

# 딥러닝 기반 영상개선 네트워크를 활용한 저조도 영상개선 기법에 관한 연구

임재성, 김진환\*  
국립과학수사연구원 디지털과

## A study on low-light image enhancement method using deep learning based image enhancement network

Jae Sung Lim, Jin-Hwan Kim\*  
Department of Digital Analysis Division, National Forensic Service

**요약** 본 논문은 저조도 영상개선을 위한 변환 함수를 추정하는 딥러닝 기반 기법을 제안하였다. 기존 모델 기반의 저조도 영상개선 기법은 단일 입력영상들의 특성을 이용하여 개선하기 때문에 다양한 저조도 계열 영상에 적용하는데 효과적이지 못했다. 반면에 제안하는 연구에서는 대용량의 데이터를 바탕으로 학습된 네트워크로 영상개선을 위한 변환 함수를 추정하였기 때문에 다양한 저조도 입력영상에 대해서도 대응이 가능하며, 특히 변환함수를 통해 입력에서 결과가 도출된 근거를 마련할 수 있다는 장점을 지니고 있다. 제안하는 기법의 효과를 평가하기 위해서 본 연구에서는 영상 품질을 확인하는 주요 지표인 PSNR과 SSIM 지표가 활용되었다. 실험 결과에서 제안된 기법이 적용한 영상개선 결과에서 밝기값을 개선시키면서도 색상의 왜곡을 억제한 결과와 함께 벤치마크 방법에 비해 PSNR과 SSIM의 정량적인 주요 지표값에서 평균적으로 50%, 17%가 개선된 것이 확인되었다.

**Abstract** This paper proposes a deep learning-based technique for estimating transformation functions to enhance low-light images. Conventional model-based methods for enhancing low-light images are not effective for various low-light scenarios as they are based on characteristics of individual input images. In contrast, the proposed approach estimates transformation functions for image enhancement using a network trained on a large dataset, allowing adaptation to diverse low-light input images and providing a rationale for the outcomes obtained from transformation functions. To evaluate the effectiveness of the proposed technique, key metrics for assessing image quality, namely PSNR and SSIM, were employed. Experimental results demonstrated that the proposed method improved brightness while suppressing color distortion in the enhanced images. Compared to benchmark methods, quantitative measurements of PSNR and SSIM showed improvements of up to 50% and 17% on average in the enhanced image results achieved by the proposed method.

**Keywords** : Low-light Image, Image Enhancement Network, Transformation Function, Forensic Science, Loss Functions

본 논문은 국립과학수사연구원(2023-디지털-03) 연구과제로 수행되었음.

\*Corresponding Author : Jin-Hwan Kim(National Forensic Service)

email: jinhwankim83@korea.kr

Received August 1, 2023

Revised August 31, 2023

Accepted September 1, 2023

Published September 30, 2023

## 1. 서론

저조도 영상이 획득되는 일반적인 환경은 주로 어두운 장소나 야외에서의 야간 촬영 시 발생하기 때문에 위와 같은 환경에서 촬영된 영상들은 일반적으로 영상 내 세부 정보나 색상이 제대로 표현되지 않는 경우가 대부분이다. 저조도 영상개선은 다음과 같은 주요 분야에서 활용될 수 있다. 첫째, 저조도 환경 하에 노이즈가 많고 낮은 대조비로 인해 관심 피사체의 구분 혹은 피사체 간 경계 구분이 모호한 경우에 개선된 영상은 더 나은 시각적 품질을 제공할 수 있다. 둘째, 저조도 조건에서 촬영된 영상은 보안 감시, 의료 영상 등 다양한 분야에서 획득되어 질 수 있기 때문에 이러한 저조도 영상개선은 특히 보안 감시 시스템으로부터 야간 환경에서 촬영된 영상이라든가 개선된 결과를 통해 피사체 즉 사람이나 물체를 식별하는데 도움을 줄 수 있다. 추가적으로 저조도 영상개선은 컴퓨터 비전 및 인공지능 분야에서 중요한 전처리 과정이다. 대부분의 컴퓨터 비전 작업은 영상 데이터에 의존하며 저조도 영상의 품질 개선은 객체 검출, 분할, 인식 등과 같은 고급 비전 작업에서 성능을 향상시킬 수 있는 필수적인 전처리 작업이 된다.

저조도 영상개선의 적용 기법은 여러 분야에 적용될 수 있지만 특히 법과학적 영상분석에서 범죄예방 및 교통사고 감시의 목적으로 설치된 CCTV, 블랙박스 등 영상 장비들에서 촬영되는 저조도 영상들은 야간 환경에서 촬영되거나, 촬영 장비의 노출에 관여하는 설정(조리개, 셔터속도, ISO)의 오류로 인해 저조도 영상을 분석해야 하는 경우가 다수 발생된다. 위 환경 하에 관심 피사체에 대한 영상 분석을 수행하게 되면 피사체에 대한 세부적인 특징을 판독하는데 상당한 어려움이 수반된다. 따라서 본 연구에서는 저조도 영상의 품질을 향상시킬 수 있는 기술에 대해 연구를 수행하였다.

기존 이론 기반의 전통적인 영상처리 기법[1-3]은 단일 영상만을 활용하여 영상 특성을 분석해야 하므로 다양한 저조도 입력영상에 대해 결과 영상이 색상 왜곡이나 잡음 등으로 인해 품질이 저하되거나 품질의 균일성이 떨어지게 된다. 반면에 제안하고자 하는 딥러닝 기반의 저조도 영상 기법은 대규모 데이터를 활용한 학습을 통해 개선된 화질을 생성할 수 있으며 또한 법과학적 측면에서 검증력을 제시할 수 있도록 본 연구에 반영하여 입력영상으로부터 결과 영상을 생성하는 과정에서 산출되는 매핑 함수(변환 함수)를 확보하고자 한다.

본 논문은 다음과 같은 순서로 구성된다. 2절 본론 중 2.1절에서는 제안하고자 하는 딥러닝 기반 네트워크 구조의 개략적인 흐름에 대한 요약 이후에 각 네트워크를 구성하는 개별 블록에 대한 내용을 기술하고 2.2절과 2.3절에서는 학습에 사용될 데이터 세트의 구성과 손실 함수의 설계 내용에 대해 기술한다. 2.4절 실험 결과에서는 영상개선 결과를 타 벤치마크 방법과 비교하여 결과 영상과 함께 제시하고자 하며, 마지막으로 위 연구에 대한 내용을 결론부에서 정리하도록 한다.

## 2. 본론

### 2.1 제안하는 딥러닝 기반 저조도 영상개선 네트워크 구조

Fig. 1는 제안하는 알고리즘의 구조를 나타내며, 제안하는 네트워크는 입력 저조도 영상으로부터 영상개선을 목적으로 매핑 함수와 저조도 영상의 특징을 추출하는 영상 특징맵을 중간 과정의 결과물로 도출하게 되며 위 결과물들은 단일 네트워크 내에서 학습을 통해 도출되는 방식으로 설계하였다.

본 제안하는 네트워크 알고리즘을 살펴보기 이전에 영상개선을 위한 매핑 함수 추정을 위해 입력영상은 화소 값을 -1과 1사이로 정규화(normalization)하여 활용하며, 또한 저조도 입력영상은 네트워크의 입력으로 들어가기 전에 영상 크기를 256×256 조정하므로 영상 크기가 다른 입력영상을 사용하더라도 균일한 크기의 특징맵을 추출할 수 있다.

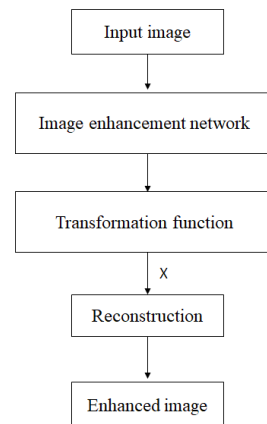


Fig. 1. A flow chart of the proposed algorithm

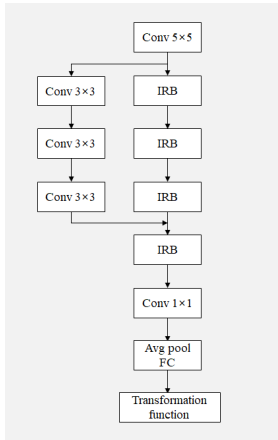


Fig. 2. Specification for image enhancement network in proposed method

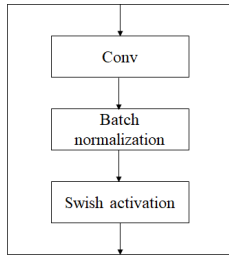


Fig. 3. Detailed structure of convolution block

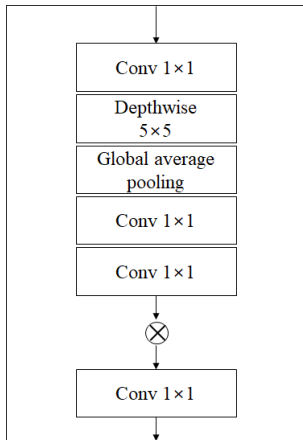


Fig. 4. Detailed structure of IRB

제안하는 네트워크를 구성하는 개별 블록에 대해서 살펴보고자 한다. Fig. 3에서 보는바와 같이 5×5 합성곱 블록(5×5 convolutional layer)은 stride 2의 5×5 합성곱 계층(5×5 convolutional layer)과 배치 정규화 층(batch normalization layer), 스위시 활성화 함수 층

(swish activation function layer)[4]으로 구성되어 있으며, 5×5 합성곱 블록을 통과하고 나면 입력영상에서 특징맵으로 도메인이 바뀌게 된다.

이때 합성곱 블록에서 사용되는 배치 정규화 층은 미니배치의 평균과 분산을 활용하여 평균이 0, 분산이 1이 되도록 정규화하여 변형된 분포가 나오지 않도록 배치 정규화 단계마다 확대(scale)와 이동(shift)을 수행한다. 또한 깊은 신경망에서 렐루 활성화 함수(relu activation function)[5]에 비해 높은 정확도를 보여주는 스위시 활성화 함수 층을 배치 정규화 층 이후에 사용하였다. 위 활성화 함수를 표현하는 식은 Eq. (1)과 같이 표현할 수 있다. 추가적으로 각 블록마다 드롭아웃(dropout)을 사용하여 네트워크 학습 과적화를 방지하고자 하였다.

$$f(x) = \frac{x}{1 + e^{-x}} \quad (1)$$

5×5 합성곱 블록에서 도출되는 영상 특징맵의 해상도는 입력으로 들어오는 특징맵 해상도의 절반으로 도출되도록 설계되었다. 5×5 합성곱 블록에서 출력된 영상 특징맵은 전역적 특징맵을 추출하는 IRB(Inverted Residual Block, Searching for mobilenetV3[6])와 지역적 특징맵을 추출하는 3×3 합성곱 블록(3×3 convolution block)의 입력으로 사용된다.

전역적 특징을 추출하는 IRB는 Fig. 4와 같이 구성되며, IRB는 특징맵에 전역적 평균 풀링(global average pooling)을 이용해서 계수들을 생성하게 되고, 이 계수들이 픽셀 단위 곱셈으로 적용되어 특징맵의 전역적 정보가 사용된 특징맵을 도출할 수 있게 된다. 또한 IRB는 적은 학습파라미터와 메모리를 활용하여 특징맵을 효율적으로 추출할 수 있다는 장점을 가지고 있다.

한편 영상의 지역적 특징을 추출하는 특성을 갖는 3×3 합성곱 블록은 stride 2의 3×3 합성곱 계층을 활용하며, 배치 정규화 계층, 스위시 활성화 함수 층으로 구성된다.

위 IRB와 3×3 합성곱 블록을 통해 생성된 전역적, 지역적 특징맵은 채널 단위의 접합(concatenation)을 통해 하나의 특징맵의 형태로 결합되게 된다. 결합된 특징맵이 다시 IRB를 거치고 나면 전역적 특징과 지역적 특징을 모두 갖는 특징맵이 만들어진다.

이후 합성된 특징맵은 1×1 합성곱 블록(1×1 convolution block)을 거쳐 768 채널의 특징맵이 되고 전역적 평균 풀링을 통해 각 채널 특징맵의 크기는 1×1 텐서가 된다. 이후 완전 연결 계층(fully connected layer)을 거친 뒤 차원 조절을 통해 3×256의 형태로 도

출하여 각 채널별로 매핑함수 추정할 수 있게 되었다.

위 단계로부터 학습된 각 채널별 매핑함수를 적용하여 시각적으로 밝기값이 개선된 영상을 도출하게 된다.

## 2.2 학습 및 테스트 데이터 세트 구성

본 제안하는 네트워크에서는 학습된 매핑함수와

그 매핑함수가 적용된 개선된 영상이 하나의 네트워크로부터 도출되는 것을 2.1.절을 통해 알게 되었다. 또한, 본 연구에서 학습데이터를 구성할 시에 저조도 영상에 대응되는 원본 데이터 세트가 제공되는 MIT Adive 5K(FiveK)[7], Low-Light(LOL)[8], Matching in the Dark(MID)[9] 데이터 세트를 활용하여 혼합된 데이터 세트를 구성하였으며, 그 중 위 언급된 세 개의 데이터 세트로부터 1050장의 학습 데이터 세트 및 테스트 데이터 세트를 구성하였다. 또한 학습데이터 중 저조도 영상을 80%, 일반조도 영상을 20% 비율로 구성하였다. 또한 위 데이터 세트에서 원본 영상(Ground Truth)을 제공하기 때문에 이를 이용하여 학습을 수행할 수 있으나, 본 연구에서 일반적인 저조도 영상에서 관찰되지 않는 세부적 정보(image details)를 판독할 수 있는 수준의 영상 밝기로 끌어올리기 위해 원본 영상의 밝기값을 개선시킨 영상[10]으로 원본 영상을 대체하여 활용하였다. 따라서 본 연구에서 위 언급된 알고리즘으로 원본 영상을 대체하여 활용한 것은 법과학적 영상 분석을 위해 저조도 영상에서 세밀한 정보를 식별(판독)하고자 하는데 이유가 있다.

## 2.3 훈련 과정(손실함수 설계)

본 네트워크를 학습시키기 위해 본 연구에서는 네 가지 파트로 이루어진 총 손실함수를 활용하고 있으며, 우리는 총 손실함수를 아래 식과 같이 나타낼 수 있다

$$L_{total\ loss} = \lambda_{img}L_{img} + \lambda_{hmse}L_{hmse} + \lambda_{exp}L_{exp} + \lambda_{var}(1-L_{var})$$

위 총 손실함수( $L_{total\ loss}$ )는 image loss, hmse loss, exposure loss, 1-variance loss로 구성했으며, 위 손실함수 항들의 가중합으로 표현된다.

$\lambda_{img}$ ,  $\lambda_{hmse}$ ,  $\lambda_{exp}$ ,  $\lambda_{var}$ 는  $L_{img}$ ,  $L_{hmse}$ ,  $L_{exp}$ ,  $(1-L_{var})$ 의 각 개별 항의 하이퍼 파라미터로 네 가지 손실 항들의 밸런스를 맞추는 역할을 수행한다.

영상 손실은  $L_{img}$ 로 표기하며, 평균제곱오차인  $L_2$ -norm을 이용하며  $L_{img}$ 는 Eq. (2)로 표현된다.

$$L_{img} = \frac{1}{N} \sum_{i=1}^N \| Y_i' - Y_i \|^2 \quad (2)$$

여기서  $N$ 은 영상의 수,  $Y'$ 은 제안된 네트워크를 통해 학습된 매핑함수를 저조도 영상에 적용하여 개선된 영상이며,  $Y$ 는 원본 영상을 의미한다. 위 식을 통해 개선된 영상  $Y'$ 과 원본 영상  $Y$ 간의 오차를 최소화하도록 학습을 수행한다. 따라서 학습되어 도출된 영상과 원본 영상에 맞추어 컬러 정보(color consistency)도 함께 보존된다.

$L_{hmse}$  손실함수는 원본 영상의 히스토그램과 추정된 영상의 히스토그램의 오차를 줄여주는 항으로 활용하였다.

$L_{exp}$  손실함수는 개선된 영상에서 과 노출 영역과 저 노출 영역이 나타나는 것을 억제하기 위해 해당 손실함수를 활용하며, 지역영역의 밝기의 평균값과 적정노출레벨( $L=0.6$ )간에 차(혹은 거리)를 Eq. (3)을 통해 계산한다.

$$L_{exp} = \frac{1}{N} \sum_{k=1}^N | Y_k - L | \quad (3)$$

$N$ 은 서로 겹치지 않는  $16 \times 16$  크기의 지역적 영역의 수이며,  $Y$ 는 개선된 영상에서의 지역적 영역의 평균 밝기값이다.

$1-L_{var}$  손실함수는 Eq. (4)로 표현되며, 영상 픽셀 휘도값을 분산시키기 위해 활용하였다. Eq. (4)는 다음과 같이 나타낸다.

$$L_{(1-var)} = 1 - \frac{1}{N} \sum_{k=1}^N (Y_k - \sigma(Y))^2 \quad (4)$$

여기서  $N$ 은 영상의 수,  $Y$ 는 개선된 영상,  $\sigma(Y)$ 는 개선된 영상의 표준편차를 의미한다.

손실함수의 밸런스를 위해서 각 계수들의 초기값을 각각  $\lambda_{img}=0.3$ ,  $\lambda_{hmse}=0.2$ ,  $\lambda_{exp}=0.2$ ,  $\lambda_{var}=0.3$ 으로 설정하여 학습을 수행하였다.

## 2.4 실험 결과

우리는 저조도 영상개선에 대한 결과 영상을 평가하기 위해서 전통적인 저조도 벤치마크 방법과 딥러닝 기반의 벤치마크 방법을 선택하였으며, 정량적인 평가를 위한 지표 네 개를 선정하여 활용하였다: Peak Signal-to-Noise Ratio(PSNR), Structural Similarity Index Measure(SSIM)[11], Measure of Enhancement(EME)[12], Pixel wise Distance(PixDIST)[13].

PSNR, SSIM은 개선된 영상과 원본영상과의 유사성 및 영상 품질을 확인하는 지표이며, EME와 PixDist는 개선된 영상의 대조비 및 히스토그램의 균일성을 보는 지표이다. 위 2.2절에서 구성한 데이터 세트 중 테스트

Table 1. Quantitative comparison on test dataset

| Method      | PSNR  | SSIM | EME   | PixDist |
|-------------|-------|------|-------|---------|
| LIME[10]    | 12.69 | 0.57 | 25.86 | 26.38   |
| ZeroDCE[14] | 12.14 | 0.58 | 20.81 | 16.97   |
| Proposed    | 18.61 | 0.67 | 18.92 | 28.56   |

영상을 대상으로 앞서 언급된 4개의 지표를 활용하였다. Table 1에서 보는 바와 같이 제안하는 기법이 기존 최신 벤치마크에 비해서 영상 품질 측면에서 우수한 PSNR과 SSIM 수치를 보여주고 있으며 안정적인 대조비와 히스토그램 균질성을 보여주고 있는 것을 볼 수 있다.



Fig. 5. Comparative image enhancement results: (a) Input image, (b) LIME, (c) ZeroDCE[14], (d) Proposed

또한 정성적으로 확인했을 때, Fig. 5에서 볼 수 있듯이 타 벤치마크 방법은 입력영상을 과도하게 밝게 만들어 피사체의 색상이 왜곡 되는 등 부자연스러운 결과를 보여주는데 비해 상대적으로 제안하는 기법은 피사체의 색 왜곡이 관찰되지 않으면서 저조도 입력영상의 휘도도 함께 효과적으로 개선시킨 것으로 확인되었다.

### 3. 결론

본 연구에서 딥러닝 기반의 단일 네트워크를 설계하여 변환함수(매핑함수)를 도출하여 저조도 영상을 개선하는 방법을 제안하였다. 제안하는 기법은 변환함수가 적용되어 결과영상을 도출하므로 입력영상에 대한 결과영상이 어떤 과정으로 생성되었는지 근거 제시가 가능하다는 추가적인 장점을 가지고 있다. 특히 법과학 분야에서 결과영상의 도출과정을 증명해야 하는 경우 본 논문의 결과

물인 변환함수를 통해 검증력을 확보할 수 있었다. 또한 제안하는 기법은 실험결과 측면에서 영상품질을 보여주는 주요 평가지표인 PSNR과 SSIM의 평균 수치에서도 타 기법에 비해 개선된 결과를 보여주며 영상개선 결과에서 저조도 영상의 밝기값을 끌어올리고 동시에 피사체의 구조적 정보는 유지하면서 색상 왜곡이 관찰되지 않는 점 등으로 보아 제안된 기법의 효과성을 확인할 수 있었다.

본 실험에서 확인된 것처럼 필요에 의해 밝기값을 크게 높여야 하는 경우가 존재하는데 이렇게 되면 색상 왜곡이 발생할 가능성이 높아지게 된다. 따라서 차후 연구에서는 색상정보와 영상정보의 왜곡 및 손실을 최대한 저감 및 보존하면서 밝기를 최대한 개선할 수 있는 방법을 확보할 수 있도록 연구를 수행하고자 한다.

### References

- [1] R.C. Gonzalez and R. E. Woods, "Digital Image Processing", Upper Saddle River, N.J., USA: Pearson 2018.
- [2] K. He, J. Sun, and X. Tang, "Single image haze removal using dark channel prior", *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.33, No.12, pp.2341-2353, Dec. 2011. DOI: <https://doi.org/10.1109/TPAMI.2010.168>
- [3] X. Ren, W. Yang, W-H. Cheng, and J.Liu, "LR3M: Robust low-light enhancement via low-rank regularized retinex model", *IEEE Trans. Image Processing*, Vol.29, pp.5862-5876, Apr. 2020. DOI: <https://doi.org/10.1109/TIP.2020.2984098>
- [4] P. Ramachandran, B. Zoph, and Q. V. Le, "Swish: a Self-Gated Activation Function", *Neural and Evolutionary Computing*, Oct. 2017. DOI: <https://doi.org/10.48550/arXiv.1710.05941>
- [5] Liu, Danqing, "A Practical Guide to ReLU", *Medium*, Nov. 2017.
- [6] A. Howard, M. Sandler, G. Chu, L. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, "Searching for MobileNetV3", *Computer Vision and Pattern Recognition*, Nov. 2019. DOI: <https://doi.org/10.48550/arXiv.1905.02244>
- [7] V. Bychkovsky, S. Paris, E. Chan, and F. Durand, "Learning photographic global tonal adjustment with a database of input/output image pair", *CVPR 2011*, June 2011. DOI: <https://doi.org/10.1109/CVPR.2011.5995332>
- [8] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep Retinex Decomposition for Low-Light Enhancement", *BMVC 2018, Computer Vision and Pattern Recognition*, Aug.

2018.

DOI: <https://doi.org/10.48550/arXiv.1808.04560>

- [9] W.Song, M. Suganuma, X. Liu, N. Shimobayashi, D. Maruta, and T. Okatani, "Matching in the Dark: A Dataset for Matching Image Pairs of Low-light Scenes", *ICCV 2021, Computer Vision and Pattern Recognition*, Sep. 2021.  
DOI: <https://doi.org/10.48550/arXiv.2109.03585>
- [10] X.Guo, Y.Li and H.Ling, "LIME: Low-light image enhancement via illumination map estimation", *IEEE Trans. Image Processing*, Vol.26, No.2, pp.982-993, Feb. 2017.  
DOI: <https://doi.org/10.1109/TIP.2016.2639450>
- [11] Z.Wang, A.Bovik, H.Sheikh, and E.Simoncelli, "Image quality assessment: from error visibility to structural similarity", *IEEE Trans. Image Processing*, Vol.13, No.4, pp.600-612, Apr. 2004.  
DOI: <https://doi.org/10.1109/TIP.2003.819861>
- [12] S.Agaian, B.Silver, and K.Panetta, "Transform Coefficient Histogram-Based Image Enhancement Algorithm Using Contrast Entropy", *IEEE Trans. Image Processing*, Vol.16, No.3, pp.741-758, Mar. 2007.  
DOI: <https://doi.org/10.1109/TIP.2006.888338>
- [13] Z.Chen, B.Abidi, D.Page, and M.Abidi, "Gray-level grouping(GLG): an automatic method for optimized image contrast enhancement-part I: the basic method", *IEEE Trans. Image Processing*, Vol.15, No.8, pp.2290-2302, July 2006.  
DOI: <https://doi.org/10.1109/TIP.2006.875204>
- [14] C.Guo, C.Li, J.Guo, C.Loy, J.Hou, S. Kwong, and M.Abidi, "Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement", *Computer Vision and Pattern Recognition*, Mar. 2020.  
DOI: <https://doi.org/10.48550/arXiv.2001.06826>

김진환(Jin-Hwan Kim)

[정회원]



- 2017년 12월 : 국립과학수사연구원 법공학부 디지털과

<관심분야>

영상신호처리, 영상개선

임재성(Jae Sung Lim)

[정회원]



- 2018년 11월 : 국립과학수사연구원 법공학부 디지털과

<관심분야>

영상신호처리, 영상개선