

Abstract

졸음 운전으로 인한 사고가 늘어나고 있다.
뭐시기 뭐시기

(사진)

개요

본 연구는 졸음 운전 감지 모델을 학습시킬 이미지와 라벨들을 생성할 때, OpenCV 와 Dlib 중 어떤 라이브러리의 정확도가 더 높은지 각 라이브러리에 같은 사진들을 labeling 하여 학습 데이터를 생성한 후, 각 학습 데이터를 같은 학습 모델에 학습시켜 실제 이미지를 투입했을때 어떤 학습 데이터를 이용했을 때 더 높은 감지율을 보이는지 비교해보고자 한다.

본론

OpenCV 와 Dlib 의 이미지 감지 방식은 다음과 같은 차이가 있다.

(사진)

OpenCV 의 경우, 라이브러리에서 눈, 코, 입을 감지하는 기능을 제공한다. 이 기능을 이용하여 눈과 입을 얼굴 사진에서 추출할 수 있다.

Dlib 의 경우, 이미지 속 사람 얼굴을 감지한 후, 사람 얼굴에 68 개의 지점을 좌표값으로 반환한다. 이렇게 얻은 좌표값을 이용하여 눈과 입을 얼굴 사진에서 추출할 수 있다.

본 연구는 Kaggle 의 데이터셋과 인터넷에서 추가로 수집한 이미지, 그리고 연구자의 얼굴 사진까지 총 6000 장의 이미지를 사용한다. 이 이미지들을 Haar Cascade 방식을 이용한 OpenCV 와 HOG + Linear SVM 방식을 이용하는 Dlib 을 이용하여 각각 labeling 을 수행한다. 이미지 labeling 은 눈 떠짐(eye_open), 눈 감김(eye_closed), 입 벌림(yawn), 입 닫음(no_yawn) 으로 나누어진 디렉토리에 각 상태에 해당하는 이미지를 저장하는 방식으로 수행한다.

이렇게 labeling 되어진 이미지들을 TensorFlow 와 Keras API 를 이용하여 작성된 학습 모델이 학습하여 실제 이미지를 입력 받았을 때 얼마나 정확하게 이미지 속 사람의 눈과 입의 열림 여부를 판단하는지 확인한다. 그리하여 OpenCV 와 Dlib 중 어떤 라이브러리로 수행한 labeling 이 더 정확한 감지를 할 수 있도록 학습 데이터를 만들었는지 비교할 수 있다.

결론

학습 데이터셋 A 를 사용하여 labeling 된 이미지를 학습한 결과와 B 를 사용하여 labeling 된 이미지를 학습한 결과를 비교하였다. 이러한 비교를 통해 학습 데이터셋 A 가 B 보다 n% 더 높은 정확도를 나타내는 것을 확인할 수 있다. 따라서 졸음 운전 감지를 위한 인공지능 모델을 구축하고 학습시키는 데 있어 A 를 사용하여 labeling 된 데이터셋이 더 정확한 학습 결과를 도출하며 우수한 성능을 발휘할 것으로 판단된다.