

논문 2024-19-04

# 저가 카메라를 이용한 스마트 장난감 게임을 위한 모형 자동차 인식 (Recognition of Model Cars Using Low-Cost Camera in Smart Toy Games)

강민혜, 홍원기, 고재필\*  
(Minhye Kang, Won-Keel Hong, Jaepil Ko)

Abstract : Recently, there has been a growing interest in integrating physical toys into video gaming within the game content business. This paper introduces a novel method that leverages low-cost camera as an alternative to using sensor attachments to meet this rising demand. We address the limitations associated with low-cost cameras and propose an optical design tailored to the specific environment of model car recognition. We overcome the inherent limitations of low-cost cameras by proposing an optical design specifically tailored for model car recognition. This approach primarily focuses on recognizing the underside of the car and addresses the challenges associated with this particular perspective. Our method employs a transfer learning model that is specifically trained for this task. We have achieved a 100% recognition rate, highlighting the importance of collecting data under various camera exposures. This paper serves as a valuable case study for incorporating low-cost cameras into vision systems.

Keywords : Smart Toy, Interactive Video Game, Low-Cost Camera, Image Recognition, Deep Learning

## 1. 서론

최근 실물 장난감을 IT 기술과 접목하여 게임에 반영하는 콘텐츠 사업이 스마트 장난감이라 불리우며 완구시장의 새로운 아이템으로 떠오르고 있다. 현재 스마트 장난감 시장은 2025년까지 연평균 15.5%의 성장이 예측될 만큼 중요한 미래 산업으로 뽑힌다 [1]. 실물 장난감을 게임과 연동하기 위해서는 장난감을 센서로 인식하는 과정이 필요하다. 장난감에 별도의 센서를 탑재하는 방법은 장난감의 가격을 높이고 고장에도 취약하다는 문제점이 있다. 이에 따라 카메라를 이용한 컴퓨터비전 기술이 대안으로 떠오르고 있다. 한편, 대표적인 컴퓨터비전 기술 중 하나인 영상인식 기술도 딥러닝 기술의 발전에 힘입어 상용화 수준에 도달하였다. 2012년 발표된 AlexNet [2]을 시작으로 VGG [3], GoogleNet [4], ResNet [5], MobileNet [6], EfficientNet [7]와 같이 매년 더 우수한 성능의 분류 모델이 발표되고 있다. 최근에는 Transformer [8] 기반의 ViT (Vision Transformer) [9] 모델이 주목받고 있다. 대용량 자연 이미지인 이미지넷 (ImageNet) [10]으로 사전 학습이 완료된 이들 모델은 공개 모델이므로 이들을 기반으로 전이학습 [11]을 적용하여 주어진 영상인식 문제를 해결할 수 있다. 그러나, 실제 서비스 환경에서는 설계 제약 조건과 외부 조명과 같은 다양한 변수가 작용하기 때문에 주어진 문제에 대해

충분히 고려하여 훈련 데이터를 수집해야 한다.

영상인식은 카메라의 성능과 주변의 조명환경에 크게 좌우된다. 이에 따라, 일반적으로 비전 시스템은 고가의 비전 전용 카메라와 잘 설계된 조명기구가 필요하다. 그러나, 스마트 장난감에서는 장난감에 부착하는 센서 비용 이상의 카메라 도입이 어렵다. 이는 카메라를 이용한 스마트 장난감 확산을 저해하는 요인 중 하나이다. 따라서, 스마트 장난감 확산을 위해서는 저가의 카메라 사용과 조명의 영향을 최소화하는 구조설계가 가능해야 한다.

본 논문은 저가의 카메라를 이용해 비디오 게임 진행 중간에 플라스틱 모형 자동차를 인식하는 방법을 제안한다. 저가 카메라의 한계, 주어진 비디오 게임의 진행 시나리오, 모형 자동차의 특성, 조명환경을 종합적으로 고려하여 다음과 같이 제안한다. 첫째, 장난감 자동차의 바닥면을 인식하도록 촬영 기구를 설계한다. 둘째, 카메라의 노출을 달리하여 다양한 조명 조건의 영상을 획득한다. 셋째, 바닥면 촬영제약을 고려한 데이터 증강을 수행한다. 마지막으로 이러한 데이터 다양성을 효과적으로 학습하기 위해 전이학습 모델을 사용한다. 주어진 문제에 대한 전이학습의 성공 여부는 적절한 훈련 데이터의 수집에 달려있다. 따라서, 저가 카메라로 인해 모형 자동차의 밝기 및 색상표현이 제대로 이루어지지 못하는 문제와 외부 조명 문제를 회피하기 위해 마련한 촬영방법이 안고 있는 모형 자동차의 일부 가림 문제를 모두 고려하여 데이터를 수집한다. 또한, 전처리로서의 히스토그램 평활화의 효과를 카메라 노출과 밝기증강 영상에 대해 PSNR (Peak Signal-to-noise ratio) [12] 관점에서 분석한다.

본 논문의 구성은 다음과 같다. 2장에서는 게임 진행 중

\*Corresponding Author (nonezero@kumoh.ac.kr)

Received: Sep. 6, 2023, Revised: Oct. 23, 2023, Accepted: Jan. 19, 2024.

M. H. Kang: Kumoh National Institute of Technology (M.S.)

W. K. Hong: Daegu University (Prof.)

J. P. Ko: Kumoh National Institute of Technology (Prof.)

\* 이 연구는 금오공과대학교 대학 학술연구비로 지원되었음 (2021).

어느 부분에서 비전 인식이 필요한지, 어떠한 제약 사항이 있는지 등의 주어진 비전 시스템이 가지고 있는 설계 제약 사항을 검토한다. 3장에서는 전이학습 모델을 위한 전처리 방법과 모델의 구조를 설명한다. 4장에서는 데이터 증강 방법과 전이학습 모델을 위한 훈련과 이에 따른 성능을 보인다. 특히, 전이학습 모델의 성능에 중요한 데이터 증강 방법과 관련하여, 저가 카메라의 노출 문제와 모형 자동차 거치에 관한 문제를 논의한다. 또한, 카메라 노출설정 변화와 밝기증강에 대해 분석한다. 마지막 5장에서는 결론을 내린다.

## II. 모형 자동차 인식 환경

### 1. 게임 진행 시나리오

개발한 카메라를 이용하는 게임기에서 영상인식과 관계된 경우는 다음과 같다. 게임이 시작되면 게임 사용자는 게임 화면의 지시에 따라 그림 1에서 빨간색 점선으로 표시된 원 안에 자신이 가진 모형 자동차를 올려놓는다. 그러면, 게임기에 설치된 카메라를 이용해 모형 자동차를 촬영하고 자동차의 종류를 인식한 후 인식한 모형 자동차에 해당하는 자동차 캐릭터 이미지를 게임화면에 표시한다. 계속해서 사용자는 게임화면의 지시에 따라 모형 자동차를 화살표 방향으로 빠르게 밀어서 이동시킨다. 이때 게임기는 모형 자동차가 놓여있던 위치를 벗어났는지를 영상인식을 통해 판정하여 게임을 진행한다.

한편, 이 게임기는 5~8세 아동을 대상으로 한다. 어린 아동의 특성상 지정된 위치에 장난감을 제대로 올려놓지 못하는 경우가 발생할 수 있다.

### 2. 카메라 설치 위치 및 모형 자동차 거치대

그림 1에서 보듯이, 게임기의 테두리 부분에는 사용자의 관심을 끌고 게임 진행 화면과 연동하여 색상이 변하는 조명이 부착되어 있다. 이 역동적인 조명에 그대로 노출된 모형 자동차를 촬영하게 되면 촬영 시점의 조명에 따라 촬영된 영상에서 모형 자동차의 밝기와 색상은 변하게 된다. 우리는 이 문제를 해결하기 위해 주변 조명의 영향을 받지 않는 부분인 모형 자동차의 바닥면을 촬영한다. 바닥면은 대체로 평면이므로 촬영 각도의 변화에 덜 민감하다는 장점이 있다. 바닥면 촬영을 위해, 모형 자동차를 올려두는 위치 아래, 게임기 안쪽에 그림 2와 같이 카메라를 설치하고 모형 자동차를 올려두는 거치대는 투명창을 둔다. 이때, 거치대 아래쪽에 조명을 설치하여 모형 자동차 바닥면을 비출 수 있도록 하고 주위는 암막으로 마감하여 외부 빛이 들어오지 못하도록 한다.

### 3. 모형 자동차 바닥 인식의 문제점

그림 3은 일부 모형 자동차의 바닥을 보여준다. 모형 자동차는 변신 모형의 구조적 특성으로 인해 모형의 바닥은 패턴과 색상이 모두 달라 시각적 구분이 가능하다.

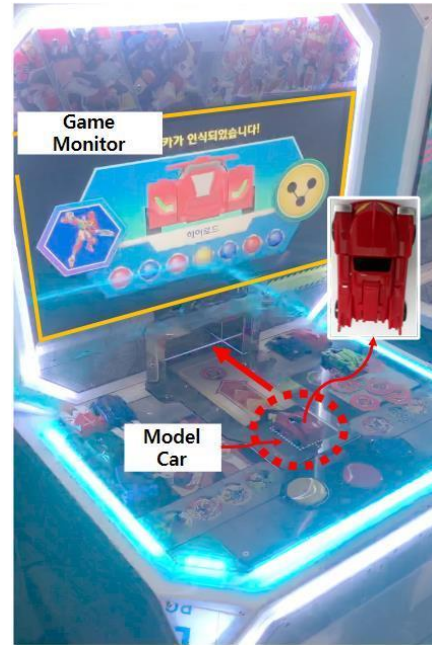


그림 1. 제안 비전 시스템이 탑재된 스마트 장난감 시제품  
Fig. 1. Smart toy prototype with the proposed vision system

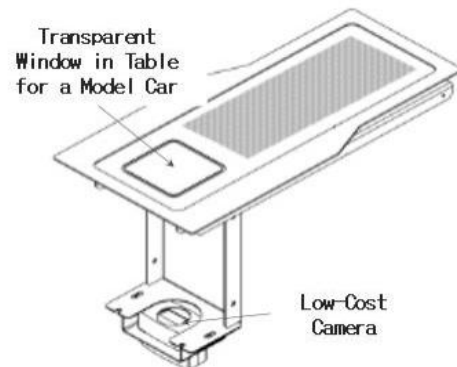


그림 2. 모형 자동차 바닥 촬영을 위한 카메라 위치 및 모형 자동차 거치 투명창  
Fig. 2. The transparent window on the table used to capture the underside of the model car



그림 3. 다양한 패턴과 색상을 가진 모형 자동차 바닥  
Fig. 3. The underside of model cars with various patterns and colors

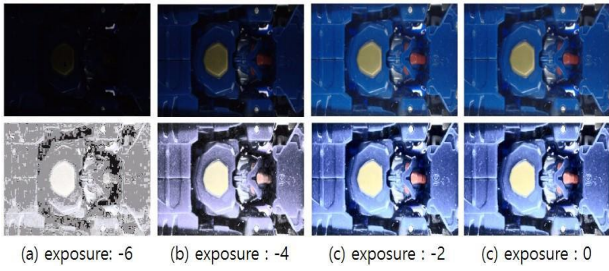


그림 4. 카메라 노출값에 따른 촬영영상 (상) 및 히스토그램 평활화 적용 영상 (하)

Fig. 4. Images captured with varying camera exposure values (top) and images with applied histogram equalization (bottom)

그러나, 실제 게임기 거치대에서는 자동차의 바닥면 전체를 촬영하지 못하도록 설계하였다. 왜냐하면, 모형 자동차 바닥면 전체가 촬영되도록 투명 거치대를 크게 제작하면 게임기가 놓인 천정에 설치된 조명이 새어들어 카메라에 직접 도달할 수 있기 때문이다. 이를 막기 위해 거치대의 투명창은 모형 자동차의 바닥 크기보다 조금 작게 한다. 그러면, 모형 자동차를 올려놓았을 때 투명창 전체가 가려질 수 있다. 대신, 이렇게 하면 모형 자동차의 일부만 촬영되는 문제가 발생하므로 인식 단계에서 이를 고려해야 한다.

한편, 어린 아동이 사용한다는 점을 고려하여 모형 자동차 거치대 뒤쪽에 지지대를 두어 길이 방향으로는 모형 자동차를 고정된 위치에 놓이도록 하였다. 그러나 거치대의 좌우에는 지지대를 둘 수 없어 모형 자동차가 좌우 방향으로 어느 정도 벗어날 수 있다. 즉, 인식대상인 모형 장난감이 촬영된 영상의 중심에 오지 않거나, 일부만 포함되는 현상이 발생할 수 있다.

### III. 모형 자동차 바닥면 인식 방법

#### 1. 영상 전처리 방법

본 논문에서 모형 자동차 인식 모델은 VGG16 [3] 기반의 전이학습 모델을 사용한다. 입력영상은 모델에 입력하기 전에 밝기 정규화 단계를 거친다. 우리는 VGG16 모델에서 훈련단계에서 사용된 정규화 대신 각 픽셀을 영상의 최대 밝기인 255로 나누어 주는 간단한 방법을 사용한다. 왜냐하면 VGG16 모델이 훈련에 사용한 이미지넷의 자연물 영상은 색상이 선명한 인공물인 모형 자동차와는 색상 특성이 다르므로 보다 범용적인 정규화 방법을 적용할 필요가 있기 때문이다.

비교적 간단한 분류 모델임에도 불구하고 전이학습 모델을 선택한 이유는 다음과 같다. 일반적으로 영상의 밝기 보정을 위해 히스토그램 평활화를 적용하면 영상의 밝기 변화를 효과적으로 줄일 수 있다. 그러나, 그림 4에서 보듯이 카메라 노출이 낮은 경우 종종 색상 정보가 사라지는데, 이는 카메라 노출이 낮은 경우 모형 자동차를 색상만으로 구분할 수 없음을 의미한다. 따라서, 자동차 바닥면에 나타나는 모양 정보를 학습할 수 있도록 깊은 층을 가진 신경망을 사용

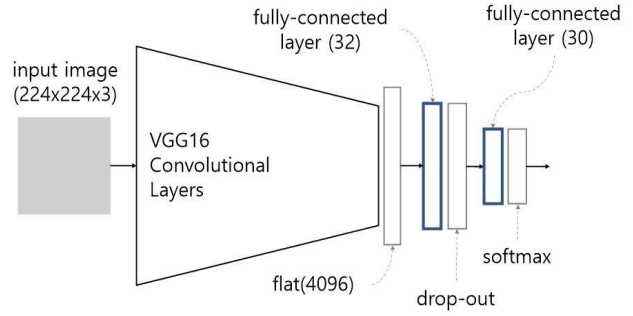


그림 5. VGG16기반 전이학습 모델 구조

Fig. 5. The architecture of transfer learning model based on VGG16

할 필요가 있다. 이러한 이유로 본 논문에서는 VGG16을 선택하였다.

#### 2. 모형 자동차 인식을 위한 전이학습 모델 구조

전이학습 모델은 크게 특징 추출부와 분류부로 나뉜다. 일반적으로 특징 추출부는 기존 모델의 것을 그대로 사용하고 분류부는 주어진 분류문제에 따라 교체한다. 훈련 방식에 따라 다시 두 가지로 나뉜다. 첫째, 분류기 부분만 훈련하는 방법과 둘째, 분류기와 특징 추출부를 모두 훈련하는 방법이다. 이때, 특징 추출부 훈련은 전체 또는 일부를 훈련하느냐에 따라 세분화할 수 있다. 본 논문에서는 분류기 부분만 훈련하는 방식을 택한다.

게임기가 구분해야 할 모형 자동차의 수는 29종이나, 다음과 같은 이유로 하나의 클래스를 추가하여 분류기는 총 30개의 클래스를 구분하도록 설계한다. 게임 진행자가 게임 화면의 지시에 따라 자신이 가지고 있는 모형 자동차 중 하나를 올려놓으면, 게임기는 모형 자동차 인식을 시도하게 된다. 게임 진행자가 언제 모형 자동차를 올릴지 알 수 없기에 이러한 시도를 등록된 모형 자동차가 인식될 때까지 반복해야 한다. 따라서, 모형 자동차가 올려져 있지 않은 상황을 별도로 구분하여 올려지지 않은 상황을 쉽게 판단할 수 있도록 해야 한다. 한편, 게임진행 중에 게임기 테두리에 설치된 조명색이 초록색, 노란색, 붉은색 등으로 수시로 바뀌므로 차가 없는 상황을 별도의 클래스로 추가한다.

제안하는 전이학습 모델은 분류할 클래스의 수가 많지 않고 인식대상이 비교적 정형화된 패턴을 가지기 때문에 합성곱층으로 구성된 특징 추출부 뒤에 2개의 완전연결층을 이어 간단하게 구성하였다. 완전연결층에는 과적합을 완화하기 위해 드롭아웃 [13]을 적용한다. 전체 모델의 구조는 그림 5와 같다.

특징의 차원은 4,096이고 분류를 위한 첫 번째 완전연결층은 여러 번의 실험을 통해 노드의 수를 32개로 설정하였다.

### IV. 실험

#### 1. 저가 카메라의 노출 문제를 고려한 데이터 증강

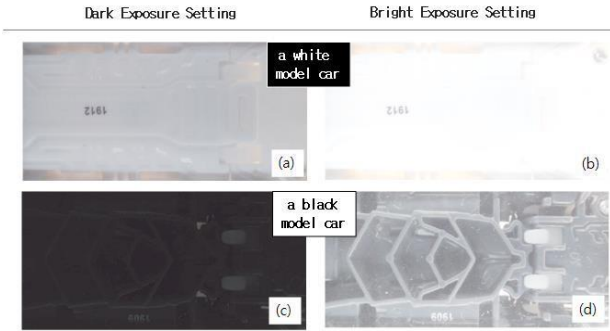


그림 6. 노출에 따른 흰색과 검은색 모형 자동차 촬영 영상  
 Fig. 6. Images of white and black model cars captured for each dark and bright exposure

저가의 UVC (USB Video Class) 카메라는 일반적으로 저해상도, 제한된 동적범위, 저속촬영의 한계를 갖고 있다. 동적범위가 좁을 경우, 고정된 카메라 노출에서는 물체의 어두운 부분과 밝은 부분을 동시에 표현하기 어렵다. 즉, 물체의 밝은 부분에 맞추어 노출을 설정하면 어두운 부분은 검게 촬영되고, 반대로 경우는 하얗게 촬영된다. 본 모형 자동차 인식에서는 이러한 제한된 동적범위가 문제가 된다. 즉, 카메라 노출을 고정하면 장난감 자동차의 색상에 따라 노출 과다 또는 노출 부족 문제가 발생한다.

그림 6은 카메라 노출을 달리하여 검은색 모형 자동차와 흰색 모형 자동차를 촬영한 영상을 보여준다. 어두운 카메라 노출에서 (a) 흰색 모형 자동차의 패턴이 명확하게 표현되고 있다. 그러나, (c) 검은색 모형 자동차의 패턴은 저노출로 인해 패턴이 표현되지 않는다. 반대로 밝은 카메라 노출에서는 (b) 흰색 모형 자동차의 패턴은 과노출로 인해 패턴이 표현되지 않지만 (d) 검은색 모형 자동차는 적정 노출로 패턴이 명확하게 표현되고 있다. 즉, 단일 고정 카메라 노출로는 UVC 카메라의 한계로 인해 검은색과 흰색 모형 자동차에 대한 적정 영상을 획득할 수 없다.

이러한 문제점을 해결하기 위하여 동적 범위 재구성 알고리즘 [14], 이를 이용한 카메라 동적 범위 영역 확장 [15] 등이 연구되었다. 그러나 이들은 스마트폰을 대상으로 연구되어, 본 논문에서 사용하는 UVC 카메라는 적용하기 어렵다.

본 논문에서는 노출 문제를 극복하기 위해, 검은색 모형 자동차와 흰색 모형 자동차가 각각 명확하게 표현되는 카메라 노출 구간을 정한 후, 노출을 바꾸어가며 딥러닝 모델 훈련을 위한 모형 자동차의 바닥면 영상을 수집하였다. 이때, 노출은 주어진 카메라의 노출 조정 범위인 0~-6까지 7단계로 하였으며, 이 중 훈련용은 짝수인 0, -2, -4, -6으로 하고, 테스트용은 홀수인 -1, -3, -5로 구분하였다. 영상의 밝기와 관련 있는 게임은 175로 동일하게 설정하였다.

2. 모형 자동차 거치대 문제를 고려한 데이터 증강

앞서 2.3절에서 설명하였듯이 모형 자동차 투명창은 자동차 크기보다 작게 제작하였다. 왜냐하면, 투명 거치대를 크게 제작하면 그림 7과 같이 틈새가 생겨 천정에 설치된 조

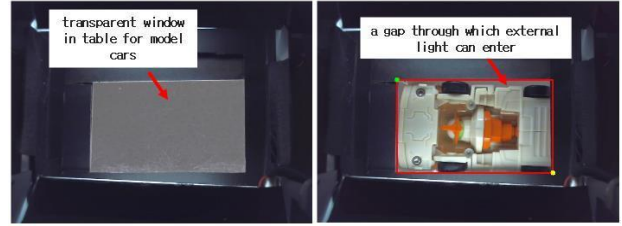


그림 7. 거치대 투명창을 올려 촬영한 영상  
 Fig. 7. Images captured upward through the transparent window

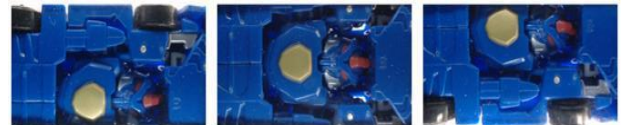


그림 8. 작은 투명창에서 촬영된 영상  
 Fig. 8. Images captured upward through the transparent narrowed window

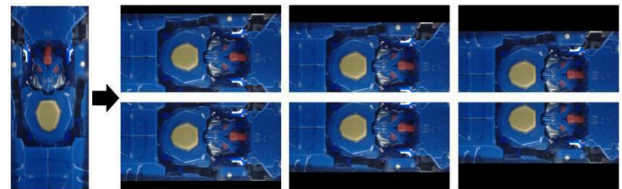


그림 9. 전이학습 모델 훈련용 데이터 픽셀 이동 증강 영상  
 Fig. 9. Images with pixel shift augmentation for training the transfer learning model

명이 새어들어 카메라에 직접 도달할 수 있기 때문이다.

투명창의 크기는 자동차의 한쪽 바퀴가 모두 보이면 다른 쪽 바퀴는 보이지 않는 수준으로 하였다. 게임 사용자가 자연스럽게 모형 자동차를 투명 거치대 위에 올려놓는 것을 가정했을 때, 그림 8과 같이 왼쪽 바퀴가 보이는 경우, 두 바퀴가 모두 보이지 않는 경우, 오른쪽 바퀴가 보이는 경우가 발생할 수 있다.

촬영된 영상을 보면 같은 모형 자동차임에도 불구하고 이들 간의 차이가 크게 나는 것을 볼 수 있다. 실제 게임 진행 상황에서는 그림 8과 같이 촬영될 수 있으므로 이를 고려하여 데이터 증강을 적용해야 한다. 노출을 7단계로 바꾸어가며 동시에 차량 위치를 다양하게 이동시키면서 촬영하는 것은 번거롭다. 우리는 훈련 데이터 수집의 용이성을 위해 그림 9와 같이 두 바퀴가 보이지 않는 영상을 촬영한 후 영상 처리를 통해 상하로 이동시키고 이동에 따라 생긴 영역은 검은색으로 채워 넣는 방식으로 딥러닝 모델 훈련용 데이터를 증강하였다. 이동 픽셀의 수는 모형 자동차 영상의 높이를 기준 삼아 0.1배, 0.2배, 0.3배에 해당하는 값으로 하였다.

훈련 영상은 7단계의 노출과 픽셀 이동 증강을 적용하여 총 844장을 수집하였다. 여기서, 모형 자동차에 대해서는 노출 4단계, 픽셀 이동 7단계, 모형 자동차 29종에 대해 812장이고, 모형 자동차를 올려놓지 않지 않은 경우는 노출 4단



그림 10. 카메라 노출 차이에 따른 PSNR 비교  
Fig. 10. Comparison of PSNR at various camera exposures



그림 11. 밝기증강에 따른 PSNR 비교  
Fig. 11. Comparison of PSNR with varying brightness adjustments

계, 8가지 상황에 대해 32장이다. 테스트 영상은 285장을 수집하였다. 여기서, 모형 자동차에 대해서는 노출 3단계, 그림 8과 같이 위치변화 3단계, 모형 자동차 29종에 대해 261장이고, 모형 자동차를 올려놓지 않은 경우는 노출 3단계, 8가지 상황에 대해 24장이다.

### 3. 전이학습 모델 훈련 및 성능

전이학습 모델을 훈련하기 위한 하이퍼 파라미터는 다음과 같다. 여러 차례의 실험을 거쳐 에폭은 100, 배치 크기는 8, 최적화 기법은 Adam [16], 학습률은 0.0005, 드랍아웃의 유지 비율은 0.7로 설정하였다. 최적의 분류를 위한 첫 번째 전연결층의 노드 수는 8개에서 96개까지 바꾸어가며 테스트 성능을 측정하였고, 실험결과 32개에서 100% 인식률을 달성하였다.

### 4. 카메라 노출변화 및 밝기증강 효과 분석

물체 분류를 위한 신경망 학습 시 밝기 증강을 적용하는 것이 일반적이다. 그러나, 본 논문에서는 카메라 노출 조절을 통해 다양한 밝기의 영상을 확보하였고 추가로 히스토그램 평활화를 적용하므로 밝기 증강은 적용하지 않았다. 카메라 노출 조절과 밝기 증강의 효과를 수치적으로 확인하기 위해 수집한 모든 영상들 간의 PSNR을 사용하였다. PSNR은 영상의 화질손실양을 평가하는 지표로 널리 사용되고 있다. 우리는 노출 바뀔과 밝기 증강에 따라 화질이 어느정도 변하는지를 수치화하는데 사용하였다.

그림 10은 카메라 노출 설정에 따른 화질 차이를 비교하고, 그림 11은 주어진 영상의 밝기를 10%에서 80%까지 증가하였을 때 원본과의 화질 차이를 비교한 그래프이다. 히스토그램 평활화를 적용한 경우 (W HE)와 그렇지 않은 경우 (W/O HE)를 함께 제시하였다. 그래프에서 PSNR이 클수록 비교 영상 간의 화질 차이가 작다는 것을 의미한다.

첫째, 히스토그램 평활화를 적용하지 않은 경우는 다음과 같이 해석할 수 있다. 노출변화에 대한 그림 10에서 영상간 노출 차이가 2일 때 30db 수준을 보이고, 노출 차이가 4, 6일 때 28db 이하로 낮아진다. 즉, 노출 차이에 따라 화질이

변한다는 것을 나타낸다. 밝기증강에 대한 그림 11에서는 전체적으로 28db 수준에 머물러 있다. 즉, 밝기증강에 따라 화질 변화가 거의 없음을 나타낸다.

둘째, 히스토그램 평활화를 적용한 경우는 다음과 같이 해석할 수 있다. 노출변화 및 밝기증강 두 경우 모두에서 히스토그램 평활화를 적용하지 않은 경우보다 상대적으로 높은 db를 유지하고 있다. 즉, 히스토그램 평활화는 두 경우 모두 영상의 화질을 일정하게 유지하는데 도움이 된다는 것을 의미한다. 한편, 노출변화에 대한 그림 10에서는 노출차이에 따라 29db 이하로 내려 가는데 반해 밝기증강에 대한 그림 11에서는 밝기증강에 따라 대체로 29db 이상을 유지하고 있다. 즉, 노출증강에서는 히스토그램 평활화의 효과가 상대적으로 크다고 볼 수 있다.

## V. 결론

본 논문에서는 스마트 장난감 제품, 구체적으로 비디오 게임 컨트롤러를 위해 저가의 UVC 카메라를 이용한 모형 자동차 인식 비전 시스템을 제안하였다. 장난감 자체에 센서를 부착하기 어려운 경우, 비전 인식은 훌륭한 대안이지만 먼저, 고가의 비전 전용 카메라 대신 저가의 UVC 카메라를 사용할 수 있는 기술을 확보해야 한다.

본 논문에서는 저가의 UVC 카메라를 사용할 수 있도록 광학설계와 전이학습 기반 모델을 제안하였다. 광학부 설계에서는 게임기가 사용되는 조명환경을 고려하여 자동차의 바닥면을 인식할 수 있도록 하였다. 인식 대상인 모형 자동차는 검은색에서 흰색까지 다양한 색상으로 구성되어 저가 카메라의 동적 범위를 넘어섰고, 더 나아가 외부 조명을 피하기 위한 기구 설계상 인식 대상의 일부만 촬영되는 문제도 있었다. 이러한 문제를 극복하기 위한 데이터 취득 방법과 데이터 증강 방법을 제시하였고, 이를 통해 훈련한 전이학습 모델은 우리가 수집한 데이터에 대해 100% 인식률을 달성하였다. 또한, PSNR 분석을 통해 카메라 노출을 달리하여 훈련 영상을 수집하는 것이 필요함을 보였다.

본 논문은 비전 기반 스마트 장난감 제품 확산을 위해 풀어야 할 문제인 저가 카메라 도입 방법을 보여주었으며, 전이학습의 실제 적용 과정을 살펴볼 수 있는 하나의 사례연구로서 의미가 있다.

## References

- [1] 김한결, K-스마트 토이 '루미' 중동시장 진출, 환경미디어, 2020.
- [2] A. Krizhevsky, I. Sutskever, G. Hinton, "Imagenet Classification with Deep Convolutional Neural Networks," *Advances in Neural Information Processing Systems*, Vol. 25, 2012.
- [3] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-scale Image Recognition," 3rd International Conference on Learning Representations, 2015.
- [4] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, "Going Deeper with Convolutions," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-9, 2015.
- [5] K. He, X. Zhang, S. Ren, J. Sun, "Deep Residual Learning for Image Recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016.
- [6] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam, "Mobilenets: Efficient Convolutional Neural Networks for Mobile Vision Applications," *CoRR*, Vol. abs/1704.04861, 2017.
- [7] M. Tan, Q. Le, "Efficientnet: Rethinking Model Scaling for Convolutional Neural Networks," *International Conference on Machine Learning*, pp. 6105-6114, 2019.
- [8] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, I. Polosukhin, "Attention is All you Need," *Advances in Neural Information Processing Systems*, Vol. 30, 2017.
- [9] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," *International Conference on Learning Representations*, 2021.
- [10] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, "Imagenet: A Large-scale Hierarchical Image Database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, pp. 248-255, 2009.
- [11] S. J. Pan, Q. Yang, "A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, Vol. 22, No. 10, pp. 1345-1359, 2009.
- [12] A. Horé, D. Ziou, "Image Quality Metrics: PSNR vs. SSIM," 2010 20th International Conference on Pattern Recognition, pp. 2366-2369, 2010.
- [13] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, R. Salakhutdinov, "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," *JMLR*, pp. 1929-1958, 2014.
- [14] D. K. Choi, C. H. Han, "Illuminance Dynamic Range Expansion using Gamma & Multi-Point Knee for Smart Phone Camera," *IEMEK J. Embed. Sys. Appl.*, Vol. 8, No. 1, pp. 43-50, 2013 (in Korean).
- [15] S. W. Park, K. R. Cha, "Dynamic Range Reconstruction Algorithm for Smart Phone Camera Pulse Measurement

*Robust to Light Condition,*" *IEMEK J. Embed. Sys. Appl.*, Vol. 10, No. 1, pp. 1-6, 2015 (in Korean).

- [16] D. P. Kingma, J. Ba, "Adam: A Method for Stochastic Optimization," *ICLR*, pp. 1-15, 2015.

### Minhye Kang (강민혜)



2022 Computer Engineering from Kumoh National Institute of Technology (B.S.)

2024 Computer-AI Convergence Engineering from Kumoh National Institute of Technology (M.S. Student)

Field of Interests: Computer Vision, Machine Learning  
Email: mh9907@kumoh.ac.kr

### Won-Kee Hong (홍원기)



1997 Computer Science from Yonsei University (M.S.)

2001 Computer Science from Yonsei University (Ph.D.)

2004~Department of Information Security from Daegu University (Prof.)

Career:

2002~2004 LG LM Research (Senior Research Engineer)

Field of Interests: Embedded System, Information Security

Email: wkhong@daegu.ac.kr

### Jaepil Ko (고재필)



1998 Computer Science from Yonsei University (M.S.)

2004 Computer Science from Yonsei University (Ph.D.)

2004~Computer Engineering from Kumoh National Institute of Technology (Prof.)

Field of Interests: Computer Vision, Machine Learning

Email: nonezero@kumoh.ac.kr