

웨이블릿 기반 확산 모델을 활용한 고효율 저조도 이미지 복원 및 국방 분야 적용 방안

김범규, 권현

육군사관학교 AI데이터과학과

khs998199@gmail.com, hkwon.cs@gmail.com

Efficient Low-Light Image Restoration using Wavelet-based Diffusion Models and Defense Applications

Kim Beom Gyu, Hyun Kwon

Korea Military Academy

요약

현대 전장 환경과 감시 정찰 시스템에서 저조도 이미지 개선(Low-Light Image Enhancement, LLIE)은 핵심적인 기술이다. 최근 확산 모델(Diffusion Models)은 뛰어난 복원 성능을 보이나, 반복적인 샘플링 과정으로 인한 느린 추론 속도와 높은 연산 비용으로 인해 국방 엣지(Edge) 디바이스 적용에 한계가 있었다. 본 연구는 이러한 문제를 해결하기 위해 웨이블릿 변환(Wavelet Transform)을 활용한 새로운 확산 모델 프레임워크인 DiffLL을 분석하고 검증한다. DiffLL은 이미지를 주파수 대역별로 분리하여 저주파수 영역에만 확산 모델을 적용하고, 고주파수 영역은 경량화된 모듈로 복원하는 이원화 전략을 취한다. 실험 결과, 기존 SOTA 모델인 WeatherDiff 대비 약 300배 이상의 추론 속도 향상을 확인하였으며, LOLv2-real 데이터셋에서 PSNR 28.857dB를 달성하여 화질 면에서도 압도적인 성능을 입증하였다. 본 연구 결과는 고성능 컴퓨팅 자원이 제한된 국방 환경에서 실시간 고화질 야간 영상 확보를 가능케 하여 감시 정찰 및 무인 체계의 작전 효율성을 크게 제고할 것으로 기대된다.

I. 서론

야간이나 악천후와 같은 저조도 환경에서 시각 정보를 확보하는 것은 국방 감시 정찰(ISR) 및 자율 무인 체계의 생존성과 직결된다. 전통적인 히스토그램 균일화(HE)나 Retinex 이론 기반의 방법들은 노이즈 증폭 및 색상 왜곡 문제를 야기했다[1]. 최근 딥러닝의 발전과 함께 등장한 생성형 확산 모델(Diffusion Models)은 데이터 분포를 정밀하게 학습하여 고품질의 이미지를 생성하는 데 성공했으나, 픽셀 공간 전체에 대해 수천 단계의 노이즈 제거 과정을 거쳐야 하므로 실시간 처리가 불가능하다는 치명적인 단점이 존재한다. 특히 WeatherDiff[2]와 같은 최신 모델은 패치 단위 처리로 인해 2K급 이상의 고해상도 이미지 처리 시 메모리 부족(OOM) 현상이 발생하며, 추론에 수십 초가 소요되어 시급을 다투는 전술 환경에 부적합하다. 본 연구는 이러한 확산 모델의 비효율성을 극복하기 위해 웨이블릿 변환을 도입한 DiffLL(Low-Light Image Enhancement with Wavelet-based Diffusion Models)[3] 방법론을 적용 및 분석한다. 본 연구의 목적은 1) 웨이블릿 기반의 공간 축소가 확산 모델의 추론 속도와 안정성에 미치는 영향을 정량적으로 분석하고, 2) 이를 국방 분야의 저조도 객체 탐지 및 상황 인식 시스템에 적용하기 위한 실질적인 방안을 고찰하는 데 있다.

II. 관련연구

II-1. 저조도 이미지 개선 기법

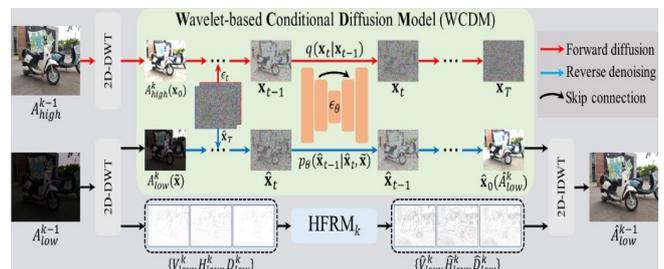
초기 딥러닝 기반 기법인 RetinexNet[4]은 조명 맵과 반사 맵을 분리하여 학습했으나 결과물이 부자연스러운 경향이 있었다. 이후 Zero-DCE[5], EnlightenGAN[6] 등 비지도 학습 기반 방법이 제안되었으나, 과도한 노출이나 학습 불안정성 문제가 여전히 존재한다.

II-2. 저조도 이미지 개선 기법

DDPM(Denoising Diffusion Probabilistic Models)의 등장 이후, Palette[7]나 WeatherDiff[2]와 같은 모델이 제안되었다. 이들은 기존 GAN 모델 대비 우수한 시각적 화질(Perceptual Quality)을 제공하지만, 반복적 연산에 따른 막대한 계산 비용과 랜덤 샘플링으로 인한 콘텐츠 일관성 부족 문제가 실용화의 걸림돌로 지적되어 왔다.

III. 제안방법론 :DiffLL

본 연구에서 분석하는 DiffLL의 핵심은 '주파수 분할 정복(Divide and Conquer)' 전략이다. 전체 파이프라인은 다음과 같다.



(그림1) 모델 파이프라인

입력 이미지를 2D 이산 웨이블릿 변환(DWT)을 통해 저주파 성분과 고주파 성분으로 분리한다. K=2 레벨의 변환을 수행하여 처리 해상도를 원본의 1/16로 축소함으로써 연산 효율을 극대화한다. 저주파 성분은 이미지의 전역적 구조와 밝기 정보를, 고주파 성분은 텍스처와 에지 정보를 포함한다.

축소된 저주파 성분은 WCDM을 통해 복원된다. 이 모델은 저조도 이미지를 조건(Condition)으로 입력받아 밝기와 색상을 보정한다. 특히 학습 단계에서 Forward와 Reverse 과

정을 모두 수행하여, 추론 시 단 10단계(S=10)의 샘플링만으로 도 안정적인 결과를 생성하도록 설계되었다.

고주파 성분은 확산 모델을 거치지 않고, 별도로 설계된 경량 CNN 모듈인 HFRM을 통해 복원된다. 이 모듈은 Cross-Attention 메커니즘을 활용하여 수직/수평 성분으로 대각선 성분을 상호 보완하며, 점진적 팽창 컨볼루션(Progressive Dilatation)을 통해 디테일을 선명하게 복원한다

IV. 실험 결과 및 분석

IV-1. 실험 환경 및 데이터셋

실험은 google colab T4 GPU 환경에서 수행되었으며, 저조도 환경 비교를 위해 다양한 조도 환경이 구성되어 있는 LOLv1 표준 데이터셋을 사용하였다. 비교군으로는 WeatherDiff를 선정하여 공정한 비교를 위해 동일 조건에서 재학습(Retraining)하였다. Batch size는 16으로 고정하고, epoch를 10부터 5씩 단계별로 올려가며 실험을 진행했다.

모델의 핵심 파라미터인 Wavelet Scale(K)은 2로 Time Step(T)은 50으로 설정해 시간을 단축하고자 했다.

IV-2. 정량적 성능 평가(Quantative Evaluation)

본 연구에서는 제안 모델(DiffLL)의 성능 우수성을 검증하기 위해 기존의 최신 확산 모델인 WeatherDiff와 4가지 핵심 평가지표를 비교 분석하였다. 평가 지표는 물리적 신호 왜곡을 측정하는 PSNR 및 SSIM과 인간의 시각적 만족도를 반영하는 LPIPS 및 FID로 구성된다. 구체적인 정의와 실험 결과는 [표 1]과 같다.

	WeatherDiff	DiffLL
PSNR	17.913	26.336
SSIM	0.811	0.845
LPIPS	0.272	0.217
FID	73.903	48.114

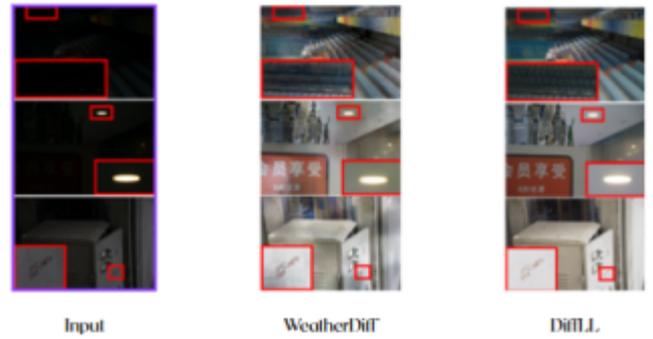
(표1) PSNR, SSIM, LPIPS, FID에 대한 성능 분석 결과 표

첫째, 왜곡 지표(Distortion Metrics) 분석 결과이다. 노이즈가 얼마나 적은지를 나타내는 PSNR에서 DiffLL은 26.336dB를 기록하여 WeatherDiff(17.913dB) 대비 약 8.4dB의 압도적인 성능 향상을 보였다. 또한, 이미지의 구조, 밝기, 대비가 원본과 얼마나 유사한지를 측정하는 SSIM에서도 0.845를 기록하여, WeatherDiff(0.811)보다 원본의 구조적 정보를 더 충실히 보존함을 입증하였다.

둘째, 인지 지표(Perceptual Metrics) 분석 결과이다. 최근 생성형 모델 평가에서 중요시되는 LPIPS는 사람이 보기에 얼마나 차이가 나는지를 수치화한 것으로, 값이 낮을수록 우수하다. DiffLL은 0.217을 기록하여 WeatherDiff(0.272) 대비 더욱 자연스러운 복원 결과를 보였다. 특히 생성된 이미지 집단과 실제 이미지 집단 간의 분포 차이를 나타내는 FID 점수에서 DiffLL은 48.114를 기록, WeatherDiff(73.903)와 비교하여 약 25점 이상의 큰 격차를 보이며 실제 야간 환경과 유사한 고품질 이미지를 생성함을 확인하였다

IV-3. 정성적 성능 평가(Qualitative Evaluation)

정량적 수치뿐만 아니라 실제 복원된 이미지의 시각적 품질을 비교 분석하였다. 그림2는 저조도 입력 영상(Input)에 대해 WeatherDiff와 DiffLL을 적용한 결과를 비교한 것이다.



(그림2) input에 대한 WeatherDiff와 DiffLL 이미지 비교

첫 번째 행(관중석)을 살펴보면, 입력 영상의 어두운 영역을 복원하는 과정에서 WeatherDiff는 좌석의 경계가 뭉개지거나 색상이 탁해지는 현상이 관찰된다. 반면, DiffLL은 좌석의 에지와 색감을 선명하게 복원하여 물체의 식별이 용이하다.

두 번째 행(실내 조명)에서는 과도한 빛 번짐(Blooming) 현상 억제 능력을 확인할 수 있다. WeatherDiff는 광원 주변이 뿌옇게 처리되어 선명도가 떨어지는 반면, DiffLL은 광원의 형태를 유지하면서도 주변부의 밝기를 자연스럽게 개선하였다.

세 번째 행(문자 식별)은 국방 감시 정찰에서 가장 중요한 세부 정보 복원 능력을 보여준다. 배전반의 붉은색 텍스트와 표면의 질감을 복원함에 있어, WeatherDiff는 글자가 흐릿하게 뭉개져 가독성이 떨어진다. 그러나 DiffLL은 고주파 복원 모듈(HFRM)의 적용으로 텍스트의 획과 윤곽선을 또렷하게 복원해 냈다.

V. 국방분야에 대한 고찰

현대전은 전장 가시화와 정보 우위(Information Superiority) 확보가 승패를 좌우한다. 본 연구에서 검증된 DiffLL 모델은 국방 분야에 다음과 같은 혁신적 가치를 제공한다.

첫째, 엣지 디바이스 기반의 실시간 야간 정찰이다. 기존의 고성능 서버 의존형 모델과 달리, DiffLL은 경량화된 구조로 무인기(UAV)나 개인 전투 장비(웨어 플랫폼)의 칩셋에서도 구동 가능하다. 이는 통신 단절 상황에서도 독립적인 작전 수행을 보장한다.

둘째, AI 표적 식별 시스템의 신뢰도 향상이다. 본 연구에서 DiffLL이 보여준 우수한 에지 복원력과 텍스처 보존 성능은 객체 탐지 모델의 전처리 단계에서 중요한 역할을 할 수 있다. 비록 본 연구 범위 내에서 직접적인 검출 실험은 수행하지 않았으나, 선명하게 복원된 객체 윤곽선은 야간 경계 감시 시스템(GOP/GP)의 오경보율(False Alarm)을 낮추고 적 식별 능력을 강화하는 데 이론적으로 기여할 수 있을 것으로 판단된다.

셋째, 합성 데이터 생성의 고도화이다. 국방 M&S 체계 구축 시 부족한 야간 학습 데이터를 고품질로 생성하여, AI 모델의 강건성을 높이는 데 기여할 수 있다. 향후 적외선(IR) 센서와 융합한 멀티모달 복원 기술로 확장한다면 전천후 감시 체계 구축이 가능할 것이다.

VI. 결론 및 제언

본 연구는 웨이블릿 변환을 통해 확산 모델의 연산 효율성을 극대화한 DiffLL 기법을 분석하였다. 실험을 통해 해당 기법이 기존 SOTA 모델 대비 수백 배 빠른 속도와 탁월한 화질 복원 성능을 가짐을 입증하였다. 이는 컴퓨팅 자원이 제한된(SWaP-C) 국방 환경에서 야간 작전 능력을 획기적으로 개선할 수 있는 실용적 기술이며, 향후에는 실제 표적 탐지 모델과의 연동 성능 검증 및 실시간 영상 스트리밍 처리를 위한 파이프라인 최적화 연구가 추가적으로 요구된다.

참고문헌

- [1] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," *IEEE Trans. Image Process.*, vol. 26, no. 2, pp. 982–993, 2016.
- [2] O. Özdenizci and R. Legenstein, "Restoring vision in adverse weather conditions with patch-based denoising diffusion models," *IEEE TPAMI*, vol. 45, no. 8, pp. 10346–10357, 2023.
- [3] H. Jiang et al., "Low-Light Image Enhancement with Wavelet-based Diffusion Models," *ACM Trans. Graph.*, vol. 42, no. 6, Dec. 2023.
- [4] C. Wei et al., "Deep Retinex Decomposition for Low-Light Enhancement," in *BMVC*, 2018.
- [5] C. Guo et al., "Zero-reference deep curve estimation for low-light image enhancement," in *CVPR*, 2020.
- [6] Y. Jiang et al., "Enlightengan: Deep light enhancement without paired supervision," *IEEE Trans. Image Process.*, vol. 30, pp. 2340–2349, 2021.
- [7] C. Saharia et al., "Palette: Image-to-image diffusion models," in *ACM SIGGRAPH*, 2022.