# 사람의 움직임 감지를 측정한 학습 능률 확인 시스템

김석현, 이진성, 유은상, 박선우, 김응태
한국산업기술대학교 전자공학부
{josephsh97, asdfppp, rhymus314, dragon61012, etkim} @kpu.ac.kr

# Learning efficiency checking system by measuring human motion detection

Sukhyun Kim, Jinsung Lee, Eunsang Yu, Seon-u Park, Eung-Tae Kim
Dept. of Electronics, Korea Polytechnic University

## 요 약

본 논문에서는 공부하는 사용자의 상황을 감지하여, 학습의욕을 고취시키고 집중력 향상을 도와주기 위한 학습능률 확인 시스템을 구현하고자 한다. 이를 위해 실시간 카메라를 통해 사용자의 얼굴이나 몸의 움직임을 추출하여 학습 태도, 집중력에 대한 데이터를 측정한다. 실시간 임베디드 시스템 구현을 위해 Jetson 보드를 사용하였으며, 영상인식을 위한 CNN(Convolution Neural Network)를 구현하였다. CNN 을 이용해 대상의 특징 부분을 검출한 후 움직임 검파를 수행한다. 캡처한 영상을 PYQT5 로 작성된 GUI 에서 영상을 보여주며, 각각 방해되는 행동을 했을 때 푸시메시지를 보내며 데이터를 수집한다. 또한 GUI 로 만든 메인 화면에서 각각의 기능들을 실행 가능하며, 수집한 데이터를 산출해주는 통계그래프와 작업관리 목록, 화이트 노이즈 등의 기능을 수행한다. 구축된 학습능률 확인 시스템을 통해 대상의 데이터를 수집 및 분석을 비롯한 다양한 기능을 사용자에게 제공하였다.

**키워드**: 키워드: 딥러닝, CNN, OpenPose, 학습 능률

## Abstract

In this paper, we implement a learning efficiency verification system to inspire learning motivation and help improve concentration by detecting the situation of the user studying. To this aim, data on learning attitude and concentration are measured by extracting the movement of the user's face or body through a real-time camera. The Jetson board was used to implement the real-time embedded system, and a convolutional neural network (CNN) was implemented for image recognition. After detecting the feature part of the object using a CNN, motion detection is performed. The captured image is shown in a GUI written in PYQT5, and data is collected by sending push messages when each of the actions is obstructed. In addition, each function can be executed on the main screen made with the GUI, and functions such as a statistical graph that calculates the collected data, To do list, and white noise are performed. Through learning efficiency checking system, various functions including data collection and analysis of targets were provided to users.

**Keyword**: DeepLearning, CNN, OpenPose, Learning efficiency

## 1. 서론

최근 들어 COVID- 19 로 인한 비대면 상황에 따라 가정학습 시간 증대와 같은 사회적 이슈와 더불어 자기 주도 학습이 증가하고 있다, 이에 따라 실시간 영상을 이용한 Zoom , 각종 매체의 강의, 유튜브 등의 동영상 플랫폼이 대두되고 있다. 하지만 그와 반면에 학습자에게 직접적으로 자기 학습관리를 도와주는 방법들은 아직 부족하다.

KOSIS 국가 통계포털[1]에 따르면 2015 년 부터 현재까지 사교육비는 2015 년 178,346 억원에서 2016 년 180,606 억원 등 2019 년에는 209,970 억원에 달하여 지속적으로 증가하는 추세가 되었으며 이에 따라 가정환경과 여러 요인들로 인한 학습 격차가 증가될 수 있다는 것이 큰 문제로 보여진다.

이처럼 사교육 문제와 더불어 현재의 Covid-19 상황에서의 학습격차와 그에 따른 사회적 교육 이슈를 해소하고자 한다. 요즘 대두되는 실시간 동영상의 형태를 활용하여 사용자로 하여금 학습 능률을 측정하고 이에 따른 방해 요소를 그래프화 하여 체크하고 학습에 도움이 되는 부분들을 통해 스스로 쉽게 취약점을 보완하고 학습 능률을 올리도록 돕는다.

따라서 본 논문에서는 실시간 동영상으로 촬영하는 카메라에 사용자의 얼굴과 몸의 움직임을 Convolution Neural Network (CNN) 기법을 활용하여 이를 통해 사용자의 학습 간에 실시간 데이터를 추출한다. CNN 기법을 통해 영상에서 이미지를 인식하여 패턴을 분류하는데 사용한다. CNN 은 이미지에서의 공간정보를 유지한 채 자동으로 특징을 추출해내는 장점을 가지고 있다.

이와함께 오픈소스인 OpenPose 를 사용하는데 OpenPose 는 신체 부분들을 키포인트로 하여 인식 할 수 있는 모델이다. 이를 통해 카메라로부터 입력된 이미지 또는 비디오로부터 신체의 마디와 같은 키포인트들을 추출하여 배경 이미지와 키포인트가 합쳐진 이미지를 영상데이터로 저장할 수 있다[2]. 여기서 학습적인 집중을 저하하는 행동 5 가지(목이나 입 만짐, 왼쪽 눈 만짐, 오른쪽 눈 만짐, 오른쪽 어깨만짐, 왼쪽 어깨만짐)로 선정 후 OpenPose 와 CNN 을 기반으로 서비스를 개발했다.

이때 모션인식은 영상의 프레임에서 이전 프레임과 지금 프레임 간의 차이는 모션에 의해서 생겼다고 가정한다. 이 영상의 차이가 모션에 의해서만 발생하는 것이 아니라 외부환경의 변화가 모션으로 감지되는 경우가 있다. 따라서 전체 화면이 아닌 인체의 특정 부분만을 통해 감지하는 식으로 처리하도록 하며, 움직임에 따라 어떤 모션인지 판별해야 하므로 영상에서 사람의 위치와 각도도 중요하게 다뤄진다.

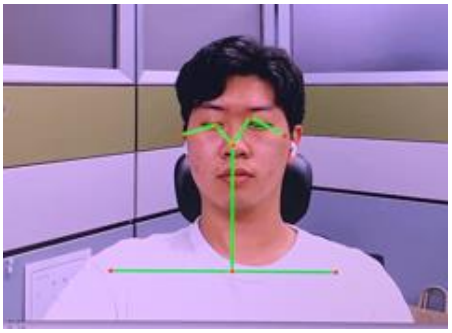
따라서, 이러한 방법을 통해 본 논문에서는 PYQT5 를 통해 GUI 로 만든 메인 화면에서 각각의 기능들을 실행 가능하도록 한다. 또한 모션인식을 통해 수집한 데이터를 산출해주는 통계 그래프와 To Do List, 화이트 노이즈 등의 기능을 수행하여 실질적인 학습 시간을 정확히 파악하고 학습능률을 향상시켜 교육 이슈를 줄이고자 한다.

## 2. 제안하는 학습집중 통계 서비스

학습용 데이터는 eye Blink Detector[3]에 공개된 눈 깜박임 데이터 셋을 사용하려 하였으나, 한쪽 눈 하나당 420 개의 데이 수가 있었으나 아주 어두운 환경과 매우 밝은 환경에서의 측정한 사진을 추가하여 80 개의 이미지 데이터를 직접 촬영하여 추가 하였습니다. 학습자가 학습 중간에 눈을 감은 일정시간 계속 감고 있을 경우 졸음으로 판별하였다. 졸음을 판별하기 위해 shape_predictor_68_face_landmarks.dat [4]데이터를 이용해 사용자의 눈의 위치를 추출하여 각각 500 개의 데이터를 학습시켰다. 판별하기 위해 ear 알고리즘을 이용하여 눈의 크기가 특정비율이 되면 졸음을 판별하도록 구별하였다. 눈의 크기가 특정비율이 되면 졸음으로 판별하도록 구현하였다. 또한 하품은 윗입술과 아랫입술의 간격이 특정 간격일 경우로 측정하여 추출하였다. 총데이터는 그림 1 에서 확인할 수 있다.

**》그림 1. Face_landmarks 를 통한 눈깜빡임과 하품 데이터추출**

또다른 학습용 데이터를 OpenPose[5]를 이용해 상체부위의 데이터를 사용하기 위해 정면에서 50 개, 측면에서 50 개씩 집중도 저하의 행동을 직접 영상의 데이터를 촬영했다. 또한 상체를 측정하기 위해 OpenPose 에서 중요한부분을 뽑았다. 그림 2 처럼 상체에서 필요한 부위만 추출하기 위해 5 개의 데이터만 사용했다.



》그림 2. OpenPose 에서 추출한 중요한 부분 예

촬영되는 카메라 영상에서 1500 개의 데이터를 쉽게 다루기 위해 줄여야했다. 추출과정에서 배경부부의 잡음을 제거하기 위해 가우시안 블러링을 수행했다. 그리고 카메라의 화질은 640*480 으로 고정했다. 촬영하는 카메라 영상의 프레임은 10,30 으로 나누어서 추출했다.

본 논문에서 인식이 좋고 정확한 모델을 사용하기 위해 모델을 나누어 보았다. 최적의 모델을 아래 표 1 과 같이 나타냈다. 각각 어떠한 방법으로 데이터를 추출하여 프레임별로 모델을 제안하였으며, 성능평가를 통해 model_ C 가 우수한 성능을 보였음을 확인했다.

표 1. 모델

| | | |
|---|---|---|
| model_A | face_detector,body_detector | frame10 |
| model_B | face_detector,body_detector | frame30 |
| model_C | face_detector | frame10 |
| model_D | face_detector | frame30 |
| model_E | body_detector | frame10 |
| model_F | body_detector | frame30 |

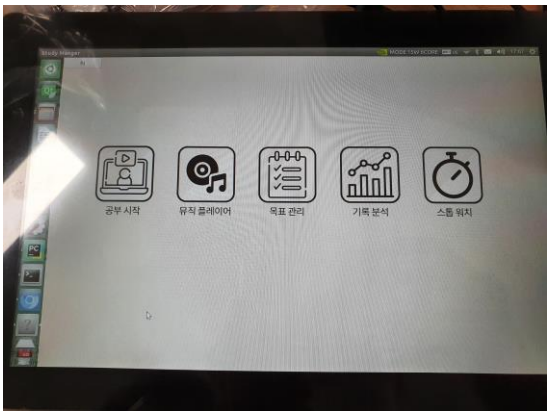
## 3. 성능평가

표 3 은 논문에서 보여준 6 가지 모델의 성능을 보인다. 데이터 셋은 총 1500 개로 각각 500 개씩 테스트를 진행했다. 모델별로 50 차례씩 시도하며 10 분이라는 일정시간동안 테스트하여 나온 결과들을 정리해보았다. model_C 가 얼굴로서 학습자의 집중도를 판단하는 정확도나 인식률이 가장 높게 보여 최종 모델로 선택했다.

그림 3 은 이 시스템을 구현하기 위한 jetson xavier NX 리눅스환경에서 PyQT5 를 통해 GUI 로 사용한 그림이며 개발환경은 표 3 에서 보이고 있다.

표 2. 모델과 통계로서 나온 정확도와 인식률

| 모델 | Recognition rate | Accuracy |
|---|---|---|
| model_A | 0.93 | 0.81 |
| model_B | 0.88 | 0.78 |
| model_C | 0.95 | 0.91 |
| model_D | 0.86 | 0.89 |
| model_E | 0.97 | 0.73 |
| model_F | 0.86 | 0.68 |




》그림 3. Jetson xavier NX 와 연결된 터치 디스플레이에서의 GUI

표 3.개발환경

| 규격 | 종류 혹은 버전 |
|---|---|
| board | jetson xavier NX |
| linux | ubuntu 18.04 LTS |
| python | python3.7 |
| GUI | PYQT5 |
| tensorflow | v.2.0 |
| camera | 1080p HD |
| screen | 10.1in touch LCD |

## 4.결론

본 논문에서는 실시간 동영상을 촬영하는 카메라를 통해 실시간으로 사람의 움직임을 감지하도록 하였다. 이를 위해서 딥러닝의 CNN 모델 학습과 OpenPose 기술을 활용하면서 어두운 환경, 너무 밝은 환경 등 다양한 환경에서의 데이터를 학습하도록 구현하여 외부환경 변화에서 민감하게 작용했던 정확도와 인식율 부분을 해결했다.

또한 학습자의 눈깜빡임, 하품 등의 여러가지 행동을 감지하여 실시간으로 체크하도록 하여 이러한 집중력이 흐트러지는 행동을 할 경우 바로 체크가 가능하도록 화면에서 푸시메세지를 띄워 가시성을 더했다.

따라서, 본 논문에서는 공부하는 학습자의 얼굴상황, 태도 등을 영상 인식하여 학습자에게 학습데이터, 시간 등 필요한 정보를 제공할 수 있는 임베디드 기반 학습능률 확인 시스템을 구현하였다. 실시간 영상인식 및 데이터 추출을 위해 CNN 기반 영상 처리를 Jatson 보드에 구현하고 성능을 평가하였다.

이를 활용하여 향후 연구에서는 OpenPose 를 통해 키포인트를 더 늘리고 모델을 더 세밀하게 나누어 다양한 환경에서 데이터를 학습하도록 하여 정확도와 인식률을 높이면 에러율을 더 줄일 수 있으리라 사료된다. 더 나아가 손동작이나 자세 등에도 적용하게 된다면 더욱 다양한 분야에서도 적용이 가능하며 차량의 졸음 운전 등의 여타의 상황에서도 좋은 기능을 수행할 수 있을 것으로 본다.

## 참고문헌


[1]https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1PE003&vw_cd=MT_ZTITLE&list_id=H1_10_001_001&seqNo=&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE

[2] 장재호, 지준환, 김두환, 최민기, 윤태진. (2020). 인체 자세 인식 딥러닝을 이용한 운동 자세 훈련 시스템 개발. 한국컴퓨터정보학회 학술발표논문집 , 28(2), 289-290.

[3]https://github.com/kairess/eye_blink_detector

[4]류세열,유재천.(2021). 딥러닝 다중 네트워크를 이용한 졸음 운전감지 및 안전벨트 착용 여부 확인 29{1),75-77

[5] https://github.com/CMU-Perceptual-Computing-Lab/openpose_train