

## 딥러닝기반 실내와 실외 환경에서의 광원 추출

이지원<sup>o</sup>      서광균      이하늬      유정은      노준용<sup>\*</sup>

카이스트 비주얼미디어연구실

{mpm11<sup>o</sup>, skg1023, honeylee, jey920, junyongnoh<sup>\*</sup>}@kaist.ac.kr

## Deep Learning-Based Lighting Estimation for Indoor and Outdoor

Jiwon Lee<sup>o</sup>      Kwanggyoon Seo      Hanui Lee      Jung Eun Yoo      Junyong Noh<sup>\*</sup>

KAIST, Visual Media Lab

### 요약

본 연구에서는 딥러닝을 기반으로 하여 실내와 실외 이미지 모두에서 알맞은 광원을 추출하는 방법론을 소개한다. 네트워크는 단일 LDR 이미지로부터 실내 혹은 실외 배경에 맞는 광원을 low dynamic range (LDR) 환경 맵으로 추출하는 Crop-to-PanoLDR 네트워크와 추출된 LDR 환경 맵을 빛의 정보를 담은 high dynamic range (HDR) 환경 맵으로 생성하는 LDR-to-HDR 네트워크 두 단계로 구성된다. 이와 같은 과정을 통해 최종적으로 생성된 HDR 환경 맵은 주어진 이미지에서 가상 객체를 렌더링할 때 적용되어 가상 객체를 조명하는 빛의 방향과 주변광 등을 확인함으로써 자연스러운 렌더링을 가능하게 하는지 검증한다. 본 연구에서 제안한 방법론의 우수성은 실내를 배경으로 한 이미지로만 구성된 데이터로 학습한 결과와 실외를 배경으로 한 이미지로만 학습한 결과 등과 비교하여 검증하였다. 또한, 실내와 실외를 구분하는 역할을 수행하는 손실 함수가 학습 결과에 미치는 영향을 실험, 검증하였다. 최종적으로 본 연구에서 생성된 환경 맵을 기존의 연구 결과와 비교 실험하는 사용자 테스트를 진행하였고 더 좋은 결과를 확인할 수 있었다.

### Abstract

We propose a deep learning-based method that can estimate an appropriate lighting of both indoor and outdoor images. The method consists of two networks: Crop-to-PanoLDR network and LDR-to-HDR network. The Crop-to-PanoLDR network predicts a low dynamic range (LDR) environment map from a single partially observed normal field of view image, and the LDR-to-HDR network transforms the predicted LDR image into a high dynamic range (HDR) environment map which includes the high intensity light information. The HDR environment map generated through this process is applied when rendering virtual objects in the given image. The direction of the estimated light along with ambient light illuminating the virtual object is examined to verify the effectiveness of the proposed method. For this, the results from our method are compared with those from the methods that consider either indoor images or outdoor images only. In addition, the effect of the loss function, which plays the role of classifying images into indoor or outdoor was tested and verified. Finally, a user test was conducted to compare the quality of the environment map created in this study with those created by existing research.

**키워드:** 광원 추출, 딥러닝, HDR 환경 맵, 실내 광원, 실외 광원

**Keywords:** lighting estimation, deep learning, HDR environment map, indoor lighting, outdoor lighting

## 1. 서론

최근 가상현실 (VR, virtual reality), 증강현실 (AR, augmented reality) 및 혼합현실 (MR, mixed reality) 생성 기술의 발전으로 실

제 객체와 가상 객체와의 상호작용이 가능해졌다. 여기서 가상 객체는 실제 객체와 가까운 모습을 가지도록 하여 더욱 실감나는 경험을 제공하기도 한다. 이를 가능하게 해 주는 것이 바로 이미지 기반 조명 (IBL, image based lighting) 기법으로, 실제 환경의

\*corresponding author: Junyong Noh/KAIST(junyongnoh@kaist.ac.kr)

광원을 이미지로 표현한 HDR 환경 맵을 이용하여 가상 객체를 조명하게 해주는 방법이다 [1]. HDR 환경 맵은 카메라와 어안렌즈, 거울 공 등 특수 장비를 사용하거나 노출을 조절한 여러장의 사진을 찍어 이를 합성하는 것으로 최종 생성하게 된다.

가상 객체를 조명하기 위해 매번 특수 장비를 사용하여 HDR 환경 맵을 생성하는 것은 비효율적이기에 최근에는 딥러닝을 사용하는 방법들이 제안되고 있다. 딥러닝 방법을 사용하게 될 경우 데이터를 이용한 네트워크 학습을 통해 이미지로부터 HDR 환경 맵을 추정할 수 있다는 장점이 있다. 그리하여 지금까지 딥러닝 방법을 이용한 광원 추출 연구는 일반적인 카메라로 촬영한 제한된 시야의 단일 LDR 이미지로부터 가상 객체를 조명할 수 있게 하는 HDR 환경 맵을 자동적으로 생성해내는 연구로 발전되어왔다 [2, 3, 4, 5, 6, 7].

지금까지 딥러닝을 사용하여 HDR 환경 맵을 구하는 광원 추출 연구 방법은 광원을 추출할 이미지의 배경이 실내인지 실외인지에 따라 구분되어 다른 방법으로 연구되어왔다. 이는 광원의 특성에 따라 구분된 것으로, 먼저 실외의 경우 주광원인 태양만을 고려하여 하늘과 태양의 특성을 담은 광원을 추출할 수 있다 [5, 6, 7]. 반면 실내에서는 광원의 종류가 다양하며 분포하는 위치도 다양하다는 특성에 따라 어떤 특정한 부분만 고려해 광원을 추출하기 어렵다 [2]. 게다가 테이블 아래 그림자가 생기는 공간과 같이 기하학적인 구조를 고려해야 하는 경우가 존재하기 때문에 실내 배경에서의 광원 추출은 실외 배경과 비교하여 좀 더 복잡한 방법을 요구하게 된다 [3, 4]. 그러나 이렇게 실내와 실외를 구분한 연구 방법은 광원을 추출할 대상인 이미지가 실내 혹은 실외로 구분되기 어려울 경우 적용되기 어렵다는 문제가 발생한다. 또한 동영상 혹은 실시간 영상을 배경으로 한 가상 객체와의 상호작용이 요구되는 가상현실과 증강현실, 혼합현실 어플리케이션의 사용이 활발해짐에 따라 실내와 실외 모두에서 광원을 추출할 수 있는 연구가 필요하게 되었다.

본 연구에서는 광원을 추출할 이미지의 실내외 여부에 상관없이 알맞은 광원을 추출할 수 있는 딥러닝 방법을 제안한다. 본 연구는 실내와 실외를 구분하지 않은 단일 LDR 이미지를 입력으로 하였을 때, 딥러닝 네트워크에서 해당 이미지에 알맞은 광원을 HDR 환경 맵의 형태로 추출하는 것을 목표로 한다. 네트워크의 구성은 Crop-to-PanoLDR 네트워크와 LDR-to-HDR 네트워크 두 단계로 구성된다. 먼저, Crop-to-PanoLDR 네트워크에서는 단일 LDR 이미지로부터 실내와 실외를 구분하고, 광원의 위치와 추출한 광원 정보를 담은 환경 맵을 추정한다. 다음으로 LDR-to-HDR 네트워크에서는 앞서 추정된 환경 맵을 이용하여 가상 객체를 조명할 수 있는 HDR 이미지를 생성한다. 이와 같은 과정을 거쳐 최종적으로 생성된 HDR 환경 맵은 가상환경에서 가상 객체를 렌더링할 때 적용되어 가상 객체에 비춰지는 빛의 방향과 주변광 등을 확인하게 된다.

## 2. 관련 연구

### 2.1 광원 추출

Debevec은 거울 공과 같은 라이트 프로브(Light probe)의 노출을 다양하게 하여 여러 사진을 촬영하는 것이 전방향 HDR 환경 맵을 복원하는데 사용될 수 있음을 보여주었다 [1, 8]. 이를 이용하여 가상 오브젝트를 실제 장면에서 사실적으로 렌더링하는 데 사용하였다. 또한, 금속 (Metallic) 혹은 분산 (Diffuse) 물질을 포함한 하이브리드 구를 촬영한 단일 이미지를 사용하여 전방향 HDR 환경 맵을 복원하는 동일한 작업을 수행할 수 있음을 증명하였다 [9]. 하지만 이와 같이 특수 장비를 사용하여 촬영하는 경우 시간과 비용이 많이 드는 단점이 존재한다. 또한 증강현실과 같은 실시간 어플리케이션에서는 새로운 환경 정보를 바로 처리해야하지만 위의 소개된 방법으로는 어려움이 있다. 따라서 최근 딥러닝을 이용하여 이러한 단점을 보완하는 연구가 활발히 진행되고 있다.

실내에서의 광원은 천장의 조명과 램프, 창문, 가구 표면에 반사된 밝은 영역과 거울에 반사된 광원 등 여러 종류가 있으며 위치 또한 정해져 있지 않아 기하학적인 구조를 포함한 다양한 점을 고려해야 한다. Gardner와 그의 동료들은 원본 파노라마에서 광원을 추정할 실내 이미지에 해당하는 부분에 카메라를 배치하여 새로운 파노라마 이미지를 생성하는 파노라마 재조정 왜곡 연산자를 도입하였다. 이를 통해 제한된 시야를 가진 실내 이미지에 해당하는 기하 정보를 얻어 최종적으로 추출된 기하에 적합한 광원을 추정하였다 [2]. Song과 그의 동료들은 광원을 추정하고자 하는 실내 이미지의 3D 기하 정보를 추정하기 위해 구형 투영으로 관찰된 왜곡된 픽셀을 활용하고, 관찰되지 않은 픽셀 영역에 대한 색상 정보를 예측한 뒤 HDR 환경 맵으로 생성하는, 세 단계로 분할하여 해결하는 방법을 제안하였다 [10]. 또한, 실내 이미지의 기하학적 정보와 함께 3D 광원을 묶어 최종적으로 광원을 추정하는 연구도 진행되었다 [3]. 이와 반대로, Garon과 그의 동료들은 광원을 추출할 이미지의 기하학적인 구조 정보를 구하는 대신 이미지를 여러 패치로 분할하여 네트워크가 각 패치에 알맞은 광원을 추정하도록 학습하는 방법을 제안 하였다 [4].

실외 광원 추출 연구는 주광원인 태양을 중심으로 태양과 하늘의 특성을 담은 모델을 생성하는 방향으로 진행되어왔다. Hold-Geoffroy와 그의 동료들은 컨볼루션 신경망 (Convolutional neural network, CNN) 구조의 네트워크를 이용하여 단일 LDR 이미지로부터 하늘 모델을 구성하는 여러 파라미터들을 추정하고 최종적으로 가상 객체를 조명할 수 있는 HDR 환경 맵의 형태로 변환하는 방법을 제안하였다 [5]. 하지만 실외 장면의 빛 환경은 날씨의 조건에 따라 크게 달라지기 때문에 Zhang과 그의 동료들은 다양한 날씨 조건에 적합한 광원을 추출할 수 있는 방법을 제안 하였다 [7]. 더 나아가 Hold-Geoffroy와 그의 동료들은 날씨 조건 뿐만 아니라 시간의 변화에 따른 적합한 광원을 추정하는 방법을 세 단계의 네트워크로 접근하였다 [6]. 먼저 HDR 하늘 파노라마

데이터를 이용하여 하늘 모델을 학습하는 오토인코더 (Autoencoder)와, LDR 파노라마 데이터를 이용하여 첫 번째 네트워크의 잠재 벡터 (Latent vector)를 획득하기 위한 학습 네트워크, 마지막으로 단일 이미지로부터 앞서 학습된 네트워크의 잠재 벡터를 추정하고 이를 HDR 환경 맵으로 매핑하는 방법을 학습하는 네트워크로 구성했다.

본 연구에서 제안하는 방법은 기존의 실내와 실외를 구분하여 다른 방법을 사용한 연구와 달리, 하나의 딥러닝 방법을 통해 실내와 실외를 구분하고 그에 맞는 광원을 추출할 수 있다. Gardner와 동료들은 [2] 하나의 인코더와 두 개의 디코더로 이루어진 네트워크 구조를 사용하여 실내 광원을 추출하는 연구를 진행하였다. 이와 다르게 본 연구에서는 실내와 실외를 구분하는 분류 헤드 (Classification head)를 추가하여 다중 작업을 수행할 수 있는 네트워크 구조로 설계함으로써 최종적으로 Crop-to-Pano 네트워크를 통해 실내와 실외를 구분하여 그에 알맞은 광원을 추출할 수 있도록 하였다.

## 2.2 HDR 재구성을 위한 딥러닝

일반적으로 8비트 LDR 이미지를 32비트 HDR로 변환하는 방법은 역 톤 매핑 (Inverse tone mapping)이라고 한다. HDR 이미지를 LDR 이미지로 줄이는 톤 매핑 (Tone mapping) 방법을 반대로 진행하는 방법으로, 이를 위해 특정 함수나 알고리즘이 사용된다 [11, 12, 13, 14, 15, 16]. 이 방법을 사용하여 변환된 HDR 이미지의 품질은 사용되는 특정 함수나 알고리즘에 따라 크게 달라지게 된다. 따라서 최근에는 복잡한 역 톤 매핑 방법이 아닌 딥러닝을 이용한, LDR 이미지에서 HDR 이미지로의 매핑 방법을 네트워크가 학습하도록 하는 연구가 활발히 진행되고 있다.

Lee와 그의 동료들은 역 톤 매핑 문제를 해결하기 위해 생성적 적대 신경망 (generative adversarial network, GAN) 기반의 딥러닝 네트워크 구조를 제안하였다 [17]. 이는 조건부 생성적 적대 신경망 (conditional generative adversarial network, cGAN)을 사용하여 추정된 다중 노출 축적을 기반으로 HDR 이미지를 생성하는 최초의 프레임워크로 제안되었다. 이와 유사하게 Ning과 그의 동료들에 의해 생성적 적대 정규화를 사용하는 역 톤 매핑 네트워크 연구도 진행되었는데 이들이 제안한 네트워크는 U-Net 기반 HDR 이미지 생성기 (Generator)와 간단한 컨볼루션 신경망 기반 판별기 (Discriminator)로 구성되며 콘텐츠 관련 손실 함수를 이용하여 학습하였다. 하지만 역 톤 매핑을 위해 일반적인 손실 함수로 컨볼루션 신경망을 학습하게 될 경우 LDR 이미지와 HDR 이미지 간의 비선형 관계로 인해 학습이 되지 않는 문제가 발생하게 된다. 따라서 Kinoshita와 그의 동료들은 HDR 이미지를 정규화 할뿐만 아니라 LDR과 HDR 이미지 간의 비선형 관계를 줄일 수 있는 손실 함수를 제안하였다 [18]. 반면에 Liu와 그의 동료들은 동적 범위 클리핑, 비선형 매핑 및 양자화로 구성된 3단계의 HDR-to-LDR 이미지 형성 파이프 라인을 모델링한 뒤 각 단계를 역으로 학습하여 LDR 이미지 형성 파이프 라인을

모델링한 방법을 제안하였다 [19]. 또한, Zhang과 그의 동료들은 단일 LDR 360 파노라마에서 실외 조명의 HDR를 추정하는 딥러닝 기반 역 톤 매핑 방법을 연구하였다 [20]. 이는 실제 HDR 하늘 광원을 이용하여 현실적인 가상 도시 모델에 조명을 비추어 합성된 화면을 대량으로 캡처한 데이터셋을 사용하여 네트워크를 학습하였으나 일반적인 역 톤 매핑 방법과 같이 톤 매핑 방법을 반대로 진행하여 학습하였다.

이와 반대로, Gardner와 그의 동료들 그리고 Song과 그의 동료들은 특정 함수나 알고리즘 없이 로그 스케일로 계산된 빛의 강도를 사용하여 딥러닝 학습을 통해 LDR 이미지에서 HDR 이미지로 재구성하는 연구를 진행하였다 [10]. 본 연구는 이러한 HDR 재구성 방법을 3차원 공간에서 가상 객체를 조명하기 위한 HDR 환경 맵을 생성하는데 적용하였다. 이를 통해 제안한 LDR-to-HDR 네트워크가 LDR 이미지와 HDR 이미지 간의 비선형 관계를 학습하여 생성된 HDR 환경 맵으로 가상 객체를 렌더링하여 알맞은 조명을 비추는지 확인하였다.

## 3. 제안하는 방법

본 논문에서 제안하는 네트워크는 Crop-to-PanoLDR과 LDR-to-HDR 총 두 단계로 구성된다. 먼저, Crop-to-PanoLDR 네트워크를 활용하여 단일 LDR 이미지로부터 실내와 실외를 구분하고, 광원의 위치와 추출한 광원 정보를 담은 LDR 환경 맵을 생성한다. 다음으로 LDR-to-HDR 네트워크를 활용하여 앞서 생성된 LDR 환경 맵을 가상 객체를 조명할 수 있는 형태인 HDR 환경 맵으로 최종 생성하게 된다. 본 절에서는 각 네트워크의 구조를 설명하고 학습에 사용된 데이터셋과 손실 함수에 대해 설명한다.

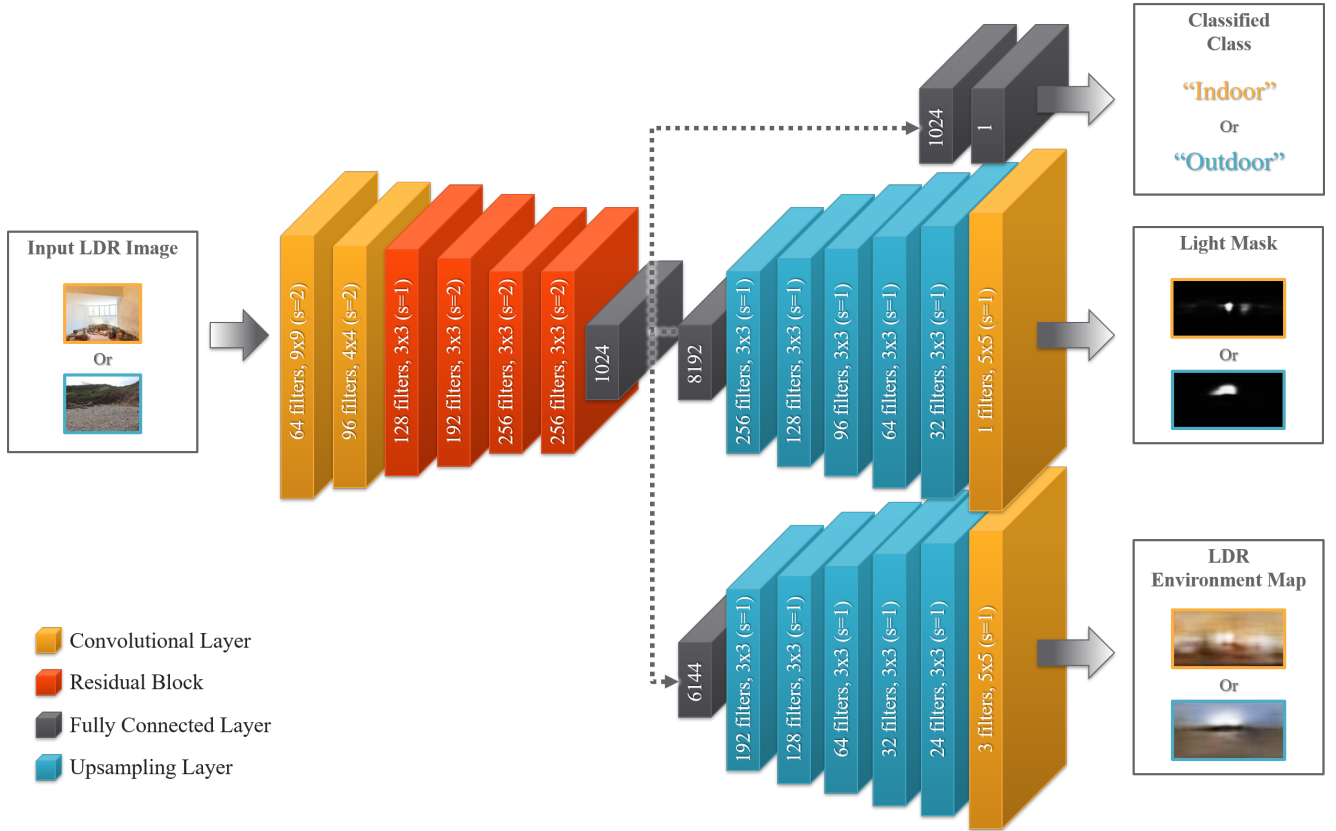
### 3.1 Crop-to-PanoLDR 네트워크

#### 3.1.1 네트워크 구조

Crop-to-PanoLDR 네트워크의 구조는 실내와 실외를 구분하여 학습하는 기존의 광원 추출 연구의 네트워크 구조 [2]와는 달리 다중 작업을 수행하는 네트워크의 구조로 디자인 하였다. Figure 1에서 확인할 수 있듯이 Crop-to-PanoLDR 네트워크는 다중 작업을 수행할 수 있도록 하나의 인코더와 두 개의 디코더, 그리고 분류 헤드로 구성되어 있다. 이는 입력 이미지를 실내 혹은 실외로 구분하여 네트워크 학습 과정 중 입력 데이터인 단일 LDR 이미지에서 추정하고자 하는 LDR 환경 맵과 빛 마스크를 구분해 맞게 생성하는데 영향을 주기 위한 디자인이다.

입력 이미지부터 순서대로 설명하자면, 제한된 시야를 가진 단일 LDR 이미지는 두 개의 컨볼루션 레이어 (Convolutional layer)와 다섯 개의 잔차 블록 (Residual block) [21]으로 구성된 인코더에 입력으로 들어가게 된다. 이후 Fully-connected 레이어를 통해 분류 헤드와 두 개의 디코더로 각각 전달된다. 먼저, 분류 헤드는 두 개의 Fully-connected 레이어로 구성되어 입력으로 들어온



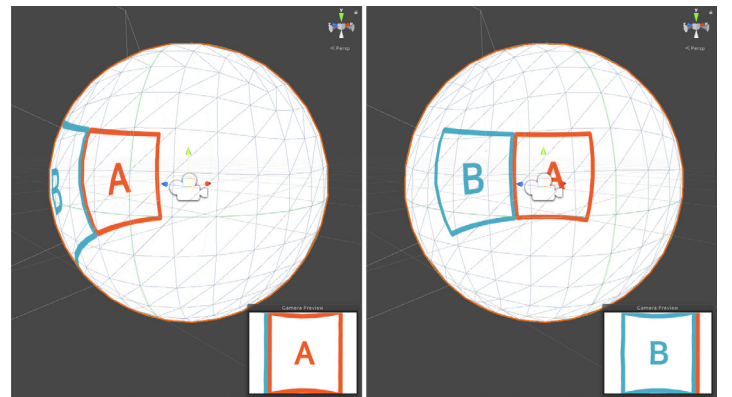


**Figure 1:** Crop-to-PanoLDR network architecture. The notation on each layer means the number of filters, the size of those filters, and the stride. The number on fully connected layers means the size of the layer.

이미지가 실내를 배경으로 한 이미지인지 실외인지를 추정하여 각각의 클래스로 구분하는 역할을 수행한다. 다음으로, 디코더는 빛 마스크를 생성하는 디코더와 LDR 환경 맵을 추정하는 디코더로 나뉘어진다. 두 디코더의 구조는 다섯개의 업샘플링 레이어 (Upsampling layer)와 한개의 컨볼루션 레이어의 같은 구조를 가졌지만 추정하고자 하는 이미지에 맞게끔 각 레이어의 필터 크기와 같은 세부적인 설정을 달리하여 디자인되었다. 그리하여 각각의 디코더는 입력으로 들어온 단일 LDR 이미지에서 광원의 위치와 범위를 추정하여 빛 마스크를 생성하고, 광원과 주변 배경의 RGB 색상을 추정하여 최종 LDR 환경 맵을 생성하는 역할을 수행한다. 추가적으로 Crop-to-PanoLDR 네트워크의 모든 레이어에는 배치 정규화를 적용하였으며 분류 헤드와 빛 마스크를 생성하는 디코더에는 Sigmoid 활성화 함수를, LDR 환경 맵을 생성하는 디코더에는 Tanh 활성화 함수를 사용하였다.

### 3.1.2 데이터셋

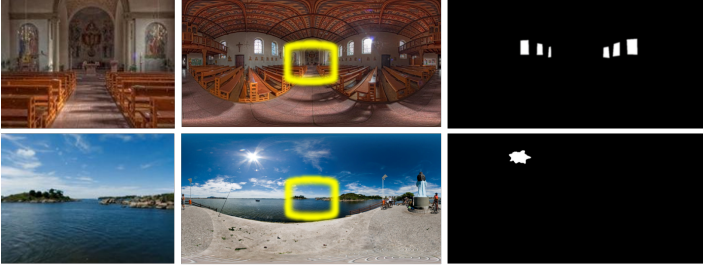
Crop-to-PanoLDR 네트워크를 학습하기 위한 데이터셋은 SUN360 파노라마 데이터베이스 [22]를 사용하여 만들었다. 해당 데이터베이스는 실내 및 실외를 배경으로 한 다양한 환경 전체를 담은 파노라마로 구성되어있으며, Crop-to-PanoLDR 네트워크를 학습하기 위해 302개의 실내를 배경으로한 파노라마와 1,029개의 실외를 배경으로 한 파노라마를 선택하여



**Figure 2:** Visual explanation of data augmentation. Each column shows the image of before rotation (left) and after rotation (right), respectively.

그라운드 트루스 (Ground truth) LDR 환경 맵으로 사용하였다. 여기서 광원의 위치와 영역을 나타내는 빛 마스크의 그라운드 트루스 데이터는 그라운드 트루스 LDR 환경 맵에서 밝은 픽셀값들을 찾아 광원으로 추정하는 방법을 사용해 생성하였다. 이 때, 실내를 배경으로 한 LDR 환경 맵의 빛 마스크는 실내 광원을 스포트라이트, 램프, 창문, 반사 등 총 네 가지의 광원으로 분류하여 위치와 영역을 라벨한 데이터 [2]를 사용하였으며, 실외의 경우 픽셀의 밝기 값이 임계 값 이상일 때 광원으로





**Figure 3:** Dataset for training the Crop-to-PanoLDR network. The top row shows indoor image and the bottom row shows an outdoor image. Each column shows a cropped LDR image (left), the ground truth LDR environment map (middle), and the ground truth light mask (right).

분류하는 방법 [23]을 사용하여 태양 즉, 주광원 하나에 대한 위치와 영역을 추정된 빛 마스크를 생성하였다. 실내와 실외의 모든 빛 마스크는 가우시안 흐림 효과 (Gaussian blur)를 적용한 뒤 네트워크 학습에 사용하였다. 마지막으로 입력 데이터로 사용되는 제한된 시야를 가진 이미지는 그라운드 트루스 LDR 환경 맵의 일부분을 제한된 시야를 가진 카메라의 뷰로 캡처하여 사용하였다.

SUN360 파노라마 데이터베이스에서 선택한 파노라마 데이터의 수가 네트워크를 학습하기에 충분한 양이 되지 않아 네트워크를 학습하기 전 데이터 증강을 수행하였다. 데이터 증강은 그라운드 트루스 LDR 환경 맵으로부터 입력 이미지를 생성하는 것과 같이, 파노라마 이미지를 3차원 구에 감싸듯이 투영하고 구의 중앙에 뷰가  $\pi \times 65.5/180$  으로 고정된 카메라를 배치한 뒤 구를 y축을 기준으로 조금씩 회전시켜 카메라 뷰에 비춰지는 달라진 화면을 캡처하는 방법으로 수행하였다 (Figure 2). 이러한 방법은 Crop-to-PanoLDR 네트워크의 입력 데이터가 제한된 시야를 가진 이미지이기 때문에 가능한 방법으로, 하나의 파노라마 이미지에서 다른 뷰를 가진 여러 이미지를 추출할 수 있다. 그러나 본 연구는 학습과정에서 입력 LDR 이미지와 추출되는 LDR 환경 맵과의 관계에 카메라 정보를 이용하지 않기 때문에 추출되는 LDR 환경 맵의 중앙 부분의 이미지라는 전제 조건을 가진다고 가정하여 y축을 기준으로 회전시키는 방법의 데이터 증강을 수행하였다. 최종적으로 본 연구의 데이터 증강 방법을 통해 12,080개의 실내와 16,464개의 실외의 제한된 시야를 가진 입력 LDR 이미지, 그라운드 트루스 LDR 환경 맵, 빛 마스크를 생성하였다 (Figure 3). 여기서 Gardner와 그의 동료들의 연구를 참고하여 [2] 입력 LDR 이미지의 사이즈는  $256 \times 192$ , 그라운드 트루스 LDR 환경 맵과 빛 마스크는  $256 \times 128$ 로 설정하였다.

### 3.1.3 손실 함수

Crop-to-PanoLDR 네트워크의 학습에 사용된 손실 함수는 세 가지로 구성된다. 먼저, 실내와 실외 클래스를 구분하는 분류 헤드에서 그라운드 트루스 클래스  $g_{class}$ 와 추정된 클래스  $p_{class}$ 를

비교하는 손실 함수  $\mathcal{L}_{class}$ 는 이진 교차 엔트로피 (Binary cross entropy) 함수를 사용하였다. 이진 교차 엔트로피 함수 사용의 이유는 Crop-to-PanoLDR에 존재하는 분류 헤드가 실내와 실외 두 개의 분류를 목적으로 하기 때문이다.

$$\mathcal{L}_{class}(p_{class}, g_{class}) = - (g_{class} \log(p_{class}) + (1 - g_{class}) \log(1 - p_{class})) \quad (1)$$

다음으로 이미지의 가로×세로×3인 총 N개의 픽셀에 대하여, 빛 마스크를 생성하는 디코더에서는 그라운드 트루스 빛 마스크  $g_{lm}$ 와 추정된 빛 마스크  $p_{lm}$  사이의 차이를 비교하는 손실 함수  $\mathcal{L}_{lm}$ , 그리고 LDR 환경 맵을 추정하는 디코더에서는 그라운드 트루스 LDR 환경 맵  $g_{LDR}$ 과 추정된 LDR 환경 맵  $p_{LDR}$  사이의 차이를 비교하는 손실 함수  $\mathcal{L}_{LDR}$ 로 L2 손실 함수를 사용한다.

$$\mathcal{L}_{lm}(p_{lm}, g_{lm}) = \frac{1}{N} \sum_{i=1}^N (p_{lmi} - g_{lmi})^2 \quad (2)$$

$$\mathcal{L}_{LDR}(p_{LDR}, g_{LDR}) = \frac{1}{N} \sum_{i=1}^N (p_{LDRi} - g_{LDRi})^2 \quad (3)$$

최종적으로, 위의 세 가지 손실 함수에 각각의 가중치를 곱한 뒤 모두 합하여 최종 손실 함수  $\mathcal{L}_{panoLDR}$ 로 구성하고 이를 학습에 사용하였다. 각각의 가중치 값은 Hyperparameter를 조절한 다양한 실험 결과를 통해 적절한 값인  $w_1 = 100, w_2 = 1, w_3 = 1$ 로 설정하였다.

$$\mathcal{L}_{panoLDR}(p, g) = w_1 \mathcal{L}_{LDR}(p_{LDR}, g_{LDR}) + w_2 \mathcal{L}_{lm}(p_{lm}, g_{lm}) + w_3 \mathcal{L}_{class}(p_{class}, g_{class}) \quad (4)$$

## 3.2 LDR-to-HDR 네트워크

### 3.2.1 네트워크 구조

LDR-to-HDR 네트워크의 역할은 Crop-to-PanoLDR 네트워크에서 생성된 LDR 환경 맵을 HDR 환경 맵으로 변환하는 것이다. 이와 관련하여 LDR 이미지에서 HDR 이미지로 재구성하는 이전 연구를 살펴보면 대부분 역 톤 매핑 방법을 사용하여 네트워크를 학습한다. 하지만 본 연구는 Song과 그의 동료들의 [10] 아이디어에서 영감을 얻어, 네트워크가 로그 스케일로 계산된 빛의 강도를 이용하여 LDR 이미지 공간에서 HDR 이미지 공간으로의 매핑 방법을 학습하도록 하였다. 또한, LDR-to-HDR 네트워크 구조는 Resnet의 생성자를 기반으로 한 구조 [24]를 적용하여 구성하였다. 그리하여 LDR-to-HDR 네트워크는 Crop-to-PanoLDR 네트워크의 출력인 LDR 환경 맵에서 HDR 환경 맵을 생성할 수 있게 된다.



**Figure 4:** LDR-to-HDR network architecture. The notation on each layer means the number of filters, the size of those filters, and the stride.

Figure 4와 같이 LDR-to-HDR 네트워크 구조는 세 개의 컨볼루션 레이어와 여섯 개의 잔차 블록, 두 개의 업샘플링 레이어, 그리고 마지막으로 한 개의 컨볼루션 레이어로 구성하였다. 여기서 두 업샘플링 레이어는 Zhu와 그의 동료들의[24] 생성자 구조에서 사용된 전치 컨볼루션 레이어 (Transpose convolutional layer)가 아닌 더 나은 결과를 얻을 수 있는 이중 선형 업샘플링 레이어를 사용하였다. 마지막으로 LDR-to-HDR 네트워크의 모든 레이어에는 ReLU 활성화 함수와 인스턴스 정규화를 사용하였다.

### 3.2.2 데이터셋



**Figure 5:** Dataset for training LDR-to-HDR network. The top row shows an indoor image and the bottom row shows an outdoor image. Each column shows the ground truth LDR environment map (left) and the ground truth HDR environment map (right), respectively

LDR-to-HDR 네트워크는 Crop-to-PanoLDR 네트워크에서 생성된 LDR 환경 맵을 입력으로 했을 때 HDR 환경 맵을 생성하는 역할을 수행한다. 하지만 Crop-to-PanoLDR 네트워크를 학습하는데 사용하는 데이터셋은 HDR 환경 맵 그라운드 트루스가 존재하지 않는 SUN360 파노라마 데이터베이스이다. 따라서, LDR-

to-HDR 네트워크를 학습시키기 위한 데이터셋으로는 실내를 배경으로 한 HDR 파노라마로 구성된 Laval Indoor HDR 데이터베이스 [2]와 실외를 배경으로 한 HDR 파노라마로 구성된 Laval Outdoor HDR 데이터베이스 [6]를 사용한다. 이러한 HDR 데이터들을 그라운드 트루스 HDR 환경 맵으로 사용하고 이를 LDR로 변환하여 입력 데이터인 LDR 환경 맵으로 사용하여 네트워크를 학습시킬 수 있게 된다. 그럼에도 불구하고 Crop-to-PanoLDR 네트워크의 출력인 LDR 환경 맵과 그라운드 트루스 HDR 환경 맵을 LDR로 변환한 데이터 사이에는 큰 차이가 존재하게 된다. 이러한 차이를 보완하기 위하여 본 연구에서는 LDR-to-HDR 네트워크를 학습하기 위해 그라운드 트루스 HDR 환경 맵으로부터 변환한 LDR 환경 맵을 Crop-to-PanoLDR 네트워크의 출력 LDR 환경 맵과 유사하게 만들기 위해 이미지 사이즈를  $256 \times 128$ 로 설정하고 가우시안 흐림 효과를 적용하였다. 또한, LDR-to-HDR 네트워크는 LDR 환경 맵이 가진 광원 정보가 HDR 환경 맵으로 적절히 반영되는 것을 목표로 하고 있으며, LDR 이미지에서 HDR로 생성하였을 때 모든 픽셀값이 정확히 생성되도록 하는 일반적인 HDR 재구성 연구와는 다른 목표를 가지고 있다. 따라서 앞서 LDR 환경 맵에 적용된 내용을 그라운드 트루스 HDR 환경 맵에도 똑같이 적용하여 LDR-to-HDR 네트워크가 학습을 통해 추정하는 HDR 환경 맵에서 광원 정보가 올바르게 추정되었는지 비교할 수 있도록 하였다. LDR-to-HDR 네트워크의 학습을 위한 데이터셋은 실내와 실외를 포함하여 2,233개의 HDR 환경 맵으로 구성하였으며 Figure 5에서 확인할 수 있다.

### 3.2.3 손실 함수

LDR-to-HDR 네트워크 학습을 위한 손실 함수  $\mathcal{L}_{HDR}$ 는 L2 함수를 사용하여 추정된 HDR 환경 맵  $p_{HDR}$ 과 그라운드 트루스 HDR 환경 맵  $g_{HDR}$  사이의 차이를 계산한다.

$$\mathcal{L}_{HDR}(p_{HDR}, g_{HDR}) = \frac{1}{N} \sum_{i=1}^N (p_{HDRi} - g_{HDRi})^2 \quad (5)$$

## 4. 실험 결과

### 4.1 네트워크 학습

본 연구의 Crop-to-PanoLDR 네트워크와 LDR-to-HDR 네트워크의 학습을 위한 모든 데이터셋은 학습용 85%, 테스트용 15%으로 분할하여 사용되었다. 두 네트워크 모두 Nvidia GeForce GTX 1080ti GPU 환경에서 100 에포크 (Epoch)로 학습되었으며 개별적으로 학습할 수 있도록 하나의 GPU에 하나의 네트워크만 배치하였다. 두 네트워크의 학습을 위한 세부적인 내용으로는 미니 배치 (Mini-batch) 크기가 64이며 초기 학습률 (Learning rate)은  $1e-4$ 인 ADAM 옵티마이저 (Optimizer) [25]를 사용하였다. 최종적으로 Crop-to-PanoLDR 네트워크를 학습하는데 약 7시간 15분이 걸렸으며 테스트시 추론에는 약 7ms가 걸렸다. LDR-to-HDR 네트워크는 학습에 약 2시간 25분을 소요하였으며 테스트시 추론에는 약 9ms가 걸렸다.

### 4.2 Crop-to-PanoLDR 네트워크 분석

#### 4.2.1 학습 데이터셋의 구성 변경

본 연구의 Crop-to-PanoLDR 네트워크가 단일 LDR 이미지에서 실내 배경과 실외 배경을 구분하여 적절한 LDR 환경 맵을 생성할 수 있는지를 검증하기 위해 Crop-to-PanoLDR 네트워크의 학습 데이터셋의 구성을 다르게 하여 학습된 결과를 확인하는 실험을 진행하였다. 데이터셋은 실내를 배경으로한 이미지만으로 구성된 데이터셋 그리고 실외를 배경으로한 이미지만으로 구성된 데이터셋, 마지막으로 실내와 실외를 모두 포함하는 데이터셋 등 총 세 가지 구성으로 나누어 각자 Crop-to-PanoLDR 네트워크의 학습에 사용되도록 하였다.

Figure 6에서 보여지듯 실내와 실외를 모두 포함한 데이터셋으로 네트워크를 학습한 결과 (Figure 6의 오른쪽에서 첫 번째 열)가 실내 데이터셋으로만 학습한 결과 (Figure 6의 오른쪽에서 세 번째 열)와 실외 데이터셋으로만 학습한 결과 (Figure 6의 오른쪽에서 두 번째 열)보다 훨씬 나은 결과를 보이고 있음을 확인할 수 있다. 실내 데이터셋으로만 학습한 결과에서는 실내 입력 이미지와 실외 입력 이미지 모두에 대하여 추정된 빛 마스크 결과가 분산되어 나타나 광원의 위치를 정확하게 추정하기 어려운 것을 확인할 수 있다. 또한, 추정된 LDR 환경 맵은 입력 이미지가 실외인 경우 일정 부분 하늘색을 가질 수는 있으나 전체적으로 실내 이미지에 가까운 결과를 생성하는 것을 확인할 수 있다. 반면에 실외 데이터셋으로만 학습한 결과에서 실내 데이터셋으로만 학습한 결과와 비교하여 비교적 형태를 갖춘 빛 마스크를 생성하는 것을 확인하였으며, 입력 이미지가 실외에 해당하여 추정된 LDR

환경 맵이 실외 배경에 알맞는 결과를 생성하는 것을 확인할 수 있다. 하지만 입력 이미지가 실내인 경우 추정된 LDR 환경 맵은 실내에 해당하는 결과를 전혀 생성하지 못하고 완전히 실외에 해당하는 결과를 생성하는 것을 확인할 수 있다.

#### 4.2.2 실내와 실외 분류를 제외한 학습 결과

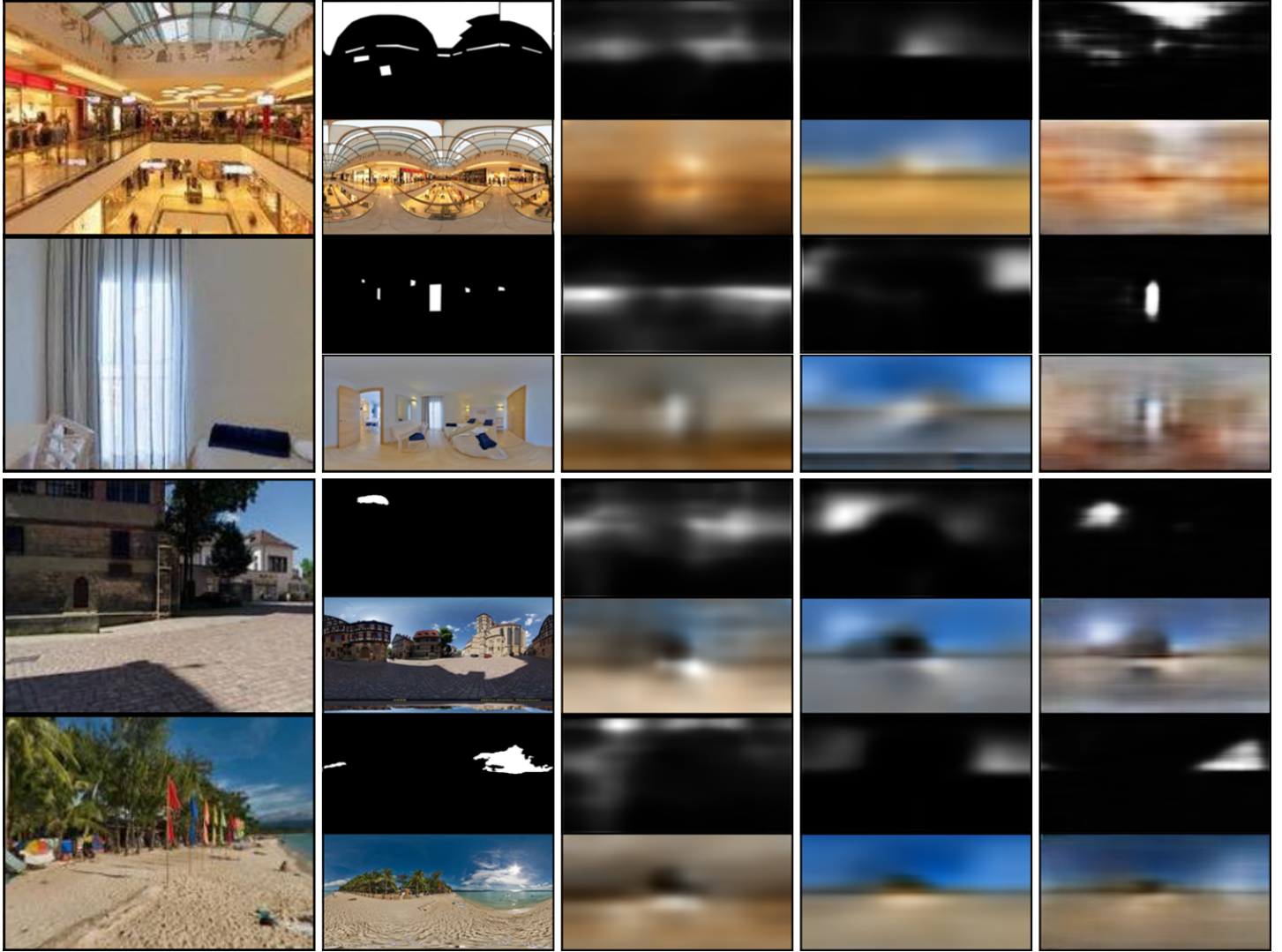
본 연구의 가장 핵심적인 부분이자 기존 연구와의 차이점인 Crop-to-PanoLDR 네트워크의 실내와 실외를 구분하는 다중 작업의 영향을 확인하기 위하여 다중 작업을 포함한 네트워크 학습 결과와 포함하지 않고 학습한 네트워크의 결과를 비교하는 실험을 진행하였다. 실험은 본 Crop-to-PanoLDR 네트워크의 손실 함수 (4)에서 네트워크 구조에 존재하는 분류 헤드를 활성화하는 손실 함수 (1)을 적용하여 네트워크를 학습하고, 반대의 경우인 분류 헤드를 비활성화하기 위해 손실 함수에 적용되는  $w_3$ 의 값을 0으로 설정하여 네트워크를 학습시킨 뒤 두 결과를 확인하는 것으로 진행되었다. 네트워크 학습에는 실내와 실외를 모두 포함한 데이터셋을 사용하였다.

실험 결과, Figure 7에서 보여지는 바와 같이 분류 헤드를 활성화하여 다중 작업을 포함한 네트워크의 학습이 다중 작업을 포함하지 않은 네트워크 학습보다 더 나은 결과를 생성하는 것을 확인할 수 있었다. 분류 헤드를 활성화하지 않은 네트워크 학습 결과 (Figure 7의 가장 오른쪽 열)에서 빛 마스크는 실내와 실외의 모든 입력 이미지에서 분산된 결과가 생성되어 광원의 위치와 영역을 정확히 구별할 수 없고, LDR 환경 맵 또한 적절하지 못한 결과를 생성하는 것을 확인할 수 있다. 따라서, Crop-to-PanoLDR 네트워크의 구조를 통한 다중 작업 학습 방법이 실내와 실외 배경 모두에 알맞은 광원 추출을 돕는 것을 확인하였다.

### 4.3 LDR-to-HDR 네트워크 분석

LDR-to-HDR 네트워크를 통해 최종적으로 생성된 HDR 환경 맵을 평가하기 위해 관련 연구와 비교하는 실험을 수행하였다. 하지만 기존의 광원 추출 연구는 실내와 실외를 구분하여 진행되었기 때문에 본 연구 결과도 실내와 실외로 구분하여 관련 연구와 비교하도록 하였다. 실험은 동일한 단일 LDR 이미지를 입력으로 하였을 때 본 연구에서 생성된 HDR 환경 맵과 관련 연구에서 생성된 HDR 환경 맵을 비교하는 것으로 진행되었다. 그러나 HDR 이미지의 픽셀별 값이 LDR 이미지와 비교하였을 때 굉장히 넓은 범위를 가지기 때문에 RMSE (Root mean squared error)와 같은 오차를 계산하는 것으로 비교하기가 어렵다. 따라서 본 실험은 입력 단일 LDR 이미지를 배경으로 하였을 때, HDR 환경 맵을 이용하여 렌더링된 가상 객체를 배치하여 배경에 알맞는 조명을 비추고 있는지 확인하는 것으로 진행하였다.





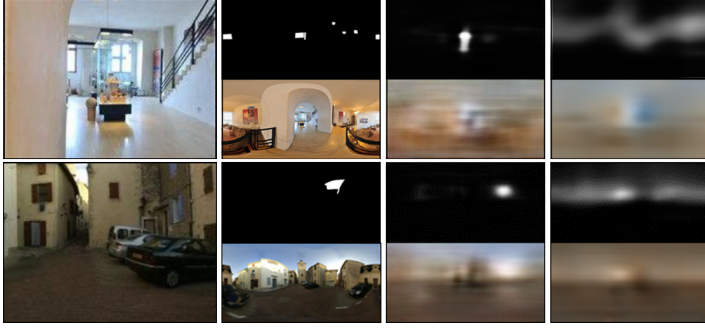
**Figure 6:** Results from training the Crop-to-PanoLDR network with different dataset compositions. Top 2 rows show indoor image and bottom 2 rows show outdoor image. From the left, each column shows the input images, the ground truth light mask (top) and the ground truth LDR environment map (bottom), an estimated light mask and LDR environment map by the trained network with only indoor dataset, an estimated light mask and LDR environment map by the trained network with only outdoor dataset, an estimated light mask and LDR environment map by the trained network with both indoor and outdoor dataset, respectively.

#### 4.3.1 실내 광원 추출 연구와 비교 실험

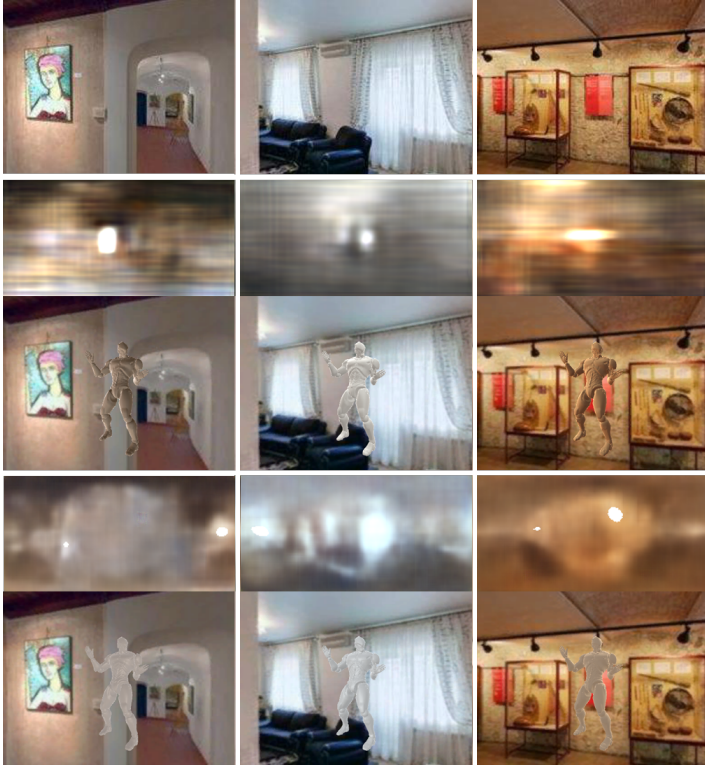
실내를 배경으로 한 단일 LDR 이미지를 입력 데이터로 하였을 때, 본 연구의 네트워크를 통해 추출된 광원인 HDR 환경 맵과 Gardner와 그의 동료들의 [2] 네트워크를 통해 추출된 HDR 환경 맵을 비교하였다. 비교 실험 결과, Figure 8에서와 같이 본 연구 결과에서 가상 객체의 표면에 조명되는 주변광의 색상이 배경인 입력 이미지의 주변 색상을 적절히 표현하고 있으며 광원의 위치에 따른 빛의 방향과 그림자 표현 또한 기존의 연구와 비교하여 뚜렷이 잘 나타나고 있음을 확인할 수 있었다. 따라서 해당 실험을 통해 본 연구는 실내를 배경으로 한 단일 LDR 이미지에 적합한 광원을 추정할 수 있다고 설명할 수 있다.

#### 4.3.2 실외 광원 추출 연구와 비교 실험

실외를 배경으로 한 단일 LDR 이미지를 입력 데이터로 하였을 때, 본 연구의 네트워크를 통해 추출된 광원인 HDR 환경 맵과 Hold-Geoffroy와 그의 동료들의 [5] 네트워크를 통해 추출된 HDR 환경 맵을 비교하였다. Figure 9에서 확인할 수 있듯이, Hold-Geoffroy와 그의 동료들의 결과와 비교하여 본 연구의 결과가 가상 객체의 표면에 조명되는 주변광의 색상을 뚜렷하게 더 잘 표현하고 있다. 이러한 이유는 Hold-Geoffroy와 그의 동료들은 실외를 배경으로 한 단일 LDR 이미지로부터 광원을 추출할 때 주광원인 태양만을 고려하여 하늘 영역만 추출하기 때문이다.



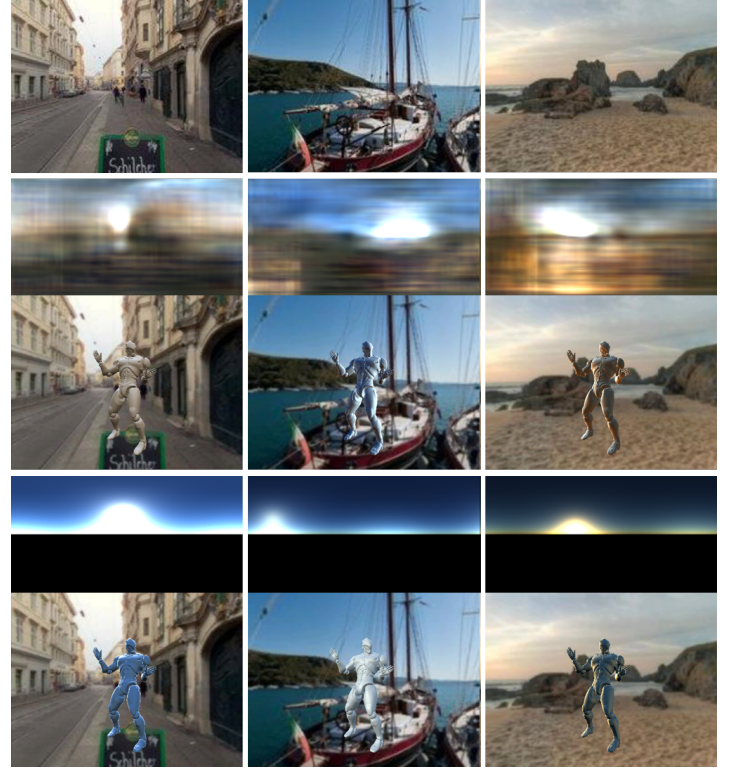
**Figure 7:** Results from training the Crop-to-PanoLDR network with and without a classification task. The top row shows an indoor image and the bottom row shows an outdoor image. From the left, each column shows the input images, the ground truth light mask (top) and the ground truth LDR environment map (bottom), an estimated light mask (top) and the LDR environment map produced by the trained network with a classification task (bottom), an estimated light mask (top) and the LDR environment map produced by the trained network without a classification task (bottom), respectively.



**Figure 8:** Indoor illumination comparison results. From the top, each row shows an input image, an HDR environment map estimated by our method and a rendered virtual object, an HDR environment map estimated by Gardner et al. [2] and a rendered virtual object, respectively.

#### 4.3.3 사용자 테스트

LDR-to-HDR 네트워크 분석을 위한 마지막 실험으로 앞서 기존의 연구 Gardner와 그의 동료들의 연구 [2]와 Hold-Geoffroy와 그의 동료들의 연구 [5]를 대상으로 비교 실험한 렌더링된 가상 객



**Figure 9:** Outdoor illumination comparison results. From the top, each row shows an input image, an HDR environment map estimated by our method and a rendered virtual object, an HDR environment map estimated by Hold-Geoffroy et al. [5] and a rendered virtual object, respectively.

체 이미지를 이용하여 사용자 테스트를 진행하였다. 본 연구의 네트워크를 통해 추출된 광원으로 렌더링한 가상 객체의 이미지와 기존의 광원 추출 연구를 통해 추출된 광원으로 렌더링한 가상 객체의 이미지를 사용자에게 동시에 보여주며 배경의 이미지에 자연스럽게 조명되고 있는 가상 객체 이미지를 선택하도록 하였다. 해당 실험을 위해 실내를 배경으로 한 가상 객체 이미지 30개와 실외를 배경으로 한 가상 객체 이미지 30개를 준비하여 사용자에게 무작위로 제시하는 방법으로 진행하였다. 본 실험에 참가한 참가자는 컴퓨터 그래픽스에 대한 지식이 없는 5명을 포함하여 총 20명으로 평균 연령이 28세인 6명의 여성과 14명의 남성으로 구성되었다.

실험 결과, 20명을 대상으로 60개의 질문에 답한 1200개의 응답 중 본 연구의 결과로 렌더링 된 가상 객체 이미지를 선택한 응답이 804개로 기존의 연구 결과로 렌더링 된 가상 객체 이미지를 선택한 응답 396개보다 두 배 이상으로 많았다. 이를 이항 검정 (Binomial test) 방법으로 분석한 결과 p값이 0.05미만 ( $p < 0.05$ )으로 유의한 차이를 확인할 수 있었다. 따라서 추출된 광원으로 가상 객체를 렌더링 하였을 때, 본 연구 결과로 렌더링 된 가상 객체 이미지가 기존의 연구 결과로 렌더링 된 가상 객체 이미지보다 더 자연스럽게 표현되며 사용자가 더 선호한다는 것을 확인할 수 있었다.



## 5. 결론 및 향후 연구

본 연구는 실내 및 실외를 배경으로 한 이미지 모두에서 적합한 광원을 추출하는 딥러닝 방법을 제안한다. 본 연구에서 제안된 딥러닝 방법은 다중 작업을 수행하는 네트워크의 구조를 가지며, 실내와 실외 이미지를 구분하는 작업을 수행하는 분류 헤드를 포함한 학습 방법과 포함하지 않은 학습 방법을 결과를 통해 비교하여 분류 헤드를 포함한 본 연구의 네트워크 구조가 실내와 실외 모두에서 광원을 추출하는데 큰 영향을 미치는 것을 확인하였다. 또한, 기존의 실내와 실외를 구분하여 광원을 추출한 연구와 비교하는 실험을 진행하여 본 연구 방법이 실내와 실외를 배경으로 한 이미지 모두에서 적절한 광원을 추출하는 것을 확인하였다.

한편, 본 연구에서 제안한 딥러닝 방법은 작은 영역을 가진 입력 이미지로부터 빛 마스크와 환경 맵을 추정하기 때문에 제한된 품질을 가지게 된다는 한계가 존재한다. 예를 들어, 입력 이미지에 광원이 존재할 경우 Crop-to-PanoLDR 네트워크를 통해 추정되는 빛 마스크와 LDR 환경 맵이 정확한 광원 정보를 담을 수 있으나 반대로 입력 이미지에 광원이 존재하지 않을 경우 비교적 정확한 광원 정보를 추정하기 어려울 수 있다. 게다가 본 연구에서 LDR 환경 맵을 추출할 때 카메라 방향이나 위치를 고려하지 않기 때문에 입력 이미지는 추출되는 LDR 환경 맵의 중앙 부분이어야 한다는 전제 조건을 가지고 있다. 또한, 본 연구의 딥러닝 방법을 통해 최종적으로 생성되는 HDR 환경 맵은 표면에 주변 환경이 정확히 반사되어야 하는 거울과 같은 텍스처를 가진 가상 객체를 렌더링하는데 적합하지 않다는 한계가 존재한다. 이러한 한계들은 최근 생성적 적대 신경망 기반의 딥러닝 네트워크 구조를 이용하여 정확한 픽셀 값을 생성하는 연구를 참조하고 카메라 방향과 같은 추가적인 정보를 추출함으로써 보다 정확한 HDR 환경 맵을 생성할 수 있을 것이다.

본 연구에서 제안된 딥러닝 방법은 실내와 실외를 구분하여 광원 정보를 추정하는 작업을 수행하는 Crop-to-PanoLDR 네트워크와, LDR 환경 맵을 HDR 환경 맵으로 추출하는 작업을 수행하는 LDR-to-HDR 네트워크로 구성된다. 그리고 두 네트워크의 학습에 사용된 데이터셋은 각 네트워크가 수행하는 작업에 알맞는 데이터로 구성되었으며 효율적인 학습 결과를 기대하기 위해 두 네트워크를 독립적으로 학습하였다. 하지만 이러한 학습방법을 통해 생성된 결과는 LDR 이미지로부터 추출된 HDR 환경 맵으로 그라운드 트루스 HDR 환경 맵과 직접적인 비교 실험을 진행하는데 어려움이 발생한다. 따라서 향후 연구로는 end-to-end 학습이 가능한 네트워크 설계를 통한 광원 추출 연구를 진행하여 효율적인 학습 결과를 본 연구와 비교하고 추출된 HDR 환경 맵의 정확한 정량적 평가를 진행할 수 있도록 할 계획이다.

## 감사의 글

이 논문은 2021년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No.2020-0-

00450, AR 실감 콘텐츠 제작을 위한 심층학습 기반 단일 이미지의 입체적 움직임 자동 생성 및 대화형 증강 기술 개발)

## References

- [1] P. Debevec, "Rendering synthetic objects into real scenes: Bridging traditional and image-based graphics with global illumination and high dynamic range photography," 1998.
- [2] M.-A. Gardner, K. Sunkavalli, E. Yumer, X. Shen, E. Gambaretto, C. Gagné, and J.-F. Lalonde, "Learning to predict indoor illumination from a single image," 2017.
- [3] M.-A. Gardner, Y. Hold-Geoffroy, K. Sunkavalli, C. Gagné, and J.-F. Lalonde, "Deep parametric indoor lighting estimation," 2019.
- [4] M. Garon, K. Sunkavalli, S. Hadap, N. Carr, and J.-F. Lalonde, "Fast spatially-varying indoor lighting estimation," 2019.
- [5] Y. Hold-Geoffroy, K. Sunkavalli, S. Hadap, E. Gambaretto, and J. Lalonde, "Deep outdoor illumination estimation," in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2373–2382.
- [6] Y. Hold-Geoffroy, A. Athawale, and J.-F. Lalonde, "Deep sky modeling for single image outdoor lighting estimation," 2019.
- [7] J. Zhang, K. Sunkavalli, Y. Hold-Geoffroy, S. Hadap, J. Eisenmann, and J.-F. Lalonde, "All-weather deep outdoor lighting estimation," 2019.
- [8] P. Debevec, "Image-based lighting," *IEEE Computer Graphics and Applications*, vol. 22, no. 2, pp. 26–34, 2002.
- [9] P. Debevec, P. Graham, J. Busch, and M. Bolas, "A single-shot light probe," in *ACM SIGGRAPH 2012 Talks*, ser. SIGGRAPH '12. New York, NY, USA: Association for Computing Machinery, 2012. [Online]. Available: <https://doi.org/10.1145/2343045.2343058>
- [10] S. Song and T. Funkhouser, "Neural illumination: Lighting prediction for indoor environments," *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6911–6919, 2019.
- [11] F. Banterle, P. Ledda, K. Debattista, and A. Chalmers, "Inverse tone mapping," in *Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia*, ser. GRAPHITE



- '06. New York, NY, USA: Association for Computing Machinery, 2006, p. 349–356. [Online]. Available: <https://doi.org/10.1145/1174429.1174489>
- [12] P. E. Debevec and J. Malik, “Recovering high dynamic range radiance maps from photographs,” in *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques*, ser. SIGGRAPH '97. USA: ACM Press/Addison-Wesley Publishing Co., 1997, p. 369–378. [Online]. Available: <https://doi.org/10.1145/258734.258884>
- [13] M. Fan, D.-H. Lee, S. Kim, and S.-J. Ko, “An optimization framework for inverse tone mapping using a single low dynamic range image,” *Signal Process. Image Commun.*, vol. 78, pp. 274–283, 2019.
- [14] P. Kuo, C. Tang, and S. Chien, “Content-adaptive inverse tone mapping,” in *2012 Visual Communications and Image Processing*, 2012, pp. 1–6.
- [15] X. Li and Y. Huo, “Hdr image reconstruction using locally weighted linear regression,” in *2018 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*, 2018, pp. 1–6.
- [16] A. G. Rempel, M. Trentacoste, H. Seetzen, H. D. Young, W. Heidrich, L. Whitehead, and G. Ward, “Ldr2hdr: On-the-fly reverse tone mapping of legacy video and photographs,” *ACM Trans. Graph.*, vol. 26, no. 3, p. 39–es, July 2007. [Online]. Available: <https://doi.org/10.1145/1276377.1276426>
- [17] S. Lee, G. H. An, and S.-J. Kang, “Deep recursive hdri: Inverse tone mapping using generative adversarial networks,” in *Proceedings of the European Conference on Computer Vision (ECCV)*, September 2018.
- [18] Y. Kinoshita and H. Kiya, “itm-net: Deep inverse tone mapping using novel loss function considering tone mapping operator,” *IEEE Access*, vol. 7, pp. 73 555–73 563, 2019.
- [19] Y.-L. Liu, W.-S. Lai, Y.-S. Chen, Y.-L. Kao, M.-H. Yang, Y.-Y. Chuang, and J.-B. Huang, “Single-image hdr reconstruction by learning to reverse the camera pipeline,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020.
- [20] J. Zhang and J.-F. Lalonde, “Learning high dynamic range from outdoor panoramas,” 2017.
- [21] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” 2015.
- [22] J. Xiao, K. A. Ehinger, A. Oliva, and A. Torralba, “Recognizing scene viewpoint using panoramic place representation,” in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 2695–2702.
- [23] T. Rhee, L. Petikam, B. Allen, and A. Chalmers, “Mr360: Mixed reality rendering for 360° panoramic videos,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 23, no. 4, pp. 1379–1388, 2017.
- [24] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, “Unpaired image-to-image translation using cycle-consistent adversarial networks,” 2020.
- [25] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” 2017.
- [26] S. Ning, H. Xu, L. Song, R. Xie, and W. Zhang, “Learning an inverse tone mapping network with a generative adversarial regularizer,” in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2018, pp. 1383–1387.

## 〈 저 자 소 개 〉



### 이 지 원

- 2012~2018 부산가톨릭대학교 소프트웨어학과 학사
- 2018~2021 한국과학기술원 문화기술대학원 석사
- 관심분야: 컴퓨터그래픽스, 컴퓨터 비전, VR/AR
- <https://orcid.org/0000-0001-6848-5110>



### 서 광 균

- 2016년 한국과학기술원 전기 및 전자공학부 학사
- 2018년 한국과학기술원 문화기술대학원 석사
- 2018년~현재 한국과학기술원 문화기술대학원 박사과정
- 관심분야: Image and Video Editing
- <https://orcid.org/0000-0003-0570-4915>



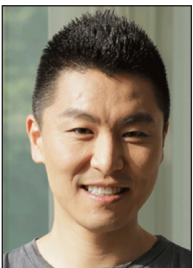
### 이 하 니

- 2009~2014 동국대학교 멀티미디어공학과 학사
- 2015~2017 한국과학기술원 문화기술대학원 석사
- 2017~현재 한국과학기술원 문화기술대학원 박사과정
- 관심분야 : 컴퓨터 그래픽스, 컴퓨터 비전
- <https://orcid.org/0000-0002-3612-3462>



### 유 정 은

- 2010~2015 한국과학기술원 전산학부 학사
- 2015~2018 한국과학기술원 문화기술대학원 석사
- 2018~현재 한국과학기술원 문화기술대학원 박사과정
- 관심분야: 컴퓨터 그래픽스, 컴퓨터 비전, VR/AR, 콘텐츠 저작 등
- <https://orcid.org/0000-0002-6276-452X>



### 노 준 용

- 2002년 University of Southern California Computer Science 박사
- 2003년~2006년 Rhythm and Hues Studio, Graphics Scientist
- 2006년~현재 카이스트 문화기술 대학원 교수
- 2011년~2014년 카이스트 석좌 교수
- 2016년~2020년 카이스트 문화기술대학원 학과장
- 관심분야: 컴퓨터 그래픽스, 컴퓨터 비전, 얼굴 애니메이션, 캐릭터 애니메이션, VR/AR, 몰입형 디스플레이
- <https://orcid.org/0000-0003-1925-3326>