정보통신공학과 졸업 논문

CNN을 이용한 차종 판별 시스템 구현

Implementation of Car Classification System using Convolutional Neural Network

한국외국어대학교

정보통신공학과

201402750

임광효

지도 교수 : 한 희 일

목 차

[1. 서론 4](#_Toc12357733)

[2. 관련 기술 조사 6](#_Toc12357734)

[2.1 CNN 6](#_Toc12357735)

[3. 본론 7](#_Toc12357736)

[3.1 차종 판별 시스템 7](#_Toc12357737)

[3.1.1 시스템 구성도 7](#_Toc12357738)

[3.1.2 데이터 구성에 따른 학습 결과 8](#_Toc12357739)

[3.1.2.1 1차 성능 개선 10](#_Toc12357740)

[3.1.2.2 중간 결과 13](#_Toc12357741)

[3.1.2.3 2차 성능 개선 14](#_Toc12357742)

[3.1.2.4 2차 결과 16](#_Toc12357743)

[3.1.3 모델에 따른 학습 결과 18](#_Toc12357744)

[4.결론 및 향후 연구 과제 19](#_Toc12357745)

[참고문헌 20](#_Toc12357746)

그 림 목 차

[그림 1-1] 연도별 공공기관 CCTV 설치 현황………………………………………………………...4

[그림 1-2] 수동 영상관제 시간에 따른 위험 탐지율………………………………………………4

[그림 2-1] CNN의 부분연결 구조………………………………………………………………………….6

[그림 2-2] CNN 구조……………………………………………………………………………………………..6

[그림3-1] 시스템 구성도……………………………………………………………….……………………..7

[그림3-2] Stanford Cars Dataset…………………………………………………………………..……….8

[그림 3-3] Loss rate…………………………………………………………………………………….………..8

[그림 3-4] 학습 결과……………………………………………………………………………….……………8

[그림 3-5] 변형된 데이터 셋…………………………………………………………………….…………...9

[그림 3-6] 변형된 데이터셋 학습 결과………………………………………………………….……….9

[그림 3-7] 차종 판별 예시…………………………………………………………………….……………..10

[그림 3-8] 차종 판별 오답…………………………………………………………………………………..10

[그림 3-9] 오답 차종 이미지………………………………………………………………….……………11

[그림3-10] 차종 판별 중 특이한 오답………………………………………………….…………...…12

[그림 3-11] 불량 데이터 교체…………………………………………………………….……………….14

[그림 3-12] 기존 모델과 개선된 모델과의 성능 비교…………………………………………..15

[그림 3-13] 가장 낮은 출력을 보인 차종의 비교…………………………….…………………….16

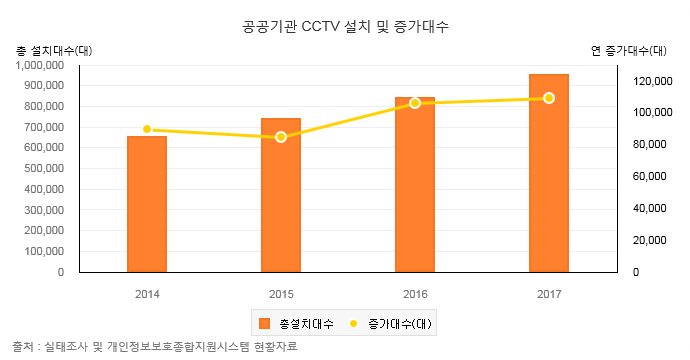
[그림 3-14] 정답이Top-5 이상의 순위로 출력된 오답 이미지……………………………..17

[그림 3-15] 모델 별 성능 비교………………………………………………………………….…………18

요 약

# 서론

오늘날 각종 범죄 예방 및 교통 정보 수집 등의 이유로 CCTV 설치 대수는 매년 증가하는 추세이나, 한정된 인력이 매년 증가하는 CCTV를 효율적으로 관리하고 모니터링하기에는 한계가 존재한다. 미 국립사법연구소(NIJ)의 보고서[1]에 의하면, 관제 요원이 직접 CCTV를 통해 상황을 감시하는 기존 CCTV 시스템에서는 상황을 감시하기 시작한 후, 22분이 지나면, 위험을 탐지하는 능력이 약 5%로 현저히 떨어진다는 실험결과가 존재한다. 따라서 이에 대한 해결 방안으로, 한정된 인력으로도 다수의 CCTV를 관리하고 모니터링 할 수 있는 시스템의 필요성이 대두되고 있다.



[그림 1-1] 연도별 공공기관 CCTV 설치 현황



[그림 1-2] 수동 영상관제 시간에 따른 위험탐지율

본 논문에서는 이러한 한계를 극복하기 위해 빈번히 발생하는 뺑소니 차량이나 도난 차량 및 범죄의 연루된 차량에 대한 수배를 조기에 처리하는데 도움을 줄 수 있는 차종 판별 시스템을 구현하는 것에 목적이 있다. 현재 CNN을 사용한 차종 판별은 시중에 나와있는 모든 차종을 판별하기에는 수가 방대하여 판별이 어렵고, 또한 모든 차종에 대한 데이터셋을 마련하기에는 무리가 있어 차종을 특정하여 차종을 판별하는 방향으로 연구가 진행되고 있다. Derrick Liu와 Yushi Wang은 196종의 차종 특정하고, 이를 판별하기 위해 SVM과 1-layer CNN를 기반으로 한 모델을 제안하였으며[2], Burak Satar와 Ahmet Emir Dirik은 Detector인 SSD[3]와 ResNet[4]를 사용한 모델을 제안하였다[5]. SVM과 1-layer CNN을 사용한 모델은 구조가 단순하고 간단하여 적은 시간으로도 학습이 가능하였으나, 100여종이 넘는 차종을 판별하기에는 데이터셋의 크기가 작고 모델의 깊이가 충분히 깊지 못해 overfitting 문제가 발생하였다. SSD와 ResNet을 조합한 모델은 Detector를 통해 먼저 차량을 인식한 후 학습한 모델로, Detector를 사용하지 않은 모델보다 높은 성능을 보였다. 하지만 Detector를 통해 먼저 차량을 인식하는 과정이 완벽하게 수행되지 않아 차종 판별에 있어 Detector의 성능이 모델의 성능에 크게 영향을 끼치는 단점이 존재했으며, 판별하는 차종의 수가 7종으로 매우 적었다.

이에 본 논문에서 구현하고자 하는 시스템에서는 차종을 한정한 후, 모델의 깊이를 조정하여 한정한 차종에 대해 적합한 모델의 깊이를 찾는 것을 목표로 하며, 학습을 진행하였을 때 발생하기 쉬운 overfitting 문제에 있어 dropout을 사용해 문제를 해결하고자 하였다.

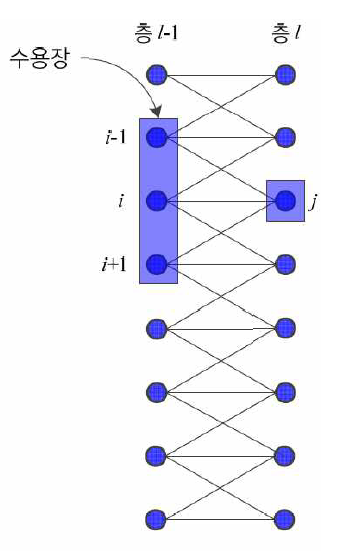
본 논문의 구현 환경은 Windows와 Linux환경에서Python기반 가상 개발 환경인 Anaconda를 사용하여 진행하며, 해당 학습을 위해 오픈 소스 이미지 데이터 셋인 Stanford Cars Dataset 224\*224 해상도로 일괄 조정한 이미지를 사용한다. 학습 방법은 ImageNet 데이터셋으로 사전 학습된 ResNet 모델을 사용해 Layer를 조정해가며 학습한다.

본 논문의 구성은 다음과 같다. 2장에서는 학습에 사용되는 핵심 기술인 CNN과 CNN모델인 ResNet에 대한 특징을 기술하였다. 3장에서는 본 논문에서 제안한 차종 판별의 설계 및 흐름도와 학습 모델을 사용하여 테스트를 진행하는 과정을 기술하였다. 4장에서는 해당 테스트의 결과를 설명하고 이를 분석하여 개선 방안을 제시한 뒤, 개선된 실험 결과와 이전 실험 결과를 비교 분석한다. 마지막으로 5장에서는 분석한 결과를 토대로 결론을 짓고 향후 연구 과제를 제시한다.

# 관련 기술 조사

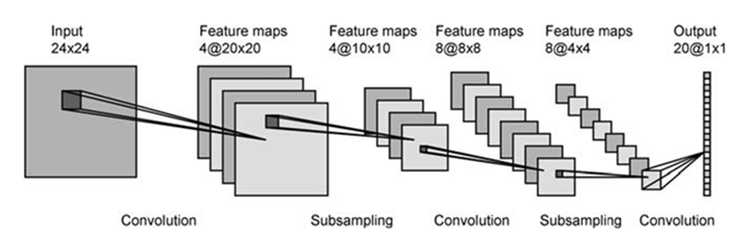
# 2.1 CNN

CNN(Convolution Neural Network)은 격자 구조(영상, 음성 등)를 갖는 데이터 학습에 적합한 신경망으로, 현재 다양한 분야에서 활용되고 있으며, 위 시스템 구현에 있어 필수 기술인만큼 해당 신경망에 대해 설명하려 한다.



[그림 2-1] CNN의 부분연결 구조

CNN은 각 Layer의 Perceptron 컨볼루션 연산을 이용한 부분연결 구조로 복잡도를 크게 낮추었다. 또한 연산을 위한 입력 데이터의 특징을 직접 추출하여 학습을 해야 하는 머신 러닝과는 달리, 입력 데이터의 특징을 해당 네트워크가 추출하여 학습에 사용하는 딥러닝 구조이다.



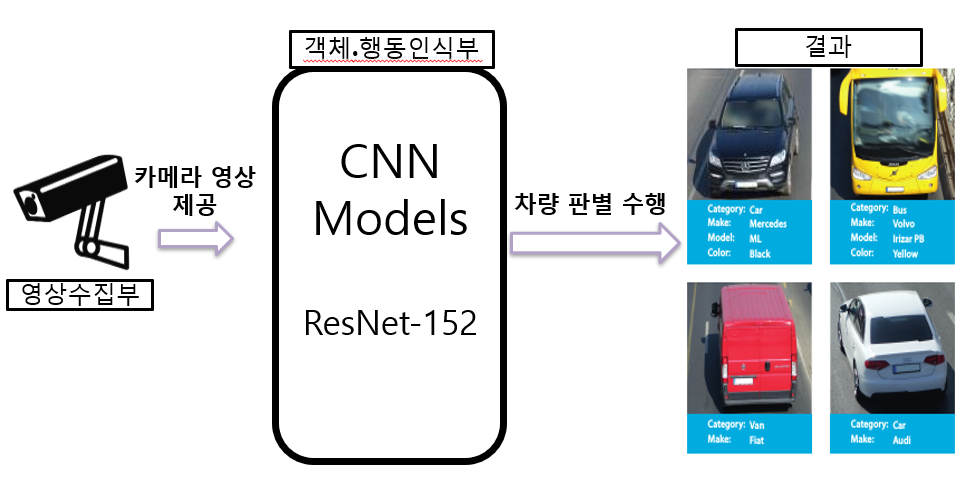
[그림 2-2] CNN 구조

# 본론

## 3.1 차종 판별 시스템

해당 장에서는 차종 판별을 가능케 하는 시스템의 설계와 구현 과정을 설명한다.

### 3.1.1 시스템 구성도



[그림3-1] 시스템 구성도

차종 판별 시스템의 구성도는 다음과 같다. 영상수집부로부터 카메라 영상을 제공하면, 객체⦁행동인식부에서 CNN 모델인 ResNet-152 모델로 차종 판별을 진행하고 결과를 도출한다.

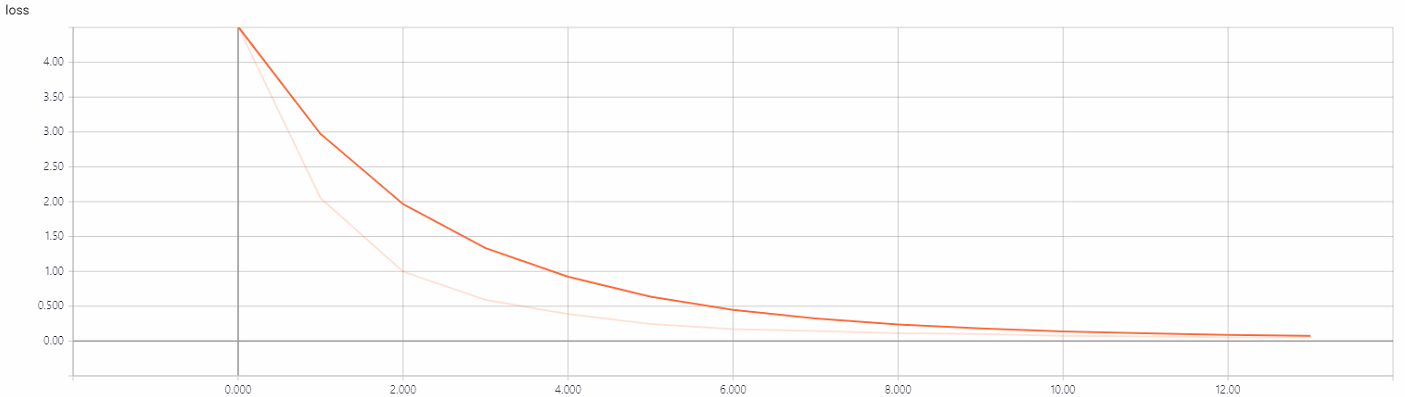
차종 판별 학습을 위해, 오픈 소스 이미지 데이터 셋인Stanford Cars Dataset을 사용하여 진행하며, 정해진 차종에 대한 새로운 이미지를 입력하여 학습이 정상적으로 진행되었는지 확인하고 문제점을 개선함으로써, 시스템을 구현하고자 한다.

### 3.1.2 데이터 구성에 따른 학습 결과

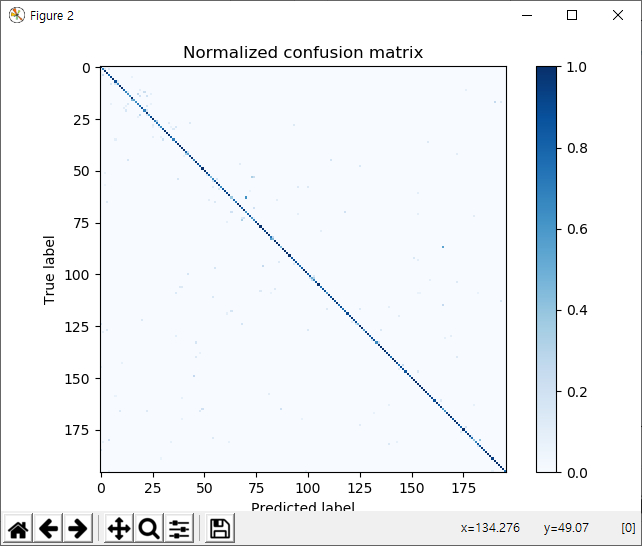
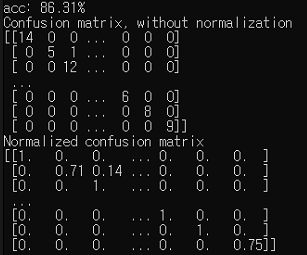


[그림3-2] Stanford Cars Dataset

3.1.1에서 명시한 dataset을 사용하여 데이터 셋을 구성하였다. 196종의 차종으로 이루어져 있으며, 6515장의 트레인 이미지와 1629장의 validation이미지, 8041장의 test 이미지로 약 50:50 비율로 트레인과 테스트 셋이 구성 되어있다. ImageNet(약 1400만 장)이나 COCO dataset(약 33만 장) 과 같은 방대한 양의 데이터셋으로 학습을 진행한다면 트레인과 테스트셋을 나누는 비율이 크게 학습에 영향을 끼치진 않지만, 위 학습에 사용하는 데이터셋은 2만장이 채 되지 않아 대형 데이터셋에 대비, 학습에 있어 트레인 테스트 데이터 셋의 비율은 훨씬 중요하다. 먼저 주어진 dataset을 변형하지 않고 학습을 진행해보았다.



[그림 3-3] Loss rate

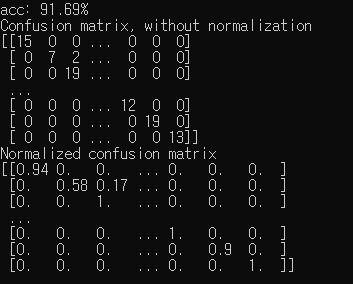


[그림 3-4] 학습 결과

해당 모델에 대한 차종 인식률을 테스트한 결과, 성능은 86.31%으로 90%에 못 미치는 결과가 집계되었다. 196종의 차종을 구별하는 만큼 준수한 성능의 인식률이라고 볼 수 있으나, 성능 개선을 위해 dataset의 트레인 테스트 데이터셋 비율을 50:50에서 80:20 비율로 변형하였다.



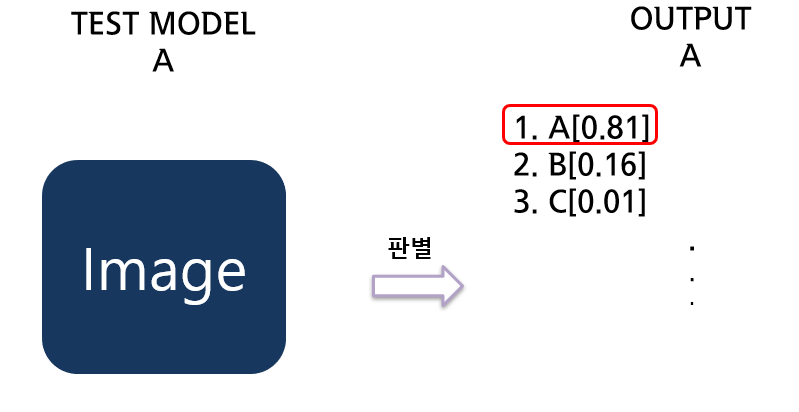
[그림 3-5] 변형된 데이터 셋

[그림 3-6] 변형된 데이터셋 학습 결과

학습 모델에 대한 테스트 결과 차종 인식률은 86.31%에서 91.69%로, 테스트 이미지 1608장 중 130장만이 오답을 출력하였고, 약 5%의 성능 개선을 보였다.

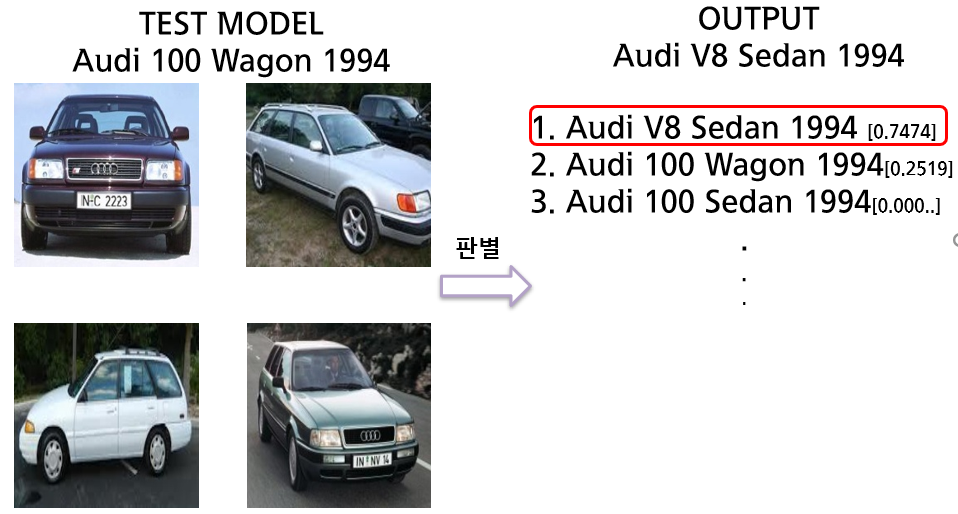
### 3.1.2.1 1차 성능 개선

3.1.2 절의 테스트 진행 중 오답을 출력한 130장의 이미지에 대해 분석하였다. 해당 오답 중 극단적인 결과를 출력한 차종을 분석하였으며, 먼저 차종 판별이 진행되는 방식을 예시를 들어 간략하게 설명 후 진행한다.



[그림 3-7] 차종 판별 예시

먼저 특정 차종 이미지를 입력하면, 시스템은 학습된 모델을 통해 196종의 차종을 유력한 순서대로 판별을 수행하며, 이중 가장 유력한 차종을 출력하는 방식이다. 그림 3-7의 예시는 A모델의 차량 이미지를 입력하였을 때 , 81%의 확률로 A 모델의 차량이라고 판별하였으며 정답을 출력한 예시이다.



[그림 3-8] 차종 판별 오답



[그림 3-9] 오답 차종 이미지

그림 3-8의 예시는 실제 130장의 오답 이미지 중 가장 많이 오답을 출력한 차종 중 하나이다. 해당 모델에 대한 정답 차종은 2위 후보에 위치하였으며, 가장 유력한 차종으로 판단되어 출력된 결과는 같은 브랜드 같은 차량 군이지만 다른 차종으로 오인하여 출력한 경우로, 실제 오답으로 출력한 차종에 대한 트레인 이미지인 그림 3-9를 보아도 판별이 모호한 것을 확인할 수 있었다. 해당 차종 이외에도 오답 이미지의 90%(117장)는 이와 같이 같은 브랜드 같은 차량 군이나 다른 차종으로 오인하여 출력한 경우이며, 정답 차종의 순위 또한 Top-5 안에 위치하였다.



[그림3-10] 차종 판별 중 특이한 오답

그림 3-10은 정답 차종의 순위가 Top-5에 위치하지 못하는 오답 이미지 중 10%를 차지한 결과 중 하나이다. 정답 차종이 9위에 위치해 있으며, 해당 차종과 출력된 차종의 구분이 육안으로 보기에 명확함에도 구분이 되지 않는 결과를 보였다.

### 3.1.2.2 중간 결과

앞서 학습을 진행한 결과 학습 모델의 인식률은 약 91.7%의 성능을 보였으며, 테스트 이미지 1608장을 대상으로 진행한 테스트에서 차종 판별에 실패한 이미지 130장을 확인한 결과, 정답이 Top-5안에 위치한 이미지가 90% (117장), 그렇지 않은 차종이 10%(13장)였으며, 해당 10%에서 같은 차종 간의 판별오류가 아닌 다른 차종(예: 세단과 SUV)간의 차종 판별이 확인되어 추가 데이터셋 확보를 통한 성능 개선을 진행할 예정이다. 또한 90%의 결과가 같은 브랜드의 같은 차량군에 위치한 차종에 대해서 판별에 어려움을 겪었으므로, 이들을 대상으로 차종을 구분하는 것은 의미가 없다고 판단하여, 이들을 차종 분류 대상에서 제거하고, 하나의 차종으로 병합하여 다시 결과를 확인할 예정이다. 번 외로, 결과가 정답으로 출력된 나머지 차종에 대해서, 정답으로 내놓은 출력이 평균 정답 출력보다 낮은 차종이 확인되었다. 판별은 문제가 없어 정상적인 결과가 출력되었지만 안정적으로 정답을 출력하는 것이 아니라고 볼 수 있기 때문에, 해당 출력이 낮은 차종에 대해서도 개선을 위한 추가 연구를 진행하는 것을 목표로 두었다.

## 3.1.2.3 2차 성능 개선

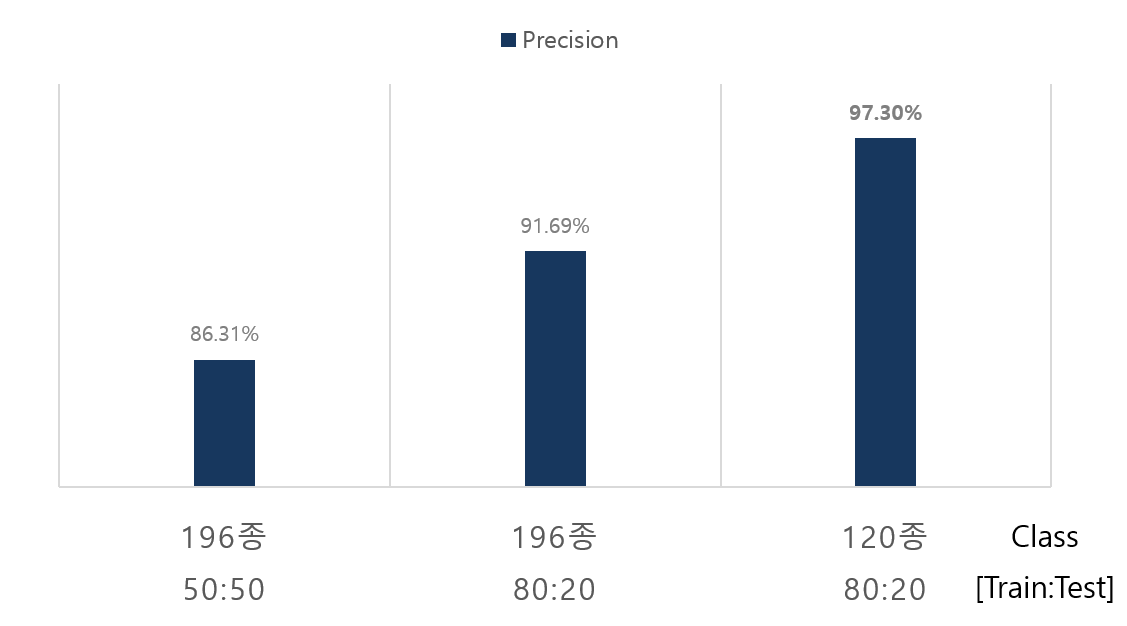
3.1.4절에서 설명한 문제점인 정답 차종의 순위가 Top-5 밖에 위치한 차종 문제, 같은 브랜드에서 같은 차종간 판별이 모호한 문제의 개선 방안을 이 절에서 설명한다.

먼저 정답 차종의 순위가 5순위를 초과한 차종에 대해 학습 데이터를 확인하였더니, 학습에 부적절하다고 판단되는 데이터가 포함되어 있었다. 따라서 정상적인 이미지로 대체하여 차종 데이터를 점검하였다.



[그림 3-11] 불량 데이터 교체

이후, 또 다른 문제로 대두된 같은 브랜드 차량간 차종 판별 문제에서, 같은 차종 (예:SUV, 세단, Coupe)간의 판별 문제의 개선 방안으로 같은 연식의 차량 또는 같은 차체를 가지는 차량을 병합하여 차종을 압축함으로써, 차종을 196종에서 120종으로 줄여 학습을 진행하였다.



[그림 3-12] 기존 모델과 개선된 모델과의 성능 비교

앞서 학습한 모델과 이 절에서 개선한 내용을 바탕으로 학습한 모델과의 성능 비교를 하였다. 차종을 196종에서 120종으로 줄인 결과 약 5.7% 가량의 성능 개선이 이루어졌으며, 기존 모델과는 약 11%의 개선이 이루어진 것을 확인할 수 있었다.

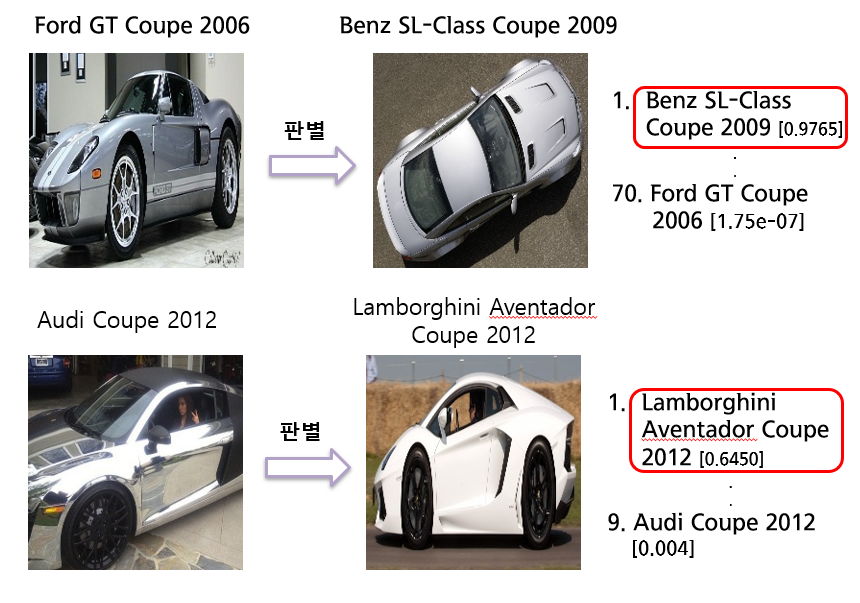
## 3.1.2.4 2차 결과

3.1.5절의 성능 개선이후, 3.1.2절에서 진행한 테스트를 동일하게 진행하였을 때, 전체 1608장의 이미지 중 오답 출력 이미지는 2차 성능 개선 전 130장에서 56장으로 감소하였고, Top-5 이상의 오답을 보이는 이미지는 2장으로 개선 전 모델 대비 안정적인 결과를 보였으나, 이 테스트 결과에서도 앞서 제기됐던 문제인 정답 출력이 평균 출력보다 낮은 차종이 동일하게 나타났다. 먼저 정답 이미지에 대해, 총 120장의 차종 중에, 16종의 차종이 평균 정답 출력(0.97)보다 낮은 출력을 보였고(0.74~0.94), 해당 차종에 대해 조사를 진행한 결과, 다른 브랜드의 차종이지만, 같은 연식의 차량 또는 같은 차종의 차량이라 디자인이 비슷하여 판별이 흐려지는 문제의 출력이 존재하였다.



[그림 3-13] 가장 낮은 출력을 보인 차종의 비교

위 문제의 예시로 120종에서 가장 낮은 출력을 보인 Chrysler사의 SUV 모델은 1순위 출력이 0.74로 정답을 출력하였으나, 2순위 출력으로 0.25를 출력하며, 두 종의 차종이 출력을 양분하는 결과를 보였다. 이 결과에 대해 조사하였더니, 2순위 출력이 Dodge 사의 SUV 모델을 기반하여 출시된 SUV 가 Chrysler사의 SUV 모델로, 육안으로 확인하였을 때, 각 회사의 특색이 되는 차의 그릴 부분을 빼고는 매우 비슷한 외형을 보였다.

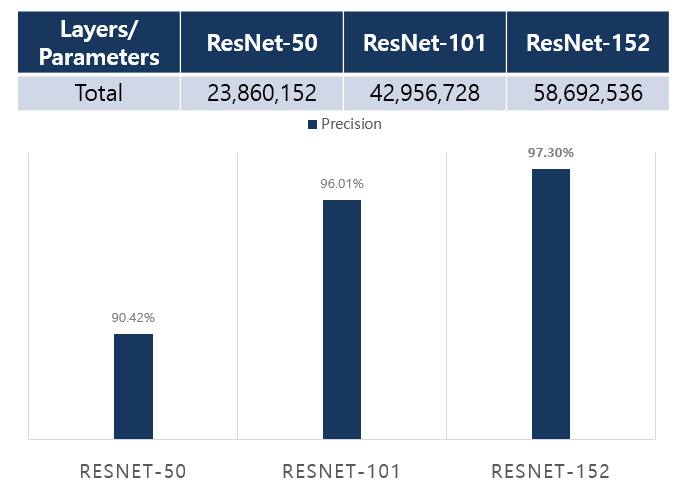


[그림 3-14] 정답이Top-5 이상의 순위로 출력된 오답 이미지

마지막으로 앞서 말한 2장의 오답 이미지를 도시한다. 해당 판별에 있어 사용되었던 이미지는, 높은 해상도의 원본 이미지를 일률적으로 224\*224 해상도로 변환함에 있어 찌그러진 외형을 가지게 되어 차종 판별의 무의미한 더미 데이터가 되었음을 확인하였다.

## 3.1.3 모델에 따른 학습 결과

3.1.2절에서 설명한 바와 같이, 학습에 기반이 되는 데이터셋이 대형 데이터셋 대비 매우 작은 Dataset이므로, 대형 데이터셋에서 쓰이는 학습 모델을 사용하는 것은, 많은 수의 파라미터 대비 적은 데이터를 가지므로 비효율적이라고 생각이 되었다. 따라서 데이터 구성에 따른 학습결과에서 도출한 최적의 데이터 구성을 바탕으로, 앞서 사용한 ResNet-152 모델을 기준으로 Layer층수 가 다른 모델인 ResNet-50과 ResNet-101 모델을 가지고 학습을 진행하였다.



[그림 3-15] 모델 별 성능 비교

학습된 결과를 비교한 결과, 3가지 모델 중에서는 ResNet-152 모델이 더 우수한 것으로 나타났으며, Layer수가 가장 적은 ResNet-50 모델에서 최저의 성능을 보였다.

## 결론 및 향후 연구 과제

위의 학습 결과들을 바탕으로 결론을 제시하면, 차종 판별에 있어 시스템을 구성할 때, 학습하는 모델에 대해 연구를 진행하였으며, 초기 설정된 데이터셋의 트레인 테스트 데이터 비율은 50대50이었으나, 80대20으로 바꾼 후 성능이 개선되었으며, 같은 브랜드의 차종 분류를 병합하여 차종을 196종에서 120종으로 압축하여 사전 구성된 데이터셋을 보강하였다. 이후 이 데이터셋을 기반으로 다른 깊이를 가지는 학습 모델을 가지고 학습을 진행하였을 때, 152 계층을 가지는 모델의 성능이 가장 좋은 결과를 보였다. 하지만, 여전히 평균 출력에 못 미치는 출력을 보이는 차종이 존재하며, 오답 이미지 데이터에 억지로 해상도를 낮추어 생긴 찌그러진 형태의 이미지를 확인하여 이를 정상적인 이미지로 대체하여야 할 것이다. 또한 크게 3가지의 학습 모델만을 비교하였으나, 데이터셋에 대한 최적의 결과를 도출하기 위해, 모델의 깊이를 101에서 152 사이에서 조절하며 성능을 비교하고, 또한 깊이가 152이상일때의 결과도 비교하여 효율적인 모델을 찾는 것을 목표로 할 것이다.

# 참고문헌

1. NIJ. <Ministry of Justice, USA>. Buyer Beware Vol.10 (11, 2002)
2. Derrick Liu, Yushi Wang. Image Classification of Vehicle Make and Model Using Convolutional Neural Networks and Transfer Learning. Stanford University. (2015)
3. Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg. SSD: Single Shot MultiBox Detector. UNC Chapel Hill , Zoox Inc. Google Inc. University of Michigan, Ann-Arbor. (29 Dec 2016)
4. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition. (10 Dec 2015)
5. Burak Satar, Ahmet Emir Dirik: ‘Deep Learning Based Vehicle Make-Model Classification’. Uludag University, Bursa, Turkey (9 Feb 2019)
6. 신동, 김병만 (2016). Convolutional Neural Network와 Tensorflow를 활용한 차량 모델 판별. 한국정보과학회 학술발표논문집, 2074-2076.
7. Linjie Yang, Ping Luo, Chen Change Loy, Xiaoou Tang. A Large-Scale Car Dataset for Fine-Grained Categorization and Verification Department of Information Engineering. The Chinese University of Hong Kong Shenzhen Key Lab of CVPR, Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China (arXiv:1506.08959v2 [cs.CV] 24 Sep 2015).
8. Adami Fatima Zohra, Salmi Kamilia, Abbas Fayçal, Saadi Souad. Detection And Classification Of Vehicles Using Deep Learning. Department Of Computer Science University of Abbes Laghrour khenchela, Algeria. International Journal of Computer Science Trends and Technology (IJCST) – Volume 6 Issue 3, May - June 2018.
9. Ms. Vijayasanthi D,Mrs. Geetha S. DEEP LEARNING APPROACH MODEL FOR VEHICLE CLASSIFICATION USING ARTIFICIAL NEURAL NETWORK, International Research Journal of Engineering and Technology (IRJET), Volume: 04 Issue: 06, June -2017
10. Jong Taek Lee and Yunsu Chung. Deep Learning-based Vehicle Classification using an Ensemble of Local Expert and Global Networks. Electronics and Telecommunications Research Institute (ETRI), Daegu, South Korea. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)
11. Jonathan Krause, Jia Deng, Michael Stark, and Li Fei-Fei. Collecting a Large-Scale Dataset of Fine-Grained Cars. Computer Science Department, Stanford University. Second Workshop on Fine-Grained Visual Categorization (FGVC2). Portland, OR. June 28, 2013.
12. Timnit Gebru, Jonathan Krause, Yilun Wang, Duyun Chen, Jia Deng, Li Fei-Fei. Fine-Grained Car Detection for Visual Census Estimation. Department of Computer Science, Stanford University. AAAI Conference on Artificial Intelligence AAAI. 2017.