378
600	0.1018	0.2423	0.2488	<u>0.0398</u>
700	0.0932	0.2384	0.2569	0.0426
800	0.1054	<u>0.2406</u>	0.2548	0.0444

* 볼드체로 표시된 값은 해당 지표에서의 최고 성능을 나타내며, 밑줄로 표시된 값은 두 번째로 높은 성능을 나타낸다.

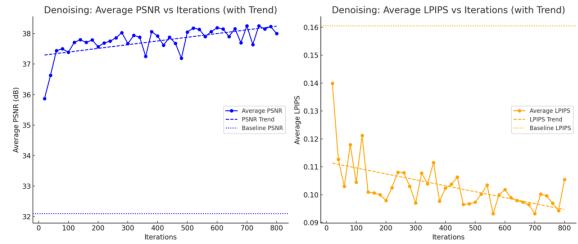


그림 2. Iteration 횟수별 Denoising 파인튜닝 결과 그래프

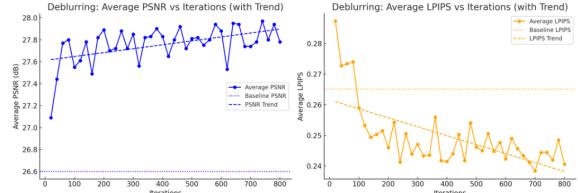


그림 3. Iteration 횟수별 Deblurring 파인튜닝 결과 그래프

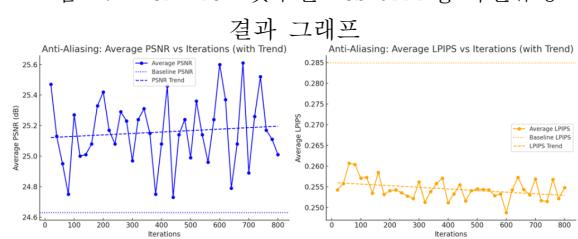


그림 4. Iteration 횟수별 Anti-aliasing 파인튜닝 결과 그래프

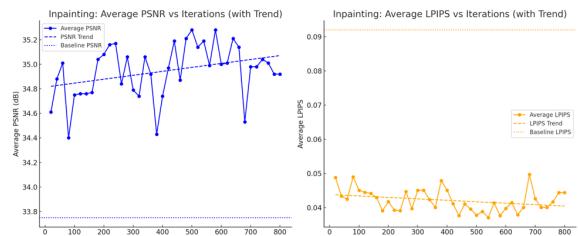


그림 5. Iteration 횟수별 Inpainting 파인튜닝 결과 그래프

Denoising과 Deblurring 태스크는 Iteration 700에서 가장 높은 성능을 기록하며, 충분한 Iteration을 통해 noise와 blur 복원이 효과적으로 이루어졌음을 보여준다. 반면, Anti-aliasing 태스크는 Iteration 600에서 최적의 성능을 보였으며, 추가적인 Iteration에서는 더 이상 성능 향상이 이루어지지 않는 경향을 나타냈다. Inpainting 태스크 또한 Iteration 500에서 최적화되었으며, 이후 Iteration에서는 성능 향상이 둔화되고, 불필요한 학습으로 인해 효율성이 저하될 수 있음을 보여준다.

이는 각 태스크별로 적합한 학습 범위가 상이하며, Iteration을 무작정 증가시키기보다는 태스크별 최적 학습 설정이 필요함을 보여준다.

2. 정성적 성과

표 6을 통해 파운데이션 모델과 각 열화 요소에 대해 파인튜닝한 모델의 복원 결과를 비교할 수 있다. 표에서 Input은 열화 이미지를 의미하며, Foundation model은 사전 학습된 모델을 사용하여 복원한 결과를 나타낸다. Fine-tuned model은 각

열화 요소에 대해 별도로 파인튜닝한 모델을 사용하여 복원한 이미지이다. 파인튜닝된 이미지는 Train 데이터셋 크기와 Iteration 횟수에 따른 성능을 비교한 후, 가장 좋은 결과를 얻었을 때의 이미지를 첨부하였다.

표 6. 이미지 복원 결과

	input	Foundation model	Fine-tuned model
Denoising			
Deblurring			
Anti-aliasing			
Inpainting			

표 7은 파운데이션 모델과 파인튜닝된 모델의 복원 성능을 시각적으로 비교하기 위해 Error Map을 사용한 결과를 나타낸다. 일반적인 시각적 비교만으로 두 모델의 성능 차이를 구별하는 데 어려운 경우가 많았기 때문에, 원본 타겟 이미지(Ground Truth)와 복원된 이미지(Restored Image) 간의 픽셀 차이를 정량적으로 나타내는 Error Map을 활용하였다. 본 연구에서는 Error Map의 시각화를 위해 jet 컬러맵을 적용하였으며, 파란색(낮은 값)은 작은 오차를, 빨간색(높은 값)은 큰 오차를 나타낸다. Error Map의 값은 최소 0에서 최대 1(혹은 다른 범위)로 정규화되어 있으며, 색상은 오차의 상대적 크기를 표현한다.

표 7. 이미지 복원 결과 Error Map

	Foundation model	Fine-tuned model
Denoising		
Deblurring		
Anti-aliasing		
Inpainting		

Error Map을 기반으로 분석한 결과, Fine-tuned Model은 모든 태스크에서 파운데이션 모델보다 전반적으로 낮은 오차를 보이며, 원본과 유사한 복원이 이루어졌음을 확인할 수 있다.

특히, Denoising 태스크에서는 Fine-tuned Model이 원본의 텍스처를 더 잘 유지하면서도 노이즈를 효과적으로 제거하는 성능을 보였다. Deblurring 태스크에서는 경계 부분에서의 왜곡이 줄어들며 보다 선명한 복원이 이루어졌으며, Anti-aliasing 태스크에서는 구조적인 패턴과 경계선에서의 계단 현상이 더욱 자연스럽게 제거되었다. 마지막으로, Inpainting 태스크에서는 결손된 영역을 복원할 때 색감과 질감이 더욱 원본과 유사하게 복원되었으며, 경계 부분에서도 부자연스러움이 감소하였다.

이러한 결과를 종합적으로 고려할 때, Fine-tuned Model은 다양한 열화 요소를 보다 효과적으로 처리할 수 있으며, 전반적인 복원 성능이 향상되었음을 확인할 수 있다.

IV. 결론

본 연구에서는 다양한 영상 열화 문제를 해결할 수 있는 통합적인 파운데이션 모델을 설계하고 개발하였다. 실험 결과, 미리 학습된 파운데이션 모델은 초기 상태에서도 네 가지 태스크(Denoising, Deblurring, Anti-aliasing, Inpainting) 모두에서 안정적인 성능을 보였으며, Fine-tuning 이후 각 태스크에 최적화된 성능을 달성하였다.

특히, 데이터셋 크기에 따른 분석에서는 Train 데이터 크기가 Denoising, Deblurring, Anti-aliasing 태스크에서 각각 40일 때, Inpainting 태스크에서 20일 때 가장 높은 성능을 기록하였다. 또한, Iteration 횟수에 따른 분석 결과, Denoising과 Deblurring 태스크는 Iteration 700에서, Anti-aliasing은 600, Inpainting은 500에서 최적의 성능을 기록하며, 각 태스크마다 최적의 학습 범위가 상이함을 확인하였다. 이는 태스크 별로 적절한 데이터 크기와 Iteration 설정이 상이 하며, 각 태스크의 특성에 맞는 학습 환경과 조건을 설계해야 함을 시사한다.

특히, 파인튜닝은 기존 파운데이션 모델의 성능을 특정 태스크에 최적화하는 과정이므로 적은 데이터셋과 학습 횟수로도 최적 성능을 낼 수 있는 효율적인 학습 전략이 필요하다.

이러한 성과는 본 연구가 다양한 열화 문제를 동시에 해결하고, 실제 환경에서도 적용 가능한 고성능 영상 복원 모델로서의 가능성을 입증했음을 의미한다. 본 연구에서 개발된 파운데이션 모델은 다양한 산업 분야에서 폭넓게 활용될 수 있다. 의료 영상 분야에서는 noise와 blur 제거를 통해 진단의 정확도를 높이고, 자율주행 시스템에서는 영상 복원을 통해 객체 탐지 및 경로 예측의 신뢰성을 강화할 수 있으며, AR/VR 콘텐츠 제작에서는 이미지의 품질을 향상시켜 몰입감 있는 사용자 경험을 제공할 수 있다.

또한, 연구 결과는 다양한 데이터셋과 열화 유형에 대한 통합적 학습 방법론을 제공하여, 향후 딥러닝 기반 영상 복원 연구에 중요한 참고 자료로 활용될 수 있다. 나아가, 본 모델은 범용성을 갖추고 있어 다양한 새로운 열화 문제에 대한 빠른 적응과 확장이 가능하며, 다양한 분야에서 신뢰할 수

있는 영상 복원 솔루션으로 자리잡을 수 있다.

참고문헌

- [1] Zamir, S. W., Arora, A., Khan, S., Hayat, M., Khan, F. S., Yang, M. H., "Restormer: Efficient Transformer for High-Resolution Image Restoration," *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5728-5739, 2022.



김지민

2025년 2월 인하대학교 정보통신공학과 졸업 예정.
관심분야는 컴퓨터 비전, 웹 개발, 컴퓨터 네트워크.



양유진

2021년~현재 인하대학교 정보통신공학과 재학 중.
관심분야는 컴퓨터 비전, AI, 웹 개발