

졸음운전 경고 시스템 개발을 위한 Open CV와 Dlib의 이미지 판별 정확도 비교

강다빈, 조우현, 송준용, 최정열

성결대학교 컴퓨터공학과

{cherifr29, jeremiahjoh, luckyjude, passjay}@sungkyul.ac.kr

Comparison of Image Discrimination Accuracy between OpenCV and Dlib for Drowsy Driving Warning System Development

Dabin Kang, Woohyun Joh, Junyong Song, JungYul Choi

Department of Computer Engineering, Sungkyul University

요약

본 논문은 졸음운전 경고 시스템 개발을 위하여 졸음 이미지 판별 모델을 구축할 때 OpenCV와 Dlib 중 어떤 전처리 방법이 정확도가 높은지 비교, 분석하였다. 데이터 세트는 눈 뜬(open), 눈 감음(closed), 하품(yawn), 입 닫음(no_yawn)으로 레이블링된 이미지를 사용했으며, 졸음이 예상되는 상태는 눈 감음(closed), 하품(yawn)으로 설정하였다. 데이터 세트를 OpenCV의 Haarcascade와 Dlib의 HOG를 사용하여 각각 전처리한 후 학습하였다. 그 결과 하품을 감지하는 모델과 눈 깜박임을 감지하는 모델 모두 Dlib이 OpenCV보다 더 높은 정확도를 나타내었다.

I. 서론

졸음운전으로 인한 교통사고에 대해 사회적 우려와 관심이 높아짐에 따라 인공지능 기술을 이용하여 운전자의 졸음을 감지 및 경고하는 연구들이 선행되어 왔다. OpenCV의 Haarcascade를 이용하여 자동차 운전자의 눈의 상태(눈 뜬, 눈 감음)와 깜박임 속도를 고려하여 졸음 여부를 감지하는 연구가 진행된 바가 있다[1]. 또한 Dlib의 HOG를 이용하여 눈의 상태(눈 뜬, 눈 감음), 하품, 머리 자세 감지를 통한 졸음 여부를 감지하는 연구가 진행되었다[2]. 선행 연구들은 OpenCV, Dlib 등과 같은 라이브러리를 이용해 얼굴을 탐지하여 졸음을 판별하였지만, 각 라이브러리에 따른 모델의 정확도는 분석된 바가 없다. 얼굴 탐지를 위한 다양한 라이브러리가 존재하므로 효율적인 졸음 경고 모델을 만들기 위해 라이브러리에 따른 정확도 분석이 필요하다. 이에 본 논문은 졸음운전 방지를 위한 얼굴 감지 모델에 대해서 OpenCV와 Dlib를 각각 사용하여 전처리한 모델의 성능을 비교하고자 한다. 동일한 데이터 세트를 OpenCV의 Haarcascade와 Dlib의 HOG[Histogram of Oriented Gradients]를 통해 각각 전처리한 후 학습을 하여 얼굴 감지 모델을 만든다. 눈과 입을 감지하여 눈 깜박임 여부(open/closed)와 하품 여부(yawn/no_yawn)를 판단함으로써 어떤 모델이 졸음 경고 이미지를 판별하는데 더 높은 정확도를 보이는지 비교하고자 한다.

II. 본론

2.1 데이터 세트

학습 및 검증 데이터 세트는 AI Hub[3]가 제공한 졸음운전 예방을 위한 운전자 상태 정보 이미지, 인터넷으로부터 크롤링한 이미지, 그리고 연구자 3명의 이미지를 바탕으로 6,000장으로 구축되었다. 학습 데이터 세트와 검증 데이터 세트의 비율은 7 : 3으로 학습 데이터 세트는 4,200장, 검증 데이터 세트는 1,800장이다. 테스트 데이터 세트는 400장의 다른 2명의 연구원 이미지로 구축되었다. 각 이미지는 눈 뜬(open), 눈 감음(closed), 하품(yawn), 입 닫음(no_yawn)을 기준으로 레이블링이 되었으며, 졸음이

예상되는 상태는 눈 감음(closed), 하품(yawn)으로 설정하였다. 4가지로 분류된 이미지는 모두 같은 비율로 나누어졌으며 정면, 측면의 얼굴 이미지로 구성되었다.

2.2 전처리 방법

OpenCV와 Dlib 라이브러리를 이용한 졸음 감지 모델의 정확도를 비교 분석하기 위해서 각각의 라이브러리로 이미지를 전처리하였다. 다만, OpenCV와 Dlib은 다른 라이브러리이므로 얼굴 감지 방식이 달라 전처리 과정이 다를 수밖에 없다. OpenCV를 이용한 데이터 세트의 전처리 과정은 다음과 같다. OpenCV의 Haar Cascade 알고리즘을 이용하여 하품(yawn), 입 닫음(no_yawn)으로 레이블링된 이미지를 흑백으로 변환하여 입 부분의 좌표를 추출한다. 이후 원본 이미지에서 추출된 좌표만큼 잘라내어 이미지 정보와 레이블링 정보(하품: 0, 입 닫음: 1)를 학습 데이터로 저장한다. 눈 뜬(open), 눈 감음(closed)으로 레이블링된 이미지 또한 흑백으로 변환하여 눈 부분의 좌표를 추출한다. 이후 원본 이미지에서 추출된 좌표만큼 잘라내어 이미지 정보와 레이블링 정보(눈 뜬: 0, 눈 감음: 1)를 학습 데이터로 저장한다.

Dlib를 이용한 데이터 세트의 전처리 과정은 다음과 같다. HOG와 Linear SVM을 이용하여 얼굴 주요 부분에 68개의 랜드마크를 표시한다. 하품(yawn), 입 닫음(no_yawn) 데이터 세트에서 입의 랜드마크 좌표를 추출하여 이미지를 잘라낸 후 이미지 정보와 레이블링 정보(하품: 0, 입 닫음: 1)를 학습 데이터로 저장한다. 눈 뜬(open), 눈 감음(closed) 데이터 세트에서는 눈의 랜드마크 좌표를 추출하여 이미지를 잘라낸 후 이미지 정보와 레이블링 정보(눈 뜬: 0, 눈 감음: 1)를 학습 데이터로 저장한다.

본 실험에서는 하품(yawn), 입 닫음(no_yawn) 데이터 세트를 OpenCV로 전처리하여 (OpenCV)yawn_no_yawn모델의 학습 데이터로 사용하였고, Dlib으로 전처리하여 (Dlib)yawn_no_yawn모델의 학습 데이터로 사용하였다. 또한 눈 뜬(open), 눈 감음(closed) 데이터 세트를 OpenCV로 전처리하여 (OpenCV)open_closed모델의 학습 데이터로 사용하였고 Dlib으로 전처리하여 (Dlib)open_closed모델의 학습 데이터로 사용하였다.

2.3 모델 학습

(OpenCV)yawn_no_yawn, (Dlib)yawn_no_yawn, (OpenCV)open_closed, (Dlib)open_closed의 학습 모델은 표 1과 같다. 학습 방법으로는 이미지 인식에 적합한 CNN을 사용하였고 batch size 16으로 지정했으며 100 epoch만큼 학습을 진행하였다.

Table. 1 학습 모델

| Layer Name | Layer Size | Activation |
|-----------------------|------------|------------|
| Conv2D, MaxPooling2D | 256 | relu |
| Conv2D, MaxPooling2D | 128 | relu |
| Conv2D, MaxPooling2D | 64 | relu |
| Conv2D, MaxPooling2D | 32 | relu |
| Flatten, Dropout(0.5) | — | — |
| Dense | 64 | relu |
| Dense | 1 | sigmoid |

2.4 실험 결과

(1) yawn_no_yawn 모델에 대한 OpenCV와 Dlib의 정확도 비교

구축한 (OpenCV)yawn_no_yawn 모델과 (Dlib)yawn_no_yawn 모델의 학습 및 검증 데이터 세트의 정확도는 그림 1과 같다. (OpenCV)yawn_no_yawn 모델의 검증 정확도는 74%이며, (Dlib)yawn_no_yawn 모델의 검증 정확도는 93%로 산출되었다.

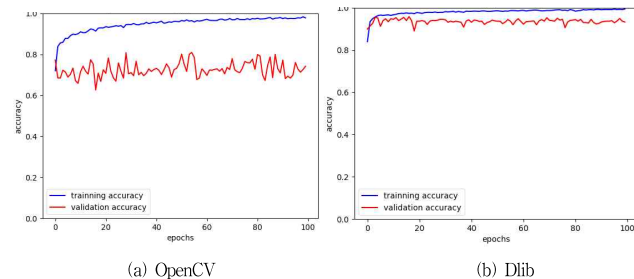


Fig. 1 yawn_no_yawn 모델의 OpenCV와 Dlib의 정확도 비교

(OpenCV)yawn_no_yawn 모델과 (Dlib)yawn_no_yawn 모델을 이용하여 테스트 데이터 세트의 하품(yawn), 입 닫음(no_yawn) 이미지 200장을 테스트하였을 때, yawn이지만 no_yawn으로 판별하는 오류(yawn error)와 no_yawn이지만 yawn으로 판별하는 오류(no_yawn error)를 비교한 결과는 표 2와 같다. OpenCV의 yawn error는 입을 감지하지 못하고 눈을 감지하는 경우가 많아 86%로, no_yawn error는 6% 산출되었다. Dlib의 yawn error는 0%로, no_yawn error는 5% 산출되었다. 검증 정확도와 테스트 데이터 세트의 오류 테스트 결과를 보았을 때 하품 감지 측면에서 Dlib이 OpenCV보다 우수한 정확도를 보였다.

Table 2 yawn_no_yawn 모델에 대한 오류율 비교

| | yawn error | no_yawn error |
|--------|------------|---------------|
| OpenCV | 86% | 6% |
| Dlib | 0% | 5% |

(2) open_closed 모델에 대한 OpenCV와 Dlib의 정확도 비교

구축한 (OpenCV)open_closed모델과 (Dlib)open_closed모델의 학습 및 검증 데이터 세트의 정확도는 그림 2와 같다. (OpenCV)open_closed모델의 검증 정확도는 82%로 산출되었으며, (Dlib)open_closed모델의 검증 정

확도는 90%로 산출되었다.

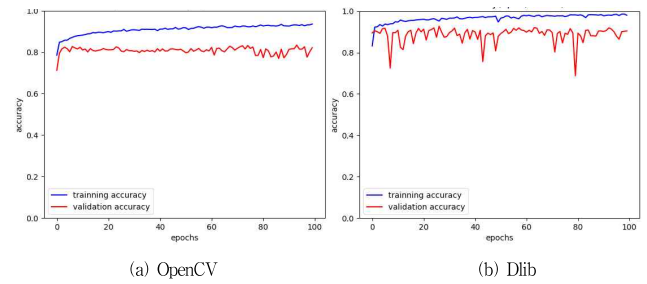


Fig. 2 open_closed 모델의 OpenCV와 Dlib의 정확도 비교

(OpenCV)open_closed모델과 (Dlib)open_closed모델을 이용하여 테스트 데이터 세트의 눈 뜸(open), 눈 감음(closed) 이미지 200장을 테스트하였을 때, open이지만 closed로 판별하는 오류(open error)와 closed이지만 open으로 판별하는 오류(closed error)를 비교한 결과는 표 3과 같다. OpenCV의 open error는 20%로 산출되었으며, closed error는 눈을 감지 못하고 콧구멍을 감지하는 경우가 많아 47%로 산출되었다. Dlib의 open error는 1%로 산출되었으며, closed error는 11%로 산출되었다. 검증 정확도와 테스트 데이터 세트의 오류 테스트 결과를 보았을 때 눈 깜박임 감지 측면에서 Dlib이 OpenCV보다 우수한 정확도를 보였다.

Table 3 open_closed 모델에 대한 오류율 비교

| | open error | closed error |
|--------|------------|--------------|
| OpenCV | 20% | 47% |
| Dlib | 1% | 11% |

III. 결론

본 논문은 동일한 데이터 세트와 학습 모델을 사용하였을 때 OpenCV와 Dlib으로 전처리한 모델 중 어느 모델이 졸음운전 탐지를 위한 이미지 판별 정확도가 우수한지를 비교하였다. 데이터 세트는 정면, 측면 이미지를 사용하였고 이미지 전처리 부분 이외의 코드는 모두 동일하게 구현하였다. 실험 결과 하품 감지와 눈 깜박임 감지 측면 모두에서 Dlib이 OpenCV보다 높은 정확도를 나타내었다. 향후 졸음운전 경고 시스템을 개발함에 있어서 Dlib을 활용할 경우 졸음이 경고되는 상황을 인식하지 못하는 오류는 적어질 것이며 더 높은 정확도로 졸음을 경고할 수 있을 것이다. 다만, 정확도는 인공지능 알고리즘의 여러 평가 지표 중 하나임으로 향후 다른 지표를 이용한 성능 비교 연구가 진행될 필요가 있다.

참 고 문 헌

- [1] Suryaprasad, J., et al. "Real time drowsy driver detection using haarcascade samples." CS & IT Conference Proceedings. Vol. 3. No. 8. CS & IT Conference Proceedings, 2013.
- [2] Babu, Athira, Shruti Nair, and K. Sree Kumar. "Driver's drowsiness detection system using Dlib HOG." Ubiquitous Intelligent Systems: Proceedings of ICUIS 2021. Springer Singapore, 2022.
- [3] "졸음운전 예방을 위한 운전자 상태 정보 영상." AI-Hub, n.d., aihub.or.kr/aihubdata/data/view.do?currMenu=&topMenu=&aihubDataSe=realm&dataSetSn=173.