

# 딥러닝을 활용한 영상 속 차량 검출 시스템 구현

김효석, 이주현  
한양대학교

kkim7549@gmail.com, joohyunlee@hanyang.ac.kr

## Vehicle detection system by using deep learning

Kim Hyoseok, Joohyun Lee  
Hanyang Univ.

### 요 약

본 논문에서는 블랙박스나 CCTV 등 불특정된 환경에서 촬영된 영상속의 차량을 검출하는 시스템을 구현하기 위해 어떠한 방법이 효과적인지 실험 및 비교하고 시스템을 직접 구현해 보았다.

### I. 서 론

최근 딥러닝 기술의 발달로 영상 속의 객체를 빠르게 파악하고 인식하는 연구가 활발히 이루어지고 있다. 이 기술을 활용해 영상 내의 번호판을 탐지, OCR 기술을 접목하여 차량의 번호를 알아내는 방식은 이미 널리 사용되고 있다. 하지만 뺑소니 등의 경찰 수사와 같은 상황에서 차량의 외견 특징만 알 뿐 번호를 특정할 수 없거나, CCTV 나 블랙박스 등 깨끗하지 못한 환경 하에서 촬영된 영상에서 특정 차량을 검출하기 위해서는 여전히 사람이 수동으로 확인할 필요가 있다. 이처럼 OCR 기술만으로는 부족한 상황이 종종 발생하고, 차량을 식별하는 데 추가적인 정보가 요구되는 경우 이미지 내의 차량의 외견적 특징인 모델이나 도색을 스스로 감지하는 시스템이 필요하다.

딥러닝 모델을 사용하기 위해서는 일반적으로 다량의 질 좋은 데이터가 필요함이 일반적으로 알려져 있다. 하지만 현실적으로 보았을 때 시간과 비용 등 제약으로 인해 모든 차량의 데이터를 확보하기에 어려움이 있다.

따라서 본 논문에서는 적은 양의 데이터를 사용하면서도, 딥러닝 기술을 적용하여 영상 속 차량의 검출과 차량의 모델과 색상을 분류하는 시스템을 구현해보고자 한다.

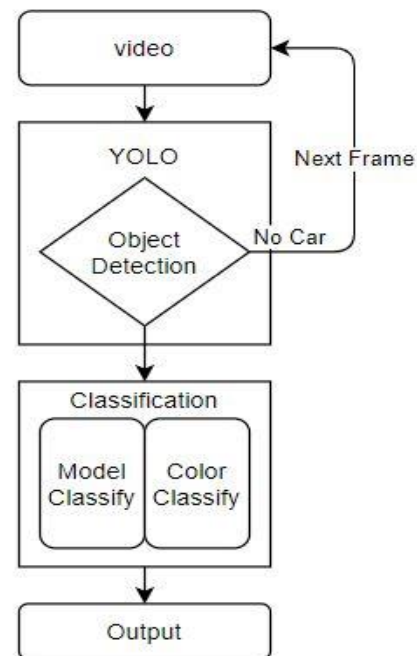
이를 위해서 CNN 모델 중에서도 YOLOv4 모델과 ResNet34 모델을 사용했으며, 적은 이미지 데이터로 원하는 성능을 얻기 위해 전이학습을 활용했다. 전이학습의 효율성을 확인하기 위해, 전이학습을 사용하지 않는 경우 또는 다른 방법과 성능을 비교해 보았다.

### II. 본론

#### 1. 시스템 개요

영상 속의 차량을 검출하기 위해서 YOLOv4 모델을 사용하며, 차량의 모델과 도색을 구분하는 분류기에서는 전이학습을 활용하여 조정한 ResNet34 기반 모델을

사용한다. 마지막으로 YOLO 모델을 한번 더 통과하여 영상에서 차량의 번호판이 찍혔는지 확인한다.



#### 2. 데이터 셋

차량과 번호판을 검출해 줄 YOLO 는 COCO 데이터셋으로 훈련된 가중치를 사용했으며, 분류기의 학습을 위해서는 국내 두 기업에서 생산한 28 개 종류의 약 2000 개 차량모델 이미지 데이터를 사용했다. 그리고 차량의 도색을 분류하기 위해서 10 가지 색상별로 분류된 차량의 이미지 데이터 셋을 사용했는데 2000 개의 이미지로 구성되어 있다.

#### 3. 차량 도색 분류기

차량의 몸체에 반사된 조명의 영향을 고려하여 두 가지 방식을 고려하였는데, 하나는 K-means clustering 을 이용해 이미지에 분포하는 색을 양자화하는 과정을 거쳐 처리된 이미지의 픽셀의 색상에 따라 개수를 계산하는 고전적인 방법, 그리고 이미지를 전처리하는 과정을 거치지 않고 CNN 모델 중 ResNet 을 사용하여 딥러닝 하는 방식이다. 이 중 CNN 모델을 사용하는 경우는 다시 두 경우로 나누어, pretrain 방식을 사용하는 경우와 사용하지 않는 경우를 비교해 실험하였다.

결과적으로 전이학습된 모델이 가장 우수한 성능을 보였다. 처음 실험에서는 400 개의 이미지로 구성된 적은 데이터 세트를 사용했을 때 과적합의 경향을 보였으나 이미지를 2000 개로 늘린 데이터세트를 사용하였더니 과적합이 해소되고 정확도가 비약적으로 상승함을 확인했다.

표 1 에 평가지표로써 Accuracy 와 F1 score 을 사용하여 세 방법을 비교한 결과를 표기하였다.

표 1. 도색 분류기의 각 방법 간 성능 비교

	Accuracy	F1 score	Loss
K-means Clustering	69.2%(avg)	X	X
Non-Pretrained	88%	0.85	0.43
Pretrained	94%	0.92	0.03

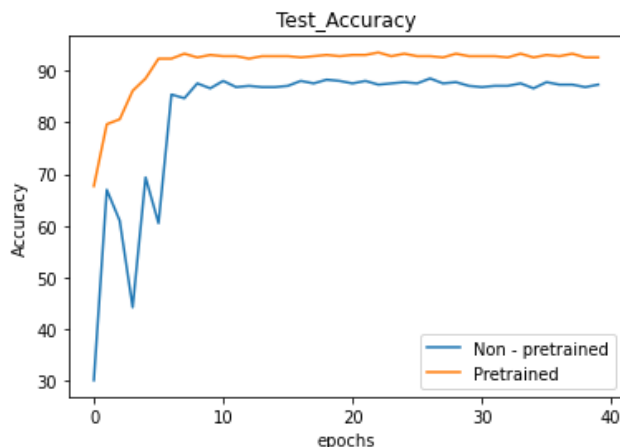


그림 2. 전이학습 모델의 훈련 결과 비교

#### 4. 차량 모델 분류기

도색 분류기와 마찬가지로, 전이학습된 ResNet 34 모델을 고정추출방식을 통해 훈련하였으며, 최종적으로 94%의 검증 정확도를 보였다. 마찬가지로 전이학습 방식을 통해 큰 향상을 보였다.

표 2. 모델 분류기의 두 방식 간 성능 비교

	Accuracy	F1 score	Loss
Non-pretrained	48%	0.365	1.48
Pretrained	94%	0.922	0.01



그림3. (좌) 모델 분류기의 Grad-Cam 히트맵. (우) 도색 분류기의 Grad-Cam 히트맵.

전이학습으로 훈련시킨 모델이 차량의 어느 부분을 특징으로 삼는지 확인하기 위해 Grad-Cam 히트맵을 표시해보면, 적은 데이터셋을 사용했음에도 불구하고 차량의 적절한 부분을 특징으로 선택하여 이미지를 분류할 수 있도록 학습되었음을 확인할 수 있다.

#### 5. 시스템의 실제 적용 결과

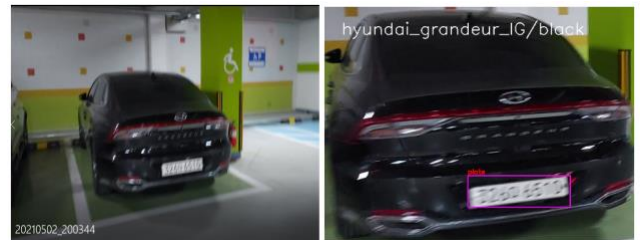


그림 4. 사용된 임의의 영상(좌)과 결과(우)

그림 4 는 실제 임의로 촬영된 영상을 시스템에 통과시켰을 시 얻을 수 있는 결과를 나타낸 그림이다. 훈련에서 사용된 이미지보다 깨끗하지 않고, 어느정도 흔들림이 있거나 절반만 촬영된 차량에 대해서도 높은 정확도를 가지고 검출해 낼 수 있음을 확인하였다.

### III. 결론

차량의 모델과 도색만을 알고 있을 때 임의의 환경에서 촬영된 영상으로부터 효과적으로 해당 특징을 갖는 차량을 분리할 수 있는 시스템을 구현하였다. 또한 전이학습을 사용하여 적은 데이터셋을 가지고도 효과적으로 훈련할 수 있음을 확인하였다.

이 방식을 응용하여 쉽게 차량의 추가적인 모델, 색상을 포함하여 분류의 범위를 넓히고 향후 OCR 기술을 접목하여 개선할 수 있을 것으로 기대된다.

### 참 고 문 헌

- [1] How many images do I need? Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring (Saleh Shahinfar, Paul Meek, Greg Falzon)
- [2] Pedestrian Classification using CNN's Deep Features and Transfer Learning (Chung Soyoung, Chung Mingyo)