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

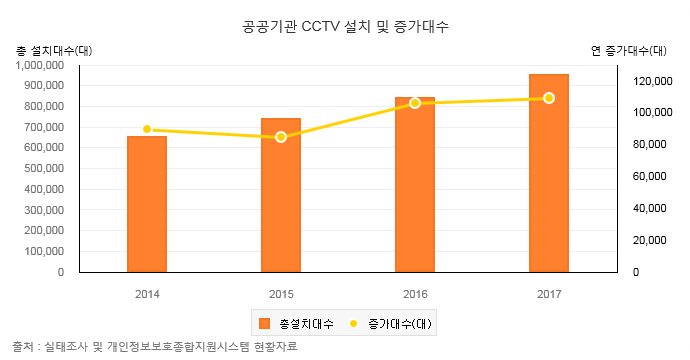
[그림 3-14] 정답이Top-5 이상의 순위로 출력된 오답 이미지……………………………..17

[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. 또한 Lee 와 Chung이 제안한 또 다른 모델[6]은 차량을 소그룹으로 묶은 데이터셋을 각각 학습한 local 모델과 차량 전체를 데이터셋으로 학습한 global 모델로 네트워크를 이루고 Softmax로 결과를 출력하는 모델이다. 성능은 최저 68.72%에서 최고 99.66%으로 매우 정확하지만, 해당 모델을 위해서 12개의 local 모델과 6개의 global 모델을 사용해 엄청난 계산 능력이 요구되어 다수의 고 사양 GPU를 필요로 하는 단점이 있다. 따라서 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.

1. 기술조사

다른 네트워크 모델을 사용했을 때 무엇이 문제인지 -> 기술설명

왜 이런 알고리즘, resnet을 사용하는지 -> 기술설명

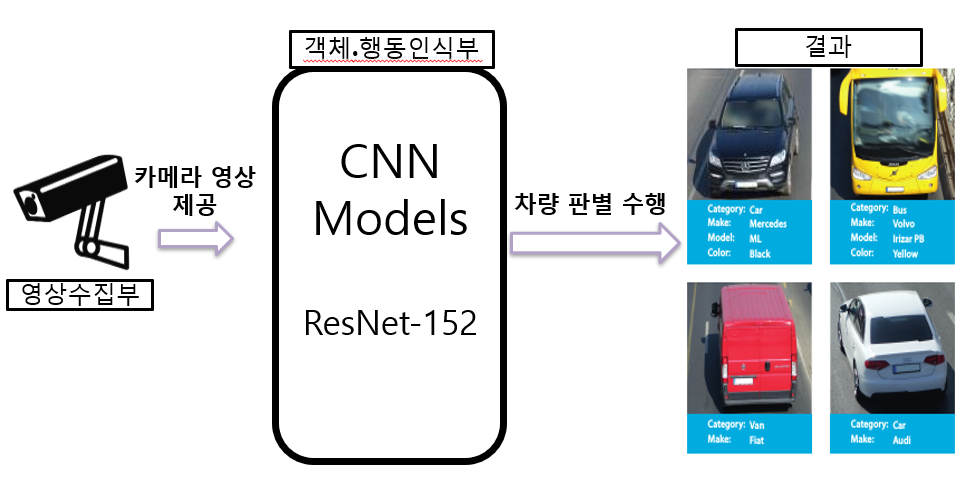
Resnet을 사용한 사례가 무엇이 있는지 ->기술설명

* 1. Cnn 설명
  2. Resnet 설명
  3. 성능비교, 책정 기준

다른 네트워크 모델과 resnet 성능 비교하여 최종적으로 renset사용하는 정당성 부여

1. 차종 판별 시스템의 구현
   1. 시스템 구성도

차종 판별 시스템의 구성도는 다음과 같다.



[그림3-1] 시스템 구성도

해당 시스템은 영상을 수집하는 영상수집부와 수집한 영상을 입력 받아 차량을 판별하는 객체⦁행동인식부로 나뉜다. 영상수집부에서 수집한 영상 프레임은 차량 판별을 위해 224X224 해상도로 변환하며, 변환된 프레임은 객체⦁행동인식부에서 CNN 사전 학습 모델인 ResNet-152 모델로 차종 판별을 진행하고 결과를 도출한다. 차종 판별을 위한 CNN 모델 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.

* 1. 데이터 셋 구성 및 모델 학습



[그림3-2] Stanford Cars Dataset

3.1에서 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 dataset(약 1400만 장)이나 COCO dataset(약 33만 장) 과 같은 대형 Dataset으로 학습을 진행한다면 트레인과 테스트셋을 나누는 비율이 크게 학습에 영향을 끼치진 않지만, 위 학습에 사용하는 데이터셋은 2만장이 채 되지 않아 학습에 있어 트레인 테스트 데이터 셋의 비율은 훨씬 overfitting에 민감하다. 따라서 주어진 dataset을 기존 50:50 비율과 다른 비율로 나누어 학습을 진행하였다.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Ratio | 50:50 | 60:40 | 70:30 | 80:20 | 90:10 |
| Train/Test | 6515/  8041 | 9711/  6474 | 11330/4855 | 12948/3237 | 14567/1618 |
| acc | 83.76% | 86.31% | 88.24% | **91.69%** | 90.89% |

[표 3-1] 비율 성능표

#####성능표에 따른 최적의 트레인 테스트 셋의 비율을 도출하고 해당 80:20 비율에서의 결과를 명시.

테스트 이미지 1608장 중 130장만이 오답을 출력하였고, 약 5%의 성능 개선을 보였다.

|  |  |  |
| --- | --- | --- |
|  | Car Classes | Top-5 acc |
| AlexNet 50:50 dataset ratio | 196 | 0.417 |
| GoogleNet 50:50 dataset ratio | 196 | 0.8 |
| ResNet-152 50:50 dataset ratio | 196 | 0.8376 |
| ResNet-152 dataset 60:40 ratio | 196 | 0.8631 |
| ResNet-152 dataset 70:30 ratio | 196 | 0.8824 |
| ResNet-152 dataset 90:10 ratio | 196 | 0.9089 |
| ResNet-152 dataset 80: 20 ratio | 196 | 0.9169 |
| ResNet-152 dataset 80:20 ratio | 120 | **0.9730** |

# 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. Jong Taek Lee, 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)
7. 신동, 김병만 (2016). Convolutional Neural Network와 Tensorflow를 활용한 차량 모델 판별. 한국정보과학회 학술발표논문집, 2074-2076.
8. 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).
9. 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.
10. 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
11. 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)
12. 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.
13. 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.