

Integrated Super-Resolution Deblurring Model Using Domain Adaptation and Multi-Task Learning

Jae-Yeob Kim*, Ji-Su Yoon**, Hyun-Sung Jang***, Seung-Wook Park***

*Senior Research Engineer, EO/IR Systems R&D, LIG Defense&Aerospace Co., Ltd., Pangyo, Korea

**Research Engineer, EO/IR Systems R&D, LIG Defense&Aerospace Co., Ltd., Pangyo, Korea

***Chief Research Engineer, EO/IR Systems R&D, LIG Defense&Aerospace Co., Ltd., Pangyo, Korea

[Abstract]

In this paper we propose an integrated super-resolution-deblurring model based on multi-task learning that employs the encoder of a deblurring network as a shared encoder. Images captured from highly dynamic platforms inevitably contain motion blur, which makes it difficult to obtain clear results when performing various image-enhancement tasks such as super-resolution or de-hazing. By jointly learning super-resolution and deblurring within a single architecture, we develop a deep-learning model capable of producing sharp, high-resolution images even in the presence of severe blur. To address the scarcity of training data in the infrared (IR) domain, we incorporate a domain-adaptation technique that aligns feature distributions between the source and target domains, thereby mitigating performance degradation caused by limited IR datasets. Finally, to achieve real-time inference (≥ 30 fps) on an on-device platform, we compress the network using knowledge distillation. The distilled model runs at real-time speed for 1K-resolution inputs on an NVIDIA RTX A4500 Embedded GPU.

▶ **Key words:** Image Restoration, Deep Learning, Super-Resolution, Deblurring, Multitask-Learning, Domain Adaptation, Knowledge Distillation

[요 약]

본 논문에서는 디블러링 모델 인코더를 공유 인코더로 활용하여 멀티태스크 러닝 기반 초해상도-디블러링 통합형 모델을 제안한다. 기동이 심한 플랫폼 환경에서 획득된 영상은 번짐 현상(Blur)이 존재하여 초해상도, 디헤이징 등 다양한 영상 개선 추론 시 번짐을 동반하여 선명한 영상을 얻기 어려운 상황이다. 따라서, 초해상도-디블러링 통합형 모델 개발로 번짐이 존재하는 환경에서도 선명한 고화질의 초해상도 영상을 획득할 수 있는 딥 러닝 모델을 개발하였다. 또한, 비교적 데이터셋이 적은 IR 도메인에서 성능 향상을 수행하기 위해 도메인 적응 기법을 사용하여 데이터셋에 의한 성능 문제를 해결하였다. 마지막으로, 온 디바이스(On-Device) 환경에서 실시간(30fps 이상)으로 추론하기 위해 지식 증류(Knowledge Distillation) 기반의 경량화를 수행하여 NVIDIA RTX A4500 Embedded에서 1K 해상도 기준 실시간을 달성하였다.

▶ **주제어:** 영상 개선, 딥러닝, 초해상도, 디블러링, 멀티태스크 러닝, 도메인 적응, 지식 증류

- First Author: Jae-Yeob Kim, Corresponding Author: Seung-Wook Park
- *Jae-Yeob Kim (jaeyeob.kim@ligdna.com), EO/IR Systems R&D, LIG Defense&Aerospace Co., Ltd.
- **Ji-Su Yoon (jisu.yoon25@ligdna.com), EO/IR Systems R&D, LIG Defense&Aerospace Co., Ltd.
- ***Hyun-Sung Jang (hyunsung.jang@ligdna.com), EO/IR Systems R&D, LIG Defense&Aerospace Co., Ltd.
- ***Seung-Wook Park (seungwook.park@ligdna.com), EO/IR Systems R&D, LIG Defense&Aerospace Co., Ltd.
- Received: 2026. 04. 28, Revised: 2026. 05. 22, Accepted: 2026. 06. 04.

I. Introduction

최근 딥 러닝 기술을 사용하여 컴퓨터 비전 분야에서 많은 연구가 이루어지면서 객체 탐지/추적뿐만 아니라 고성능의 영상 획득을 위해 영상 개선 분야에서도 초해상도(Super-Resolution), 디블러링(Deblurring), 디헤이징(Dehazing), 디노이징(Denoising) 등 다양한 연구가 이루어지고 있다. 대다수의 기존 학계 딥 러닝 모델은 각 태스크 별 초해상도, 디블러링, 디헤이징 등 단일 모델에서의 추론 성능 향상을 위한 연구가 주로 이어지고 있다. 하지만 항공기 등 플랫폼의 기동, 진동이 심한 온 디바이스(On-Device) 환경에서 일반적인 학계에서 사용하는 단일 모델 기반의 초해상도, 디헤이징 등 영상 개선을 적용할 경우 플랫폼 이동에 따른 번짐 현상(Blur)이 반영되어 선명한 초해상도 결과를 얻을 수 없다. 플랫폼의 기동, 진동이 심한 온 디바이스 환경에서도 선명한 초해상도 영상을 획득하기 위해 번짐이 제거된 영상 혹은 특징맵(Feature Map) 기반의 추론이 가능한 딥러닝 모델 개발이 필요하다.

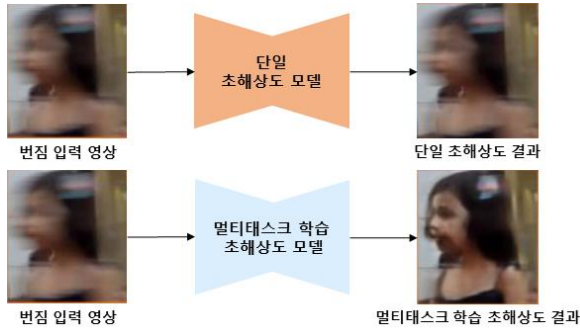


Fig. 1. Comparison of Results Between a Single Super-Resolution Model and a Multi-Task Learning Based Super-Resolution Model

본 논문에서 핵심적인 기여로 앞서 제시한 문제를 해결하도록 기동이 심한 플랫폼에서 얻어지는 번짐 입력 영상에서 번짐이 제거된 초해상도 영상을 획득하기 위해 멀티태스크 러닝(Multi-Task Learning) 기반 초해상도-디블러링 통합형 딥 러닝 모델을 제안한다. 초해상도-디블러링 통합형 모델로 기존 단일 모델에서 해결하지 못한 번짐 입력 시 번짐 초해상도 결과를 디블러링 인코더를 공유 인코더로 활용하여 문제를 해결하였다. 또한, 다중파장대역 실시간 추론 시스템에 적용하기 위해 도메인 적응과 지식 증류(Knowledge Distillation) 기반의 모델 경량화 방법을 추가로 적용하였다.

상세한 모델의 구조로는 제안 초해상도-디블러링 통합형 모델은 U-Net [1] 구조를 가지는 디블러링 모델의 인코더(Encoder)를 공유 인코더로 사용하여 출력을 초해상도, 디블러링 두 가지 태스크를 추론 가능하도록 설계, 개발하였다. 초해상도 추론 시 입력 영상은 공유 인코더를 거친 후 초해상도 브랜치에 입력 채널에 맞춰 특징맵을 전달 받아 초해상도를 추론한다. 디블러링은 추론 시 초해상도 브랜치를 거치지 않고 단일 모델처럼 활용이 되어 추론이 가능하다. 결과적으로, 단일 입력 기반의 초해상도-디블러링 다중 출력을 수행하는 통합형 모델을 개발하였다.

또한, 본 연구는 EO/IR 두 도메인에 대해 하나의 모델로 처리하도록 개발되었다. 다량의 데이터셋이 존재하는 RGB, EO 도메인과 다르게 데이터셋이 비교적 부족한 IR 도메인의 성능 향상을 위해 도메인 적응(Domain Adaptation) 기법을 적용하여 IR 도메인의 성능을 향상시켰다. 다량의 데이터셋을 가진 RGB, EO로 선행 학습을 수행한 후 적대적 학습 기반 특징맵 추출부가 특정 도메인에 일관되는 특징맵을 추출하지 않도록 하여 다중 대역 영상의 일반적인 추론 결과를 이루어냈다.

그리고, 기존 학계에서는 초해상도의 경우 모바일에서 사용할 수 있는 저전력 저비용의 많은 딥러닝 모델이 개발되었지만 디블러링 등 다른 태스크에서는 초해상도에 비해 많은 연구가 이루어지지 않았다. 본 연구를 적용하는 시스템은 온 디바이스 환경, 1K 해상도에서 실시간으로 동작해야 하는 제약 조건을 가지고 있다. 따라서, 온 디바이스 환경에서 실시간으로 동작하기 위해 Teacher-Student 기반의 지식 증류 기법을 활용하여 경량화를 수행하였다. 지식 증류 기법은 거대한 Teacher 모델로부터 가벼운 Student 모델이 지식을 습득하는 방법으로 딥러닝 모델 경량화에서 많이 사용되고 있는 기법이다. 결과적으로 멀티태스크 러닝 기반 초해상도-디블러링 통합형 모델을 지식 증류 기법으로 경량화를 수행하고 NVIDIA RTX A4500 Embedded에서 1K 영상에서 초해상도, 디블러링 두 태스크 모두 실시간을 달성하였다.

II. Preliminaries

1. Related works

1.1 MIMO-UNet [2]

본 논문에서는 멀티태스크 러닝 기반 통합형 모델의 기본이 될 딥 러닝 디블러링 모델을 최신 연구에서 추론 시간에 장점이 있는 MIMO-UNet을 기반으로 설계하였다.

MIMO-UNet은 기존 멀티 스케일(Multi-Scale) 방식으로 연구되었던 디블러링 연구들이 높은 연산 비용의 문제를 가지는 것을 해결하기 위해 하나의 UNet 내부에서 효율적으로 다중 스케일 정보를 처리하는 구조를 채택하였다. 큰 구조로는 입력 이미지를 특징맵으로 변환하는 초기단계에서 하위 스케일의 정보를 효율적으로 추출하는 SCM(Shallow Convolutional Module), 멀티 스케일의 정보를 융합하여 고해상도와 저해상도 정보를 융합하는 AFF(Asymmetric Feature Fusion), 공간적 채널(Spatial Channel)에서 윤곽선, 객체의 질감 등 중요한 특징을 추출할 수 있는 FAM(Feature Attention Module)로 구성되어 있다.

또한 번짐 영상의 경우 객체의 윤곽선, 물체의 질감 등 고주파수단(High-Frequency) 부분이 흐려지는 현상이 나타나는데 MIMO-UNet은 이를 해결하기 위해, FFT(Fast Fourier Transform) 변화를 사용한 주파수 도메인 손실 함수(Frequency Domain Loss Function)를 추가적으로 설계하여 디블러링 추론 후에 객체에 고주파수단을 보존하였다.

1.2 SPAN [3]

본 논문에서는 초해상도 브랜치에 사용할 모델로 추론 시간에 장점이 있는 SPAN 기반으로 설계하였다. SPAN은 기존 초해상도의 고질적 문제인 복잡한 어텐션 매커니즘(Attention Mechanism)으로 인한 추론 속도 저하 문제를 해결하여 추론 시간에 장점이 있다. 주요 특징으로는 Parameter-free Attention Mechanism을 사용하여 추가적인 파라미터 학습 없이 활성화 함수 시그모이드만으로 특징맵의 중요한 부분과 중요하지 않은 부분을 구분하는 어텐션 기법을 활용하였다. 본래 시그모이드는 0~1 사이의 범위로 값을 맵핑하지만 SPAN은 -1~1로 스케일링 변환을 통해 윤곽선 및 질감 등 기여도가 높은 정보는 Positive로 강화하고 노이즈 등 기여도가 낮은 정보는 Negative로 억제하도록 설계하였다. 결과적으로, CNN (Convolutional Neural Network)의 레이어 증가나 복잡한 구조 없이도 초해상도 성능을 확보하여 빠른 추론 속도를 보였다.

1.3 Domain Adaptation Using DANN [4]

본 논문에서는 RGB 혹은 EO와 같이 일반적인 파장대역의 영상보다 데이터가 다소 부족한 IR(Infra-Red) 파장대역의 영상의 추론 성능을 향상 시키기 위해 DANN [3]을 적용하여 도메인 적응 기법을 사용하였다. DANN은 데이터가 충분한 기존 데이터셋인 소스 도메인과 데이터가 충분하지 않은 타겟 도메인을 구분하여 학습을 수행한다.

DANN은 도메인 분류기를 특징 추출자(Feature Extractor)의 뒷단에 적용시켜 역전파(Backpropagation) 시 그래디언트 역 레이어(Gradient Reverse Layer)를 거쳐 그래디언트 부호를 바꾸어 역전파 시킨다. 이 때 역전된 값으로 특징 추출자 모듈을 학습하여 특징 추출자가 도메인을 잘 구분하지 못하게 한다.

결과적으로, 적대적 학습 기법으로 모델의 인코더가 EO 데이터, IR 데이터의 특징을 추출할 때, 두 도메인의 특징 맵 분포를 유사하게 맵핑하여 추론 시 다중 파장대역의 성능을 보장할 수 있도록 한다.

1.4 Knowledge Distillation Using DIPNet [5]

딥 러닝 모델 경량화에서 지식 증류 방법은 프루닝, 양자화 등과 더불어 가장 보편적으로 검증된 방법으로 거대한 Teacher 모델과 경량화된 Student 모델로 구성되어 있다. 프루닝, 양자화는 기존 모델 구조에 기반하여 불필요한 가중치(Weight)를 제거하기 때문에 성능이 심히 저하될 우려가 있다. 반면에, 지식 증류는 거대하게 학습된 Teacher 모델을 기반으로 가벼운 Student 모델의 특징맵을 학습하는 방법으로 프루닝, 양자화와 비교하였을 때 성능 저하 우려가 적다. 지식 증류에서 지식은 거대하게 학습된 Teacher 모델의 특징맵을 의미하며 이 특징맵의 분포를 Student 모델이 학습하는 것이 목표이다.

본 논문에서는 멀티태스크 러닝 기반 통합형 거대 모델을 학습 완료 후 초해상도에 적용한 DIPNet [4] 기반의 지식 증류 방법을 활용하여 경량화를 수행하였다. DIPNet은 Teacher 모델의 출력과 중간 레이어에서 추출되는 특징맵을 Student 모델이 모두 학습할 수 있도록 구성되어 얇은 레벨의 특징맵부터 깊은 레벨의 특징까지 Student 모델이 학습할 수 있는 장점이 있다.

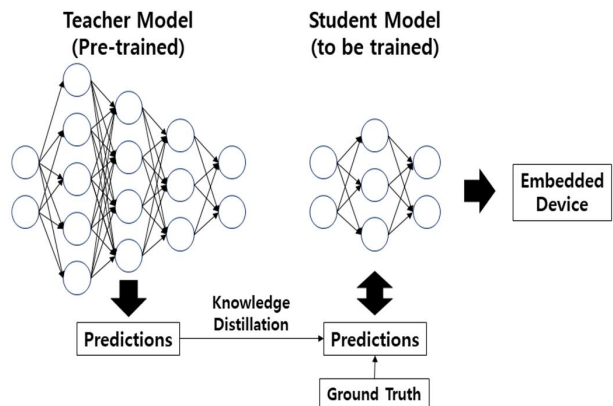


Fig. 2. Knowledge Distillation Method

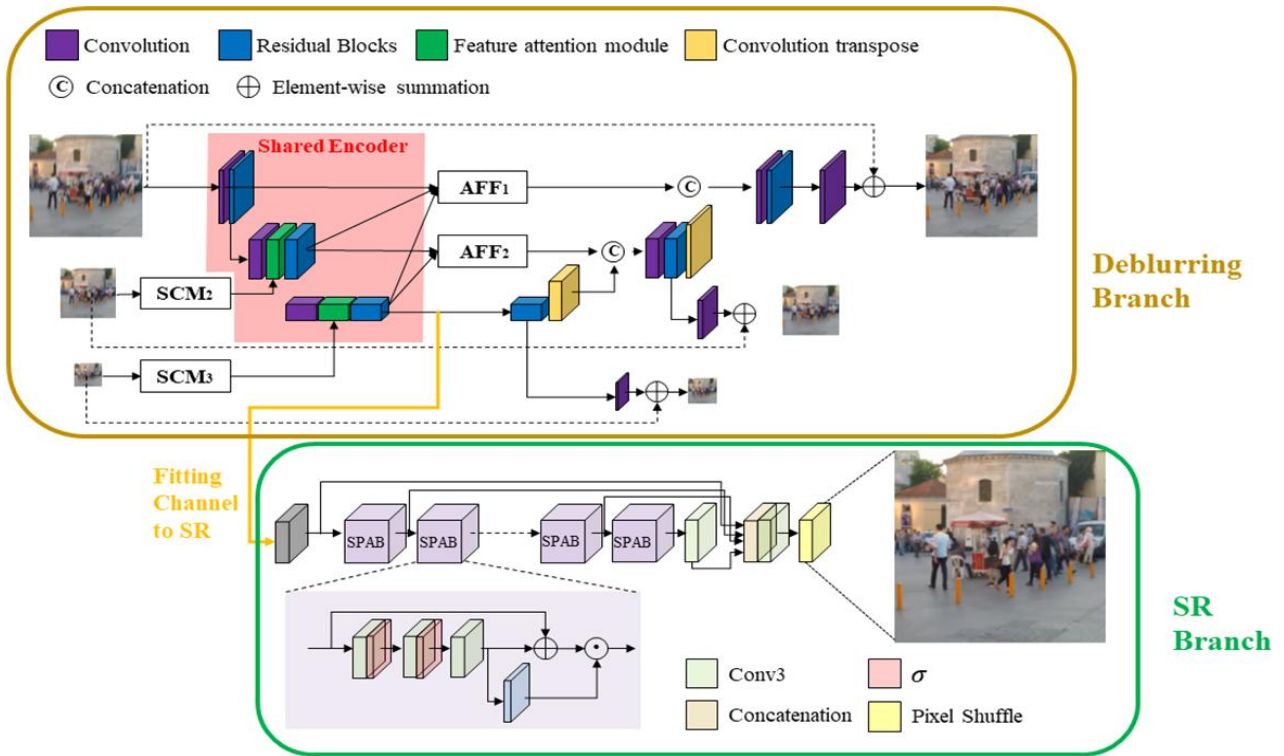


Fig. 3. Proposed Integrated Super-Resolution and Deblurring Model Architecture.

III. The Proposed Scheme

1. Proposed Model Architecture

1.1 Multi-Task Learning Based Integrated Model

본 논문에서는 멀티태스크 러닝 기반으로 MIMO-UNet과 SPAN을 결합한 통합형 딥러닝 초해상도, 디블러링 모델을 제안한다. 멀티태스크 러닝 방법을 적용하여 번짐 입력 영상에 대한 초해상도 추론 시 번짐이 제거된 초해상도 결과를 얻기 위함으로 디블러링 인코더를 공유 인코더로 사용하였다.

MIMO-UNet의 인코더를 Fig. 3의 Shared Encoder로 초해상도와 디블러링 모두 사용하는 공유 인코더로 설계하였다. 디블러링 모델 MIMO-UNet의 인코더를 공유 인코더로 사용한 이유는 해당 연구는 번짐이 제거된 초해상도 결과를 얻기 위함이기 때문이다.

MIMO-UNet은 멀티스케일의 구조로 여러 계층으로 구성되어 있다. 본 논문에서는 디블러링 디코더로 전달되는 특징맵에서 브랜치를 두어 디블러링 특쟁맵을 초해상도의 입력으로 활용한다. 공유 인코더에서 초해상도 브랜치의 입력으로 주어지는 특징맵은 초해상도 모델의 입력 특징맵에 맞도록 컨볼루션 연산을 통해 차원 및 채널 수를 맞춰주었다(Fig. 3에 Fitting Channel to SR).

초해상도 태스크를 수행할 때, 공유 인코더를 통해 디블

러링에 대한 특징맵이 추출된 후 특징맵을 초해상도 브랜치의 입력에 채널을 맞추어 입력으로 사용한다. Fig. 4의 초록색 SR Inference는 초해상도 태스크 수행 시 통합 모델이 동작하는 부분을 도식화한 결과이다.

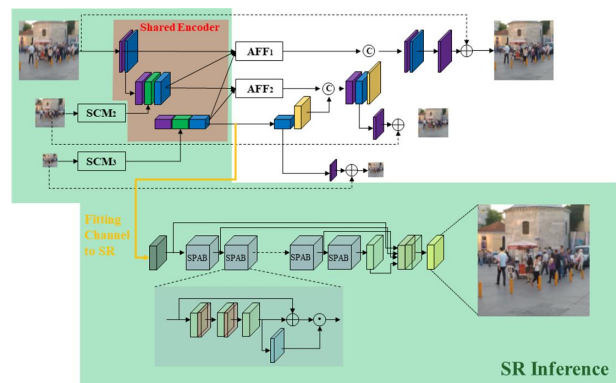


Fig. 4. Super-Resolution Inference Flow.

2. Model Training Strategy

본 논문에서는 모델을 학습시키기 위해 총 3가지 방법을 거친다. 공유 인코더를 가지는 멀티태스크 러닝 기반 초해상도-디블러링 통합형 모델을 학습하기 위한 Two-Phase training을 첫 번째 학습으로 모델을 개발한다. 그리고, EO-IR 도메인에서 모두 동작 가능하도록 타겟 도메인 IR 도메인의 성능 향상을 위해 도메인 적응 기법을

적용한다. 마지막으로, 실시간으로 임베디드 장비 플랫폼에서 동작하기 위해 Teacher-Student 기반의 지식 증류 기반 경량화를 수행한다.

Fig. 5.는 제안하는 초해상도-디블러링 통합형 모델을 학습하기 위한 방법의 전체 흐름도이다. 각각 학습된 초해상도와 디블러링 모델을 기반으로 Two-Phase training 방법으로 Fine-Tuning을 수행한다. 이후에 EO-IR 도메인에서 모두 동작 가능하도록 도메인 적응 기법을 적용하고 최종적으로 지식 증류 기법을 적용하여 경량화를 수행하는 순서로 학습한다.

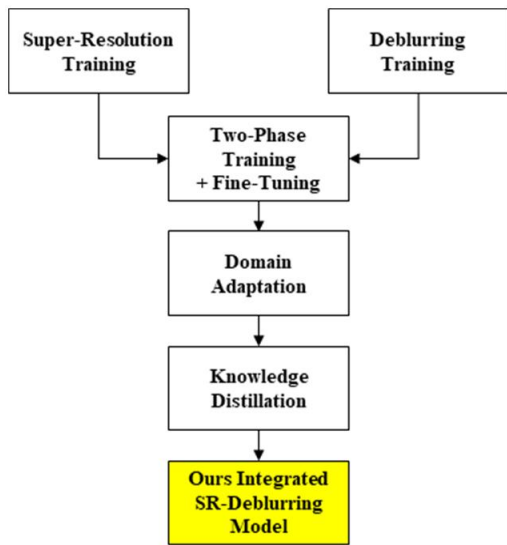


Fig. 5. Flowchart of the overall training strategy.

2.1 Two-Phase Training Method

본 연구는 기동이 심한 플랫폼에서 얻어지는 번짐 영상의 초해상도 수행 시 번짐이 반영되는 문제를 해결하기 위해 디블러링 공유 인코더를 사용하여 멀티태스크 러닝 기반 통합형 모델을 구성하였다. 따라서, 학습을 수행하는 방법에 따라 특징맵의 추출에 영향을 줄 수 있어 공유 인코더는 번짐을 개선하는 특징맵을 가질 수 있도록 학습을 수행하였다.

본 논문에서는 경험적으로 Two-Phase training 학습 방법을 경험적으로 선택하였다. Two-Phase training 방법을 사용한 근거는 다음과 같다. 디블러링과 초해상도 모두 한 번에 학습하는 기존 Single-Phase training 학습 방법은 통합 모델의 공유 인코더가 디블러링과 초해상도의 모든 특징을 학습하여 일관된 번짐 개선에 대한 특징맵을 추출하지 못하는 방향으로 학습되었다. 또한, Single-Phase training의 경우 학습 시 손실함수에 대한 역전파 과정이 전체 모델에 거쳐 수행되기 때문에 역전파

과정이 길고 두 가지 동일 태스크를 동시에 학습하는 측면에서 각 태스크 디블러링, 초해상도 모두 성능 저하가 발생하는 문제가 있었다.

Two-Phase training 학습 방법은 총 두 번의 학습 방법을 바탕으로 멀티태스크 러닝 기반 초해상도 디블러링 통합 모델을 학습하는 방법이다. Fig. 6.는 Two-Phase training 학습 방법에 대한 도식화이다. Fig. 6.의 Phase1으로 공유 인코더를 가진 디블러링 모델을 우선 학습을 수행하고 초해상도 브랜치는 사용하지 않고 학습한다. Phase1 학습이 종료되면 Phase2로 초해상도 브랜치 학습을 수행한다. 이 때, 디블러링 브랜치는 가중치를 Freeze한 후 공유 인코더로 학습이 끝난 번짐에 대한 특징맵을 초해상도 브랜치의 입력으로 주어 학습을 수행한다.

결과적으로 본 연구의 목적인 번짐 입력 영상에 대해 번짐이 제거된 초해상도 결과를 얻기 위해 번짐이 개선된 특징맵이 초해상도의 입력으로 주어져야 하는데 공유 인코더는 디블러링 태스크 시에만 학습을 수행하고 초해상도 학습 시 가중치를 고정시키는 방향으로 Two-Phase training 방법을 사용하였다.

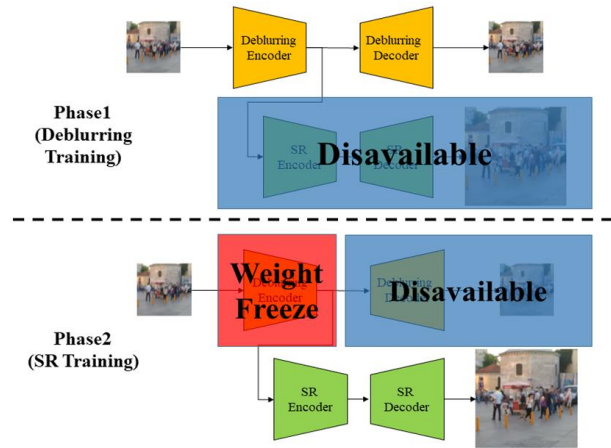


Fig. 6. Two-Phase training method.

2.2 EO/IR Optimization Based-On Domain Adaptation

본 논문에서는 대량의 학습 데이터를 가진 RGB 혹은 EO 도메인으로 학습된 모델에서 적외선 영상인 IR 도메인에서의 성능 확보를 위해 DANN 기반 도메인 적응 기법을 적용한다. 도메인 적응 적용 시에 EO를 소스 도메인, IR을 타겟 도메인으로 설정하고 IR은 Ground-Truth가 없다는 가정하여 각 태스크의 인코더 특징맵에 대해 적대적 학습을 적용한다. 따라서, 학습 환경 셋팅은 디블러링과 초해상도 모두 EO 기반에 데이터셋만 활용하여 적대적 학습을 수행한다.

2.3 Model Compression Based-On Knowledge

Distillation

본 논문에서는 거대한 파라미터를 가지는 Teacher 모델과 비교적 가벼운 파라미터를 가지는 Student 모델 기반으로 경량화를 수행하는 지식 증류 기법을 적용한다. 장비에 실시간 딥러닝 모델을 적용하기 위해서는 경량화는 필수적이다.

초해상도-디블러링 통합형 모델에서 초해상도 모델인 SPAN의 경우 이미 매우 적은 파라미터로 높은 초해상도 복원 성능을 보이기 때문에 따로 경량화를 수행하지는 않았다. MIMO-UNet은 복잡한 번짐에 대한 문제를 해결하는 디블러링 모델 특성상 많은 파라미터를 가진 무거운 모델이기 때문에 MIMO-UNet의 블록 채널 수를 절반으로 경량화하여 Student 모델을 정의하였다.

지식 증류 기법 적용 시 총 3가지 손실 함수(Loss Function)를 적용한다. Feature Loss는 특징맵의 분포를 Student 모델이 Teacher를 따라갈 수 있도록 각 태스크별 인코더 단의 특징맵을 비교한다. Pseudo Loss는 Teacher 모델의 출력 결과를 Student 모델이 모사할 수 있도록 Teacher 모델의 출력을 정답으로 하여 Student 모델의 출력과 비교한다. 마지막으로, Student 모델의 출력을 기존 데이터셋의 Ground-Truth와 비교하는 GT Loss이다. 각 손실함수는 별도의 가중치를 다르게 두지 않고 단순히 덧셈 연산을 통해 역전파를 수행하였다.

IV. Experimental Results

본 실험 결과의 하드웨어 스펙은 다음과 같다. CPU는 Intel® Xeon® W11865MRE Octa-Core 2.6GHz를 사용하였으며, 딥러닝 추론을 위한 GPU는 임베디드 환경에서 사용되는 NVIDIA RTX A4500을 사용하였다. 실시간성 확인을 위한 추론 시 영상의 해상도는 디블러링의 입출력 해상도는 1240x984이며, 초해상도는 4배 업스케일링으로 개발하여 입력 해상도는 310x246, 출력 해상도는 1240x984이다.

학습 시 사용한 파라미터는 다음과 같다. 옵티마이저는 ADAM(Adaptive Moment Estimation)을 사용하였으며, 학습률은 1e-4로 Two-Phase training과 Fine-Tuning을 수행할 때 사용하였으며 총 2000에폭을 학습하였다. 추론 시 실제 시스템에서의 추론과 같이 하기 위해 배치 사이즈는 1로 고정하여 추론하였으며 FPS 측정도 배치 사이즈 1 기준으로 측정하였다.

1. Public and Ours Test Dataset

1.1 Super-Resolution Test Dataset

본 논문에서는 모델의 초해상도 성능 검증을 위한 데이터셋으로 URBAN100 [6]과 VGTSR [7]을 사용하였다. URBAN100은 RGB, EO 대역의 성능을 평가하기 위한 데이터셋으로 도심의 각 건물, 다양한 배경으로 이루어져 있다. 또한, 패턴이 다양한 이미지로 구성되어 초해상도 성능을 검증하기에 적합하다. VGTSR의 경우 항공 뷰로 취득된 IR 대역의 이미지로 IR 대역의 초해상도 성능을 평가하기 위해 적합한 데이터셋이다. 해당 연구에서는 하나의 모델로 URBAN100과 VGTSR 데이터셋의 추론 결과로 성능을 검증한다. URBAN100은 추론용 데이터셋으로 이미지 100장, VGTSR은 학습 데이터셋은 900장, 추론용 데이터셋은 225장으로 구성되어 있다.

1.2 Deblurring Test Dataset

본 논문에서는 모델의 디블러링 성능 검증을 위한 데이터셋으로 GOPRO [8]와 자체 검증용으로 VGTSR에 모션 블러를 추가하여 VGTDB라는 자체 데이터셋을 구축하였다. GOPRO는 RGB, EO 대역의 성능을 평가하기 위한 데이터셋으로 디블러링에서 가장 보편적으로 사용되는 평가 데이터셋이며 프레임은 여러 장 겹쳐 블러 효과를 적용하였다. 해당 연구에서는 하나의 모델로 GOPRO와 VGTDB 데이터셋의 추론 결과로 성능을 검증한다.

Fig. 7.은 VGTDB 생성 과정에서 모션 블러를 생성하여 적용하기 전과 모션 블러를 생성하여 적용한 후의 비교 결과이다. 모션 블러를 생성한 결과를 모델의 입력으로 하여 추론을 수행하고 성능을 검증한다. GOPRO 데이터셋의 경우 학습용 2103장, 추론용 111장으로 구성되어 있으며 자체 데이터셋인 VGTDB는 VGTSR의 영상에 모션 블러를 추가하여 225장으로 구성되어 있다.

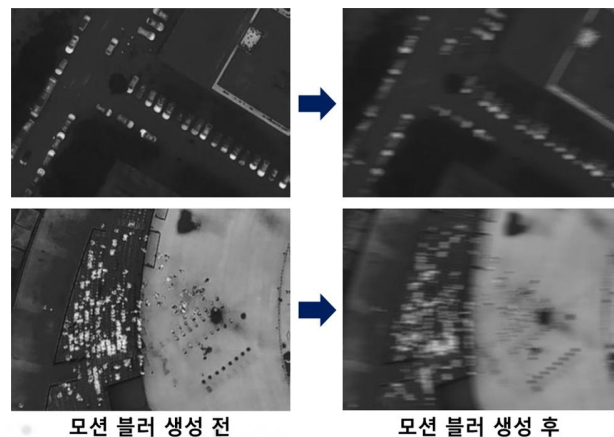


Fig. 7. Comparison Before and After Applying Motion Blur.

2. Model Optimization Performance Results

2.1 Domain Adaptation Quantitative Result

본 논문에서는 DANN 기반 적대적 도메인 적응 방법을 적용하여 타겟 도메인인 IR 대역의 성능을 확보한다. 도메인 적응의 경우 타겟 도메인의 성능 향상이 중점이기 때문에 소스 도메인의 성능이 일부 하락할 우려가 있다. 본 연구에서는 소스 도메인의 하락을 최소화할 수 있도록 실험을 통해 학습률(Learning Rate)을 경험적으로 $1e-5$ 로 설정하여 최종 학습을 수행하였다.

Table 1.은 도메인 적응 학습률별 성능 수치를 나타낸 표이다. Original을 기준으로 각각 학습률을 $1e-4$, $1e-5$ 를 적용하여 결과를 비교하였다. 결과적으로, 학습률을 $1e-5$ 로 설정하여 도메인 적응 기법 적용하였을 때 RGB 도메인의 적은 성능 하락을 보이고 IR 도메인의 결과인 VGTSR, VGTDB의 성능이 향상된 결과를 보였다. 따라서, 지식 증류와 함께 최종 모델 개발 학습 시 DANN 기반 도메인 적응 기법을 학습률 $1e-5$ 로 적용하여 개발하였다.

Table 1. Domain Adaptation Experimental Results

Method	Dataaset	Super-Resolution (x4)	Deblurring
Original	URBAN	25.86 / 0.84	-
	GOPRO	-	32.72 / 0.95
	VGTSR	30.85 / 0.92	-
	VGTDB	-	31.84 / 0.94
DANN (lr:1e-4)	URBAN	25.8 / 0.83	-
	GOPRO	-	33.16 / 0.95
	VGTSR	30.89 / 0.92	-
	VGTDB	-	32.06 / 0.94
DANN (lr:1e-5)	URBAN	25.84 / 0.84	-
	GOPRO	-	33.18 / 0.95
	VGTSR	31.00 / 0.92	-
	VGTDB	-	31.98 / 0.94

2.2 Final Model's Quantitative Result Based on Knowledge Distillation

본 논문에서는 MIMO-UNet과 SPAN을 결합하여 구성 본 논문에서는 MIMO-UNet과 SPAN을 결합하여 구성된 통합형 모델을 Teacher 모델로 정의하였다. 초해상도 브랜치로 사용한 SPAN은 충분한 경량 네트워크이기 때문에 SPAN은 제외하고 MIMO-UNet의 특징맵 파라미터를 절반으로 줄인 모델을 Student 모델로 정의하였다.

지식 증류 기법은 최종적으로 모델을 경량화하여 추론 시간을 확보하기 위한 목적이기 때문에 Teacher 모델을 정의 후 Fine-Tuning, 도메인 적응 등 모든 과정을 거친 후 최고의 성능을 확보하여 적용한다. Table 2.는

Teacher 모델을 기반으로 Fine-Tuning, 지식 증류 기법 적용과 도메인 적응을 모두 적용한 결과이다.

결과적으로 Teacher 모델의 파라미터 9.04M에서 Student 모델 파라미터 3.94M으로 약 2.3배로 절반 이상 감소하는 결과를 보였다. 추론 속도 역시 Teacher 모델은 27.9FPS, Student 모델은 37FPS로 약 1.3배로 향상되어 실시간성을 확보하였다.

Table 2. Experimental Results of the Final Model after Applying Knowledge Distillation

Method	Dataset	Super-Resolution (x4)	Deblurring	Time
Teacher Model	URBAN	25.86 / 0.84	-	27.9 FPS
	GOPRO	-	32.72 / 0.95	
	VGTSR	30.85 / 0.92	-	
	VGTDB	-	31.84 / 0.94	
Student Model + Fine-Tuning + DANN	URBAN	25.84 / 0.84	-	37 FPS
	GOPRO	-	31.64 / 0.94	
	VGTSR	30.81 / 0.92	-	
	VGTDB	-	32.51 / 0.95	

3. Final Super-Resolution Deblurring Model's Quantitative/Qualitative Results

Fine-Tuning, 도메인 적응 및 지식 증류를 포함한 최적화가 완료된 모델이 최종 모델이다. 본 장에서는 초해상도와 디블러링의 정량적/정성적 결과를 평가한다. 특히 본 연구에서 중점으로 개발했던 번짐 영상에 대해 초해상도 추론 결과를 분석한다.

3.1 Final Super-Resolution Deblurring Model's Quantitative Results

최종 초해상도-디블러링 통합형 모델의 정량적 결과 비교는 Table 2.와 같다. Student Model의 경우 디블러링 모델의 인코더를 2배 파라미터를 줄였기 때문에 기존에 학습 데이터셋이었던 GOPRO에서 성능 하락이 있었던 것으로 분석된다. 또한, Fine-Tuning 및 DANN을 적용하였을 때 자체 제작 데이터셋인 VGTDB의 성능은 반대로 향상되었는데 자체 제작 데이터셋의 추가 Fine-Tuning과 도메인 적응으로 인한 성능 향상으로 분석된다. 최종적으로 GOPRO 데이터셋에서의 일부 성능 하락이 있었지만 37FPS 이상의 실시간적인 측면, 다른 데이터셋에서의 성능이 유지 혹은 향상이 있었기에 최종 모델로 선정하였다.

3.2 Super-Resolution Qualitative Results

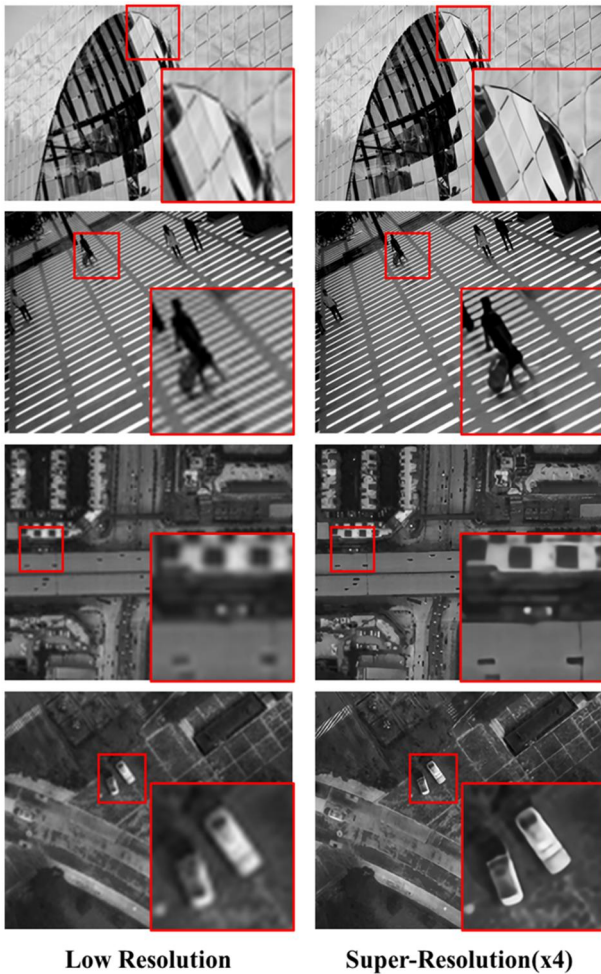


Fig. 8. Input Low Resolution image and Output Super-Resolution image.

Fig. 8.은 EO 도메인인 URBAN100(위의 두 영상) 데이터셋과 IR 도메인인 VGTSR(아래의 두 영상) 데이터셋으로 추론한 정성적 결과이다. 계단 열화현상 혹은 초점이 맞지 않는 것과 같이 번지는 영상에 대해 고화질로 영상을 복원하는 결과를 보였다. 또한, EO, IR 도메인 모두 정성적인 결과에서 성능 저하 없이 우수한 복원 품질을 보였다.

특히, 본 연구에서는 번짐 입력 영상에 대해 선명한 초해상도 영상을 얻기 위해 멀티태스크 러닝 기반 초해상도-디블러링 통합형 모델을 개발하였다. 이를 검증하기 위해 원본영상에 모션 블러 커널을 활용하여 모션 블러가 존재하는 번짐 입력 영상을 생성하였고 각각 멀티태스크 러닝 적용 전후의 결과를 비교하였다.

Fig. 9. 좌측은 멀티태스크 러닝 적용 전 단일 초해상도 모델의 결과이다. 공유 인코더를 통해 번짐에 대한 특징맵

을 가지고 있지 않기 때문에 번짐이 제거되지 않고 단순히 초해상도를 수행한 결과를 볼 수 있다. 반면, 우측은 멀티태스크 러닝 기반 통합형 모델의 초해상도 결과로 번짐 영상이 입력되어도 공유 인코더를 통해 번짐이 제거된 초해상도 영상을 얻는 결과를 볼 수 있다.

결과적으로, Fig. 9.의 결과와 같이 멀티태스크 러닝 기반 디블러링 인코더를 공유 인코더로 활용한 통합형 모델을 개발함으로써 흔들림이 심한 플랫폼에서 번짐 영상을 획득하더라도 번짐이 제거된 선명한 초해상도 영상을 얻을 수 있음을 확인하였다.

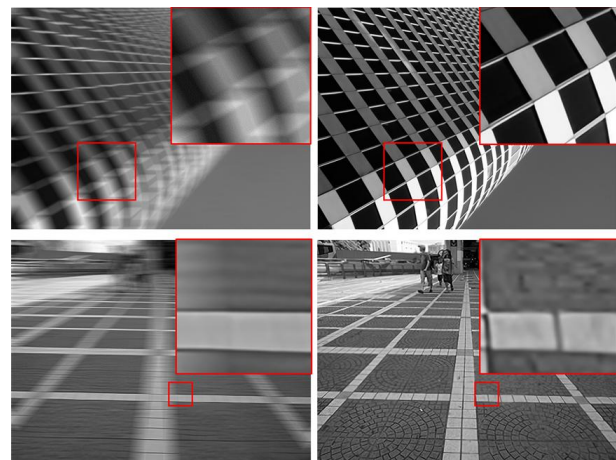


Fig. 9. Single Super-Resolution Output image(left) and Our Multi-Task Learning based Super-Resolution image(right).

3.3 Deblurring Qualitative Results

Fig. 10.은 EO 도메인인 GOPRO(위의 두 영상) 데이터셋과 블러 커널을 활용한 모션 블러를 랜덤하게 적용하여 자체적으로 제작한 IR 도메인 VGTDB(아래의 두 영상) 데이터셋으로 추론한 정성적 결과이다. 지식 증류를 통해 모델의 파라미터가 절반 이상 감소했음에도 불구하고 Table 2.의 정량적인 결과뿐만 아니라 정성적인 결과에서도 디블러링이 정상적으로 수행됨을 확인할 수 있다.

특히 디블러링 태스크에서 중요한 번짐이 존재하는 영상에 대해 글자, 건물의 질감, 차량 등을 선명하게 복원한다. 또한, EO 도메인과 IR 도메인에서 모두 정성적인 결과에서 번짐을 잘 제거하는 결과를 확인할 수 있다.



Fig. 10. Input Blurry Image and Output Deblurring Image.

V. Conclusion

본 논문에서는 기동이 심한 플랫폼에서 번짐이 제거된 초해상도 영상을 얻기 위해 디블러링 인코더를 공유 인코더로 활용하는 멀티태스크 러닝 기반 통합형 모델을 제안한다. 또한, RGB나 EO 도메인에 비해 비교적 학습 데이터가 부족한 IR 도메인에서의 추론 성능 향상을 위해 멀티태스크 러닝 기반으로 학습된 통합형 모델에 도메인 적응 기법을 활용하여 성능 향상을 수행하였다. 그리고, 30fps 이상의 실시간으로 임베디드 장비에서 추론 가능하도록 Teacher-Student 방식의 지식 증류 기반으로 모델을 경량화하였다.

결과적으로, 번짐이 존재하는 입력 영상이 주어졌을 때 공유 인코더를 거친 초해상도 결과 영상은 번짐을 제거하며 초해상도를 수행하였다. 또한, 도메인 적응에 의한 IR 도메인 성능 향상을 통해 하나의 단일 모델로 다중 도메인의 추론 가능한 결과를 보였다. 마지막으로, 비교적 거대한 모델인 디블러링 브랜치의 가중치를 절반으로 줄인 Student 모델을 기반으로 지식 증류를 수행하여 초해상

도, 디블러링 통합 모델이 실시간으로 추론할 수 있도록 경량화를 수행하였다.

본 연구를 통해 멀티태스크 러닝 기반 통합형 모델로 기동이 심한 플랫폼에서 번짐 입력 영상에 대해 초해상도 결과를 도출하였을 때 번짐이 제거된 초해상도 결과를 확인하였다. 추후 초해상도, 디블러링 뿐만 아니라 다른 영상 개선 태스크도 브랜치로 구성하여 통합형 모델을 설계하여 다양한 영상 개선의 결합을 기대해 볼 수 있다.

ACKNOWLEDGEMENT

This research was supported by the Korea Research Institute for Defense Technology (KRIT), funded by the Defense Acquisition Program Administration (DAPA), Republic of Korea, under Grant No. KRIT-CT- 22-073, in 2026.

REFERENCES

- [1] O.Ronneberger, P.Fischer, and T.Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition., 2015, DOI:10.48550/arXiv.1505.04597.
- [2] Cho, Sung-Jin, et al. "Rethinking coarse-to-fine approach in single image deblurring." Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition., 2021, DOI:10.48550/arXiv.2108.05054.
- [3] Wan, Cheng, et al. "Swift parameter-free attention network for efficient super-resolution." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2024, DOI:10.48550/arXiv.2311.12770.
- [4] Ganin, Yaroslav, et al. "Domain-adversarial training of neural networks." The Journal of Machine Learning Research vol.17, pp.1-35, 2016, DOI:10.48550/arXiv.1505.07818.
- [5] L.Ya, et al. "DIPNet : Efficiency Distillation and Iterative Pruning for Image Super-Resolution," Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition Workshops. pp-1692-1701, 2023, DOI:10.48550/arXiv.2304.07018.
- [6] Agustsson, Eirikur, and Radu Timofte. "NTIRE2017 challenge on single image super-resolution: Dataset and study." Proceedings of the IEEE/CVF conference on Computer Vision and Pattern Recognition Workshops., 2017, DOI: 1109/CVPRW.2017.150.
- [7] Zhao, Zhicheng, et al. "Thermal UAV Image Super-Resolution Guided by Multiple Visible Cues." IEEE Transactions on

Geoscience and Remote Sensing, 2023, DOI: 1109/TGRS.2023.3234058.

- [8] Nah, Seungjun, et al. "NTIRE2019 challenge on video deblurring and super-resolution: Dataset and study." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops., 2019, DOI:1109/CVPRW.2019.00251.

Authors



Jae-Yeob Kim received the M.S. degree from the Department of Electronic Engineering, Hanyang University, Seoul, South Korea, in 2022. He is currently working with the EO/IR Systems Research and Development

Laboratory, LIG Defense&Aerospace, Pangyo, South Korea. His research interests include VLM(Visual-Language Model), image deblurring, super-resolution, and object detection, object tracking in EO/IR systems.



Ji-Su Yoon received the M.S. degree from the Department of Artificial Intelligence In, Hanyang University, Seoul, South Korea, in 2025. She is currently working with the EO/IR Systems Research and Development

Laboratory, LIG Defense&Aerospace, Pangyo, South Korea. Her research interests include domain adaptation, object detection, object tracking in EO/IR systems.



Hyun-Sung Jang received the B.S. and M.S. degrees in electrical and computer engineering from Pusan National University, Busan, South Korea, in 2007 and 2009, respectively. He is currently pursuing the Ph.D. degree in

electrical and electronic engineering with Yonsei University, Seoul, South Korea. He is a Chief Research Engineer with the EO/IR Systems Research and Development Laboratory, LIG Defense&Aerospace, Pangyo, South Korea. His research interests include computer vision and machine learning.



Seung-Wook Park received the B.S. degree in electrical and electronic engineering from Korea University, Seoul, South Korea, in 2007. He is currently with the EO/IR Systems Research and Development

Laboratory, LIG Defense&Aerospace, Pangyo, South Korea. His research interests include EO/IR systems and large-aperture telescope systems.