

컨볼루션 신경망 기반 적외선 영상 이용 탑승객 검출

김상혁⁰¹

서강대학교 전자공학과

hitboy91@gmail.com

Occupant Detection System by Infrared Images Based on Convolutional Neural Network

Sanghyuk Kim⁰¹

Electronic Engineering, Sogang University

요 약

이 논문에서는 Convolutional Neural Network(CNN)을 기반으로 하여 적외선 영상을 이용하여 탑승객 검출 시스템을 제안한다. 먼저 운전자 좌석과 유사한 환경을 조성하여 적외선 카메라를 이용하여 영상 데이터 수집한다. 구축된 적외선 영상 데이터를 이용하여 MobileNet 구조에 학습을 시킨다. 학습된 필터의 Weight를 동일한 CNN 구조의 피드 포워드(Feed forward) 구조에 로드한다. 이때 실시간 임베디드 시스템을 구현하기 위하여 피드 포워드 구조는 C언어로 제작하였다. 제안된 시스템은 학습기와 검출기를 분할하여 그래픽 처리 장치(Graphic Processing Unit, GPU)없이 중앙 처리 장치(Central Processing Unit, CPU)로만 구성된 임베디드에 적용가능하도록 하여 다양한 임베디드 환경에서 시스템을 구현할 수 있도록 하였다.

1. Introduction

현재 자율 주행 자동차에 대한 연구가 진행됨에 따라 운전자뿐만 아니라 탑승객 전체의 안전에 대한 관심이 높아지고 있다. 현재 WP29에서 자율 주행 Level 2~3 자동차 내·외부 안전 표준 기준이 정립 중에 있다. 이러한 탑승객에 대한 안전과 가장 밀접한 연관이 있는 것이 바로 안전벨트 착용이다. 삼성 교통안전문화 연구소의 연구결과에 따르면 안전벨트 미착용 시 교통사고 치사율(2.4%)은 착용시(0.2%)보다 12배 높다. 이에 따라 한국에서는 2017년 6월 이후 전 좌석 안전벨트 착용 의무화가 이루어졌다. 안전 벨트 착용 의무화를 위한 첨단 운전자 지원 시스템(Advanced Driver Assistance System, ADAS)이 개발되고 있다. 대표적인 것이 차량 전 좌석에 대한 탑승 인원 및 위치를 판단하는 시스템이다.

기존 탑승객 유무를 판단하기 위해서는 중량 센서를 이용하여 판단하였으나, 좌석 위의 사물이 사람으로 인지되는 한계가 있다. 이를 해결하기 위해 컴퓨터 비전을 이용하여 영상 데이터 기반 탑승객 검출 시스템을 개발할 수 있다. 현재 컴퓨터 비전 분야에서 딥러닝 알고리즘이 널리 사용되고 있다. 하지만 기존 AlexNet[1] 및 VGGNet[2]의 경우 Weight가 많아 Memory inference 측면에서 임베디드 환경에 구현하기에 부적합하였다. 이후 Memory inference를 고려한 CNN 모델인 MobileNet[3] 및 ShuffleNet[4]등이 점차 개발됨에 따라 CNN 구조의 전체적인 Parameter 수가 축소되어 임베디드 시스템에서 활용할 수 있는 CNN 기반 시스템이 개발되기 시작하였다.

또한 기존 CNN 모델은 RGB 영상의 데이터베이스를 기반으로 제작되었다. RGB 영상은 조도 변화에 민감하다. 따라서 차량 내부와 같이 조도 변화가 빈번한 환경에서 사용할 영상 데이터 베이스로 RGB 영상은 완벽하지 않다. 이를 위해 조도 변화에 강건한 적외선 영상을 이용한 시스템을 개발할 수 있다. 예를 감시 카메라 시스템에서는 주·야간에 상관없이 동작하여야 하므로 적외선 영상이 많이 이용되고 있다. 이처럼 적외선 영상의 이용 범위가 늘어나고 있으며 자동차 산업에서도 점차 적외선 영상을 기반 시스템 개발을 진행하고 있다.

본 논문은 검증된 Convolutional Neural Network(CNN)을 기반으로 하여 차량 내부 환경에 맞는 적외선 영상을 제작 및 활용하여 탑승객 검출 시스템을 제안한다.

2. Customized infrared database construction

차량내 환경에서 적합한 시스템 구현을 위하여 자체 적외선 영상 데이터베이스 구축하였다. 이때 Driving simulator를 설치하여 실제 운전 상황과 유사한 환경을 구축하였다. 또한 차량내 선바이저와 유사한 각도로 운전자를 바라볼 수 있도록 카메라를 설치하였다. 이는 그림 1과 같다.

생성된 데이터 베이스는 차량 내에서 운전자가 할 수 있는 여러가지 상황에 대한 시나리오를 작성하여 1차적으로 수집하였다. 이때, 실험참가자에게 일정부분 움직임의 자유를 허락하여 여러가지 시나리오 이외의 예외 상황을 만들어 데이터 베이스의 다양성을

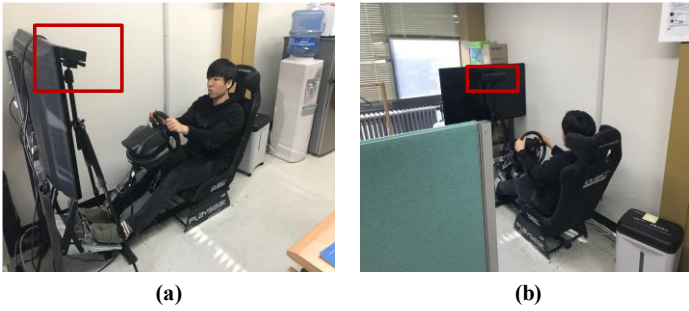


그림 1. 적외선 영상 데이터베이스 구축 환경: (a) 전면, (b) 후면.

확장하였다. 이는 그림 2와 같다. 이를 기반으로 1차적으로 탑승객 검출 시스템을 제작한 후, 데이터베이스를 추가적으로 수집하였다. 이때 탑승객을 검출하지 못한 영상(Error database)을 중점적으로 수집하여 오차율을 감소시켰다.

3. Learning by Convolutional Neural Network

제안된 시스템은 차량에서 동작해야 하므로 CNN 모델은 CPU 만으로 임베디드 환경에서 구현 가능한 모델이 선택되어야 한다. 검증된 CNN 모델 중 Memory inference와 실시간 처리가 가능한 작은 모델을 선택하였으며, 이는 MobileNet[3]이다. MobileNet은 기존 Convolutional 필터를 두개의 Seperable convolutional 필터인 Depth-wise와 Point-wise convolutional 필터를 이용하여 필터의 파라미터 수 및 연산량을 감소 시켰다.

또한 기존 MobileNet의 경우 1,000 class에 대한 결과를 출력하였으나, 제안된 시스템에서는 탑승객 유무만 판단하는 Binary 결과를 출력한다. 따라서 Model의 Width(Filter의 수)를 Reference model을 1로 잡을 경우, 제안된 시스템에서는 0.5 model을 사용하여 적외선 영상을 학습시켰다. 학습시 Gradient decent는 Adam[]을 채택하였으며, Learning rate는 아래 Polynomial decay 방법을 채택하였다.

$$base_lr_n = base_lr_{n-1} \times (1 - \frac{iter}{max_iter})^{power} \quad (1)$$

이때 $base_lr_n$ 은 n번째 학습률, $power$ 는 power, $iter$ 은 현재 반복 횟수 그리고 max_iter 은 총 반복 횟수를 의미한다.

Weight 초기화시에는 Microsoft Research Asia(MSRA) 방법[5]을 채택하여 Symmetry를 제거하는 방향으로 초기화를 실시하였다.

4. Occupant detection based on feed forward network

학습된 Weight를 이용하여 실시간 탑승객 검출을 하기 위해서 Off-line 학습만 실시한다. 따라서 Weight를 로드할 CNN 모델은 학습 모델과 동일하나, Feed forward network만 구현하였다. 이때 다양한 임베디드

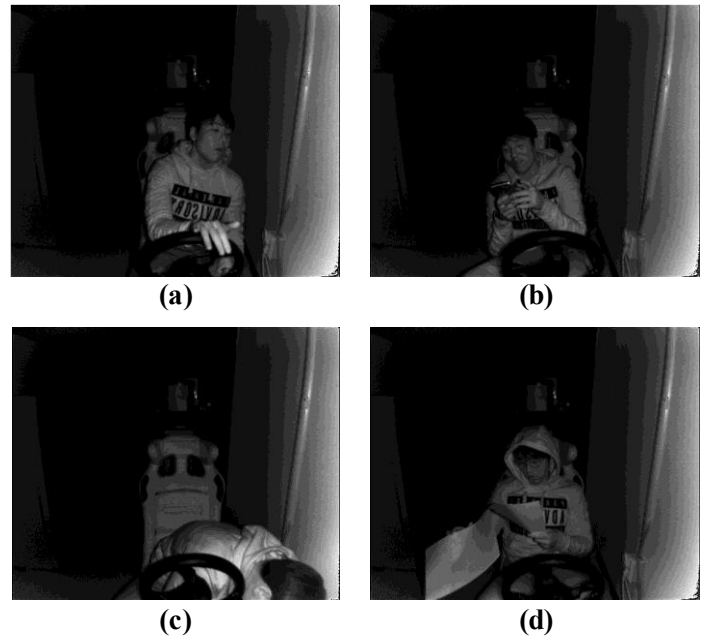


그림 2. 데이터 베이스 구성 예시: (a) 운전, (b) 휴대전화 사용, (c) 얼굴이 가려진 몸의 일부만 노출, (d) 책 읽기(예외 상황).

환경에서 추가적인 포팅을 실시하지 않게 하기 위해서 C 언어로 검출기를 제작하였다. 1차 데이터베이스를 기반으로 탑승객 검출기를 제작하여, Error 데이터베이스를 습득하는데 활용한다. 이후 Error 데이터베이스를 추가적으로 학습된 Weight를 로드하여 최종 탑승객 검출 시스템을 완성하였다. 이는 그림 2와 같다.

5. Implementation

제안한 시스템을 실험하기 위해 사용한 장비는 다음과 같다. 데이터 학습을 위해 사용한 Deep learning 라이브러리는 Caffe이며, Ubuntu 16.04, CUDA 8, cuDNN5 환경에서 학습을 실시하였다. PC의 Specification은 Intel Xeon E5-2687W v4(3.00GHz, 24 cores)이며, GPU는 GeForce GTX 1080 Ti(250W, 11GB), 128GB DDR4 RAM로 구성되어 있다.

탑승객 검출을 위해 사용한 추가적인 외부 라이브러리는 없으며, Windows 10, Visual Studio 2015 환경에서 동작하였다. PC의 Specification은 Intel Core i5-6600 (3.30GHz), 16GB DDR4 RAM으로 구성되어 있다.

적외선 영상 수집을 위해 사용한 카메라는 Kinect for Window version 20이며, specification은 최대 적외선 화질 512(가로)x424(세로) 픽셀이며, Windows 10 환경에서 동작하였다.

6. Experimental results

실험 결과는 그림 3과 같다. 그림 3의 (a)와 (b) 같이 입력 영상에서 얼굴이 일부 가려진 상황에서도 탑승객

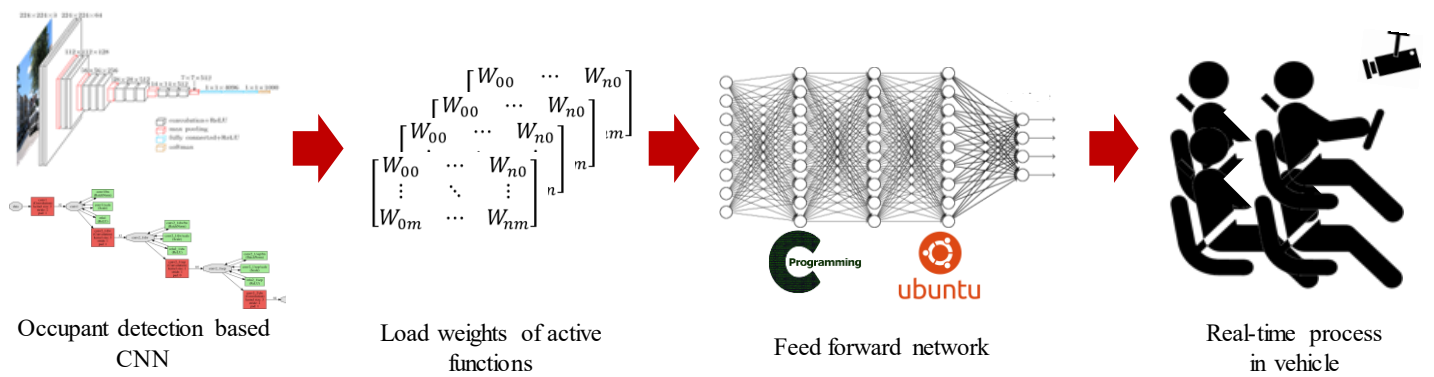


그림 3. 탑승객 검출 시스템 제작 과정



그림 4. 탑승객 검출 결과: (a) 두꺼운 옷을 착용한 상태에서 방안의 불을 켜진 상태, (b) 얇은 옷을 착용한 상태에서 방안의 불을 끈 상태

존재를 파악할 수 있다. 이는 적외선 영상 데이터베이스에 얼굴이 가려진 영상도 다수 포함됨에 따라 학습시 필터의 추상화 과정에서 얼굴에 대한 Feature뿐만 아니라 사람의 형태를 나타낼 시에는 사람으로 인지하도록 학습이 되었음을 알 수 있다.

또한 그림 3의 (a)와 (b) 같이 입력된 영상 속 인물의 옷의 전반적인 색과 두께에 관계가 없이 출력 결과가 운전자임을 알 수 있다. 이 역시 단순히 Pixel-level로 인한 판단이 아님을 알 수 있다. 이는 영상 속에서 행동이 다양해진 상황과 위치의 다름을 통해서도 알 수 있다. 즉, 탑승객이 카메라와 상대적 위치가 달라짐에 따라 발생하는 탑승객의 영상 속 크기가 달라져도 CNN 모델의 결과는 동일하게 출력됨을 알 수 있다.

그리고 방안의 불을 켜진 상태와 끈 상태에서 모두 실험한 결과 입력 영상 자체에 큰 변화가 없음을 알 수 있다. 이는 적외선 영상을 이용함에 따라 조도의 변화에 강건한 시스템을 구현할 수 있게 되었음을 알 수 있다.

7. Conclusion

이 논문에서는 CNN을 기반으로 하여 적외선 영상을 이용하여 탑승객 검출 시스템을 제안하였다. 자체 제작 적외선 영상 데이터 베이스를 활용하여 Parameter 수가 적은 CNN 모델인 MobileNet을 이용하여 필터의 Weight 학습을 실시하였다. 학습된 필터의 Weight를 동일한 CNN 구조의 피드 포워드 구조에 로드하여 실시간 탑승객 검출을 하였다. 본 연구 결과를 기반으로

하여 향후 차량 내 탑재되어 방치승객으로 생기는 사건을 미연에 방지할 수 있을 것으로 사료된다.

참 고 문 헌

- [1] A. Krizhevsky, et al., "ImageNet Classification with Deep Convolutional Neural Networks," Conference on Neural Information Processing Systems, 2012.
- [2] K. Simonyan, et al., "Very Deep Convolutional Networks for Large-Scale Image Recognition," Conference on Computer Vision and Pattern Recognition, 2015.
- [3] A. G. Howard, et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," arXiv, 2017.
- [4] X. Zhang, et al., "ShuffleNet: an Extremely Efficient Convolutional Neural Network for Mobile Devices," Conference on Computer Vision and Pattern Recognition, 2017.
- [5] K. He, et al., "Delving Deep into Rectifiers: Surpassing Human-level Performance on Imagenet Classification," International Conference on Computer Vision, 2015.