SWIM

(Study With Me)

AI 비대면 독서실 어플리케이션

*Posco AI, Big data Academy 9기 A반 3조*

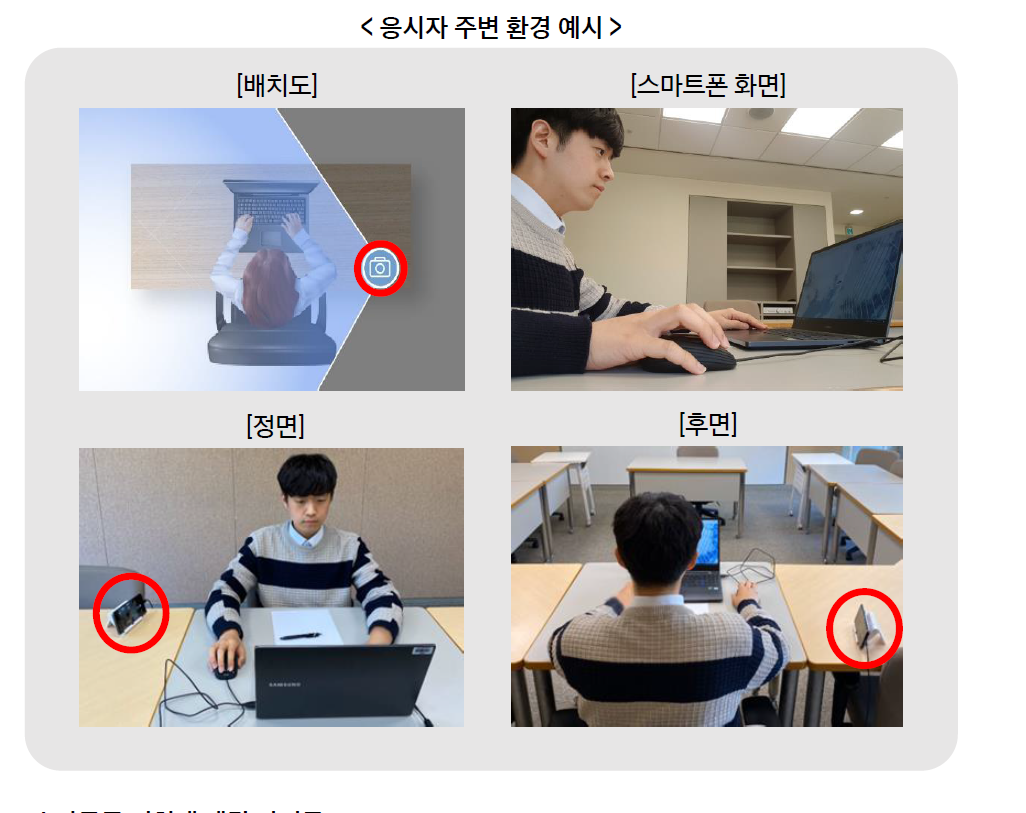
김진명, 김채은, 김한빈, 이경원, 이다연, 이상엽

**[CONTENTS]**

1. 프로젝트 개요
2. 프로젝트 소개
3. 주제 선정배경
4. 유사 기술 현황
5. 진행 과정
6. 기대 효과
7. 구현 프로세스
8. Openpose
9. 모델 생성
10. 임계치 설정
11. Pose Estimation test
12. 어플리케이션
13. 구현 결과
14. Feature extraction
15. Studying Classification
16. 어플리케이션
17. 보완사항
18. 조원 소개 및 평가
19. 참고문헌
20. **프로젝트 개요**
21. 프로젝트 소개

SWIM(Study With Me)는 Pose Estimation을 활용하여 공부 시간을 측정하는 ‘AI 비대면 독서실 어플리케이션’입니다. 학습자의 자세를 실시간으로 분석하여 분석 결과를 기반으로 공부 시간을 기록합니다.

이 어플리케이션은 Pose estimation 기술을 활용하여 학습할 때의 어깨골격과 손목을 detection합니다. 움직임이 많지 않은 어깨골격과 움직임이 많은 손목을 인식하여 학습자세 여부를 분류합니다. 이 때 카메라는 이용자의 어깨와 손목을 인식할 수 있도록 정면에서 양 측면까지의 범위 내에 위치해 있어야 합니다. 이 후 스마트폰 카메라로 학습자를 영상으로 촬영하면, 현재 학습자의 자세가 공부중인 자세인지를 분석하고 분석결과를 바탕으로 공부 시간을 기록하는 프로세스입니다. 또한 기록 통계를 관리하고 공유할 수 있도록 어플리케이션을 제작했습니다.



1. 주제선정배경

2-1) 코로나19로 달라진 학습 환경

최근 전세계를 강타한 코로나19 여파로 사회적 거리두기와 같은 생활 속 방역을 시행 중에 있습니다. 집단문화를 자제하는 시국으로 인해 사회 곳곳에서 변화가 감지되고 있습니다. 학생들에게 끼치는 영향 또한 예외가 아닙니다. 유수 대학들의 대면강의가 온라인 강의로 대체되면서 학생들이 카페로 삼삼오오 모여드는 ‘카공족’을 심심치 않게 볼 수 있습니다. 취업준비생들도 상황이 다르지 않습니다. 하지만 카페는 코로나19 감염에 취약하다는 지적이 있습니다. 또한 대부분의 시립도서관 열람실은 2월 이후 폐쇄되었고 거리두기가 시행되며 스터디를 진행하기도 쉽지 않습니다. 비대면 스터디도 있기는 하나 제한사항이 많은 실정입니다.

2-2) 현 상황의 한계점 및 극복 방안

카페는 코로나19 감염에 취약하며, 학습할 목적으로 설립한 사업체가 아니기 때문에 학습에 차질이 생길 수 있는 다양한 요소가 존재합니다. 하지만 자택에서 공부하는 환경은 학습 집중력이 저하되는 명확한 한계점이 있습니다. 자택은 PC, TV, 게임기, 가족 등 다양한 학습 방해요소가 존재합니다. 즉, 학습 분위기 조성 미흡으로 인한 집중력 저하가 크게 우려되는 실정입니다.

따라서 학습분위기를 조성하여 학생들의 학습 집중도를 향상시키고, 더불어 학습 집중도 데이터를 활용하여 체계적인 학습 전략을 수립하고자 하였습니다. 또한 스마트폰만 있다면 어느 곳이든 독서실과 동일한 학습환경을 제공할 수 있는 어플리케이션을 제작하기 위해 SWIM 프로젝트를 진행하였습니다.

1. 유사 기술 현황

3-1) 스마트홈트

Kakao VX Corps에서 개발한 AI 코칭 홈트레이닝 전문 어플리케이션입니다. 이 서비스는 체계적인 피트니스 커리큘럼에 AI 코칭을 접목하였습니다. 딥러닝 기반 AI 기술을 활용하여 이용자의 실시간 관절 움직임을 detection하고 분석 후 올바른 운동 자세를 추천하는 서비스입니다.

운동이 시작되면 이용자의 운동 자세를 전문 트레이너와 동작 일치여부를 표시합니다. 일치한다고 판단하면 녹색으로, 일치하지 않는다고 판단하면 빨간색으로 표시하는 점이 특징입니다. 또한 운동이 시작되면 이용자의 운동 자세를 트레이너 자세와 비교하면서 운동할 때 주의할 사항들을 알려줍니다.

3-2) 12PT

컨티넘에서 출시한 AI 홈트레이닝 어플리케이션입니다. 스마트폰 카메라를 이용해 사용자의 운동 자세 등 정보를 실시간으로 분석해 운동 횟수를 알려주고, 올바른 자세를 유도합니다. 12PT는 딥러닝 프레임워크인 텐서플로와 NVDIA cuDNN 라이브러리를 사용해 사람의 관절을 실시간으로 추정하는 기술을 사용합니다. 추정된 사람 관절을 다시 딥러닝으로 학습해 운동 패턴을 분석합니다. 이러한 분석을 바탕으로 운동 횟수를 알려주고 올바른 자세를 유도하는 프로세스입니다.

또한 사용자 개개인이 본인에게 적합한 운동 횟수와 종류를 직접 설정할 수 있습니다. 점핑잭, 스쿼트, 런지, 푸시업, 점프 스쿼트, 스탠딩 사이드 크런치, 스탠딩 크로스 크런치, 굿모닝, 핑거팁 투 토우 잭, 하이니 등 총 10가지 운동을 지원합니다. 또한 easy, normal, hard로 구성한 운동강도를 기본으로 추천하는 동시에, 사용자가 원하는 운동 횟수와 순서도 직접 편집할 수 있습니다.

3-3) 결론 및 참고점

상기의 2가지 서비스는 딥러닝 기술을 활용하여 이용자의 골격의 움직임을 detection하고 올바른 운동자세를 판단한다는 공통점이 있습니다. 본 프로젝트도 딥러닝 기반의 기술을 활용하며, 어깨 골격과 손목을 detection하여 학습중인 자세 여부를 판단한다는 점에서 상기의 서비스와 유사합니다. 또한 어플리케이션을 사용하여 실시간으로 자세를 분석한다는 면에서도 본 프로젝트와 유사합니다. 이러한 유사성을 참고하여 진행하면 프로젝트의 완성도를 높일 수 있겠다고 판단하였습니다.



1. 진행과정

|  |  |  |  |
| --- | --- | --- | --- |
| **주차** | **일정** | **순서** | **내용** |
| 1주차 | 2020.5.11  ~2020.5.17 | 주제 최종 선정 | * 주제 선정을 위한 자료 탐색 * 전체 일정 논의 * 관련 자료 검색 및 분석 |
| 2주차 | 2020.5.18  ~2020.5.24 | 아이디어 검토 및 개선 | * 아이디어 발전 가능성 검토(앱 개발 등) * 이전 기수 프로젝트 참고 |
| 알고리즘 관련 논문 및  참고문헌 조사 | * <패턴인식과 머신러닝>, <OpenCV4로 배우는 컴퓨터비전과 머신러닝> 등 |
| 3주차 | 2020.5.25  ~2020.5.31 | 학습 모델 구현 | * Openpose 모델 선정 * 선정 모델로 Pose estimation |
| 임계치 설정 | * Shoulder line과 Wrist 임계치 설정 |
| 적합한 모델 후보군 선정 후 테스트 | * 모델 구현 후 결과 분석 * 결과 향상 작업 * Studying과 non studying의 이진분류 모델 생성 |
| 4주차 | 2020.6.1  ~2020.6.5 | 어플리케이션 구축 | * 어플리케이션 설계 * 적합한 모델로 환경 설정 |
| 최종 발표 | * PPT 최종 발표 * 시연 영상 공개 |
| 보고서 작성 | * 최종 보고서 작성 |

1. 기대 효과

5-1) 사용자의 학습 집중도 향상

학생들이 공부한 시간을 체크할 때 집중을 하지 않거나 자리를 비워둔 시간도 총량적으로 공부시간으로 체크하는 경향이 있어 실질적인 공부 시간이 체크하기에 제한이 있습니다. 하지만 SWIM 어플리케이션은 사용자의 자세를 인식하기 때문에 실질적인 공부 시간의 자동 측정이 가능할 것으로 기대합니다.

이와 더불어 공부 시간의 측정을 통하여 학습 집중도 데이터를 축적시킬 수 있습니다. 이러한 데이터를 활용하여 체계적인 학습 전략 수립이 가능합니다. 예를 들어, 가장 집중력이 높게 유지되는 시간에 가장 우선순위가 높은 과목을 공부할 수 있도록 전략을 수립할 수 있습니다. 이 뿐만 아니라, 어플리케이션 사용 중에는 스마트폰을 사용할 수 없으므로 학습 중 스마트폰 사용을 억제하는 효과도 기대할 수 있습니다.

5-2) 새로운 학습 플랫폼

사용자 개인적인 용도 뿐 아니라 플랫폼 용도로도 활용할 수 있는 가능성이 기대됩니다. 코로나19로 인해 비대면 스터디가 활발해지고 있는 만큼 이에도 활용할 수 있을 것이라 생각됩니다. 스터디원들과의 집중시간 공유를 통해 언택트 스터디 플랫폼을 공유할 수 있을 것으로 기대됩니다. 또한 단순한 스터디를 넘어서 입시교육에서도 활용 가능할 것으로 기대됩니다. 학교, 학원 등의 교육현장과 학부모가 학생의 학습시간 데이터를 공유하고, 이를 통해 종합 입시 솔루션도 가능할 것이라 예상됩니다.

더 나아가 온라인 테스트 시 컨닝방지 목적으로도 활용할 수 있을 것이라 기대됩니다. 최근 삼성 인적성 검사인 GSAT에서 온라인 테스트가 도입되었는데 컨닝을 방지할 수 있는지에 대한 여부가 관심사였습니다. 본 프로젝트에서 제작한 어플리케이션은 자세를 인식하여 학습 여부를 판단할 수 있기 때문에 컨닝 방지 어플리케이션으로 발전할 수 있을 것이라 기대됩니다.

**2. 구현 프로세스**

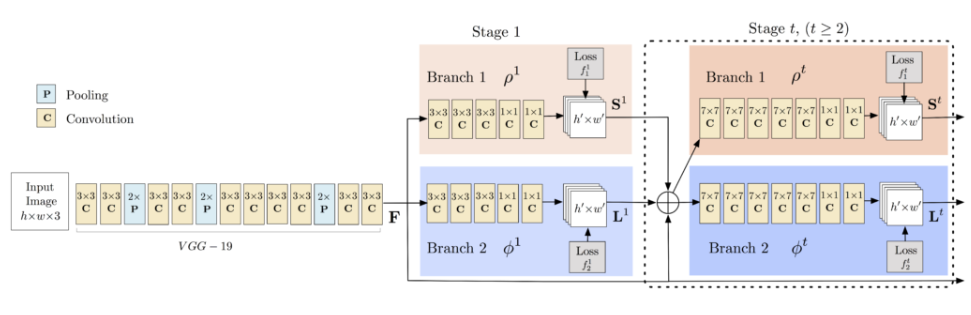
1) Openpose

1-1) 개요

Openpose는 CNN 기반의 딥러닝 기술로 이미지 혹은 영상을 input으로 하여 사람의 움직임을 감지합니다. Caffe와 OpenCV를 기반으로 구성된 손, 얼굴 등을 포함한 몸의 움직임을 추적해주는 API라 할 수 있습니다.

Openpose는 센서 필요없이 일반 카메라로 이미지와 동영상을 추출할 수 있고, 이미지 내 관절을 먼저 파악하고 자세를 추정하는 Bottom-up 방식이며, 추정한 자세로 사람의 골격인 Skeleton을 추출한다는 특징이 있습니다.

1-2) 작동원리



1-2.1) 좌측단의 입력 데이터를 VGGNet-19를 통해 수행된 Output 데이터의 feature를 강조한 상태로 출력합니다.

1-2.2) 출력된 output을 2번의 Branch를 통해 전파를 수행하게 되는데, 첫번째 분기점에서는 전반적인 신체 부위의 특정에 사용됩니다.

1-3.3) 반복되는 Stage에 따라 가지를 거쳐서 confidence map과 affinity field를 구하게 됩니다. 여기서 confidence map이란 인간의 각 부위 간 유사성에 기반하여 2D vector fields를 예측합니다. 이는 인간의 관절 구조 등을 찾는데 사용됩니다. 또한 affinity field는 추출된 관절 구조가 어떤 객체의 것인가에 대해 알아내는데 사용됩니다.

1-3.4) 연산된 confidence map과 affinity field로 영상 혹은 이미지에 존재하는 사람들의 keypoint를 예측하고 생성합니다. 이러한 1번부터 3번까지의 과정을 반복하여 학습하면 Openpose가 완성됩니다.

아래는 기학습되어 있는 모델을 OpenCV로 로드하여 구현한 결과입니다.



1-3) 적용 방안

대부분의 Openpose 모델은 전신을 대상으로 한 모델이었습니다. 하지만 본 프로젝트는 학습중인 자세를 인식해야 함으로 하반신을 인식할 수 없는 영상에는 큰 오차가 존재했습니다. 따라서 위와 같은 전신을 대상으로 한 모델은 적합하지 않았습니다.

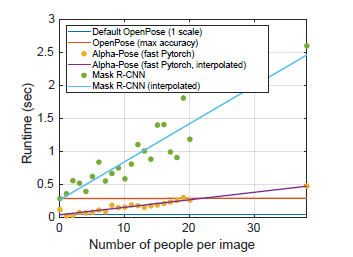
따라서 상반신만 인식하고자 하였고 그 중에서도 Shoulder line과 Wrist를 활용하는 모델을 생성하였습니다. 일반적으로 학습중일 때는 어깨가 크게 움직이지 않으며 손목은 유동적으로 움직이게 됩니다. 이러한 학습 중의 특징을 활용하였습니다. 어깨 Shoulder line과 Wrist를 인식하는 Openpose 모델은 추후 Shoulder line이 벗어나는 정도와 Wrist가 움직이는 정도에 따라 학습 여부를 판단할 수 있는 분류모델을 생성하는 데 활용하였습니다.

2) 모델 생성

2-1) 모델 선정 이유

\*[그림1], [그림2]는 "OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields." 논문에서 인용하였습니다.

Multi-person pose estimation model 중에서 가장 유명한 모델이 Openpose와 Alphapose 그리고 Mask R-CNN이 있습니다. 그 중에서 아래 [그림1]을 보시면 Openpose모델이 하나의 이미지에서 사람의 수가 늘어나도 실행속도는 낮아지는 것을 알 수 있습니다. 그래서 저희는 Frame안에 많은 사람들이 들어와도 감지를 빠른 속도에서 할 수 있는 Openpose모델을 최종적으로 선택하였습니다.

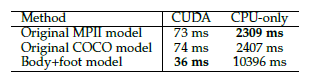


[그림 1]

2-2) BODY+foot vs. COCO vs. MPII Model 비교

Mp4형식의 공부 동영상을 Openpose모델에 input으로 넣어주어 output으로 주요 신체부위의 좌표값은 JSON파일의 형식으로 Skeleton의 영상은 AVI형식으로 나오게 됩니다.

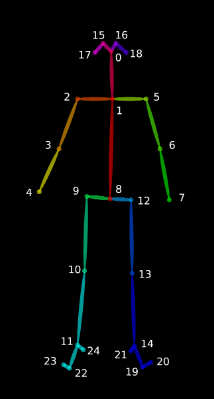
그 중 신체 특징을 뽑아내는 모델은 다음과 같이 3가지가 있습니다.



[그림2]

[그림2]를 보시면 MPII, COCO, Body+foot에서 CUDA와 CPU 환경일 때 각각 모델의 실행속도를 나타낸 표입니다. 보시다시피 Body+foot모델 즉, 신체와 발을 포함해 detection하는 모델이 나머지 두 모델들보다 속도가 약 2배정도 빠른 것을 확인할 수 있습니다. 그러한 이유로 저희는 CUDA환경내의 Body+foot모델을 최종적으로 선정하게 되었습니다.

2-3) Body Keypoints 추출

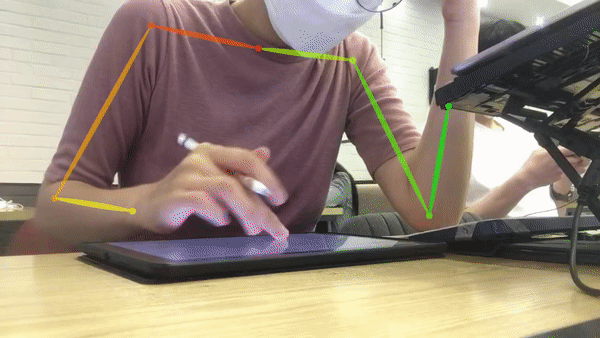
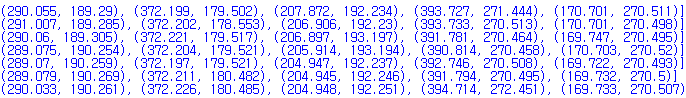


[그림3]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0 :Nose | 1 :Neck | 2 :RShoulder | 3 :RElbow | 4 :RWrist |
| 5 :LShoulder | **6 :LElbow** | **7 :LWrist** | **8 :MidHip** | **9 :RHip** |
| 10 :RKnee | **11 :Rankle** | **12 :LHip** | **13 :LKnee** | **14 :LAnkle** |
| 15 :REye | **16 :LEye** | **17 :REar** | **18 :LEar** | **19 :LBigToe** |
| 20 :LSmallToe | **21 :LHeel** | **22 :RBigToe** | **23 :RSmallToe** | **24 :RHeel** |

[그림4]

앞에서 언급했듯이 body+foot모델을 사용했고 그 구조와 신체의 keypoints들은 [그림3] [그림4]에서 볼 수 있습니다. 여기서 저희는 상반신 모델만을 이용해 공부 자세를 정의하기 때문에 필요한 keypoints의 좌표값들만 추출하였습니다.

[그림5] [그림6]

[그림5]에서 보이듯이 차례로 각각 왼쪽, 오른쪽 어깨와 목 그리고 왼쪽, 오른쪽 손목을 frame에서 위치 좌표값(x,y)의 값으로 데이터를 뽑아 [그림6]과 같이 리스트안에 튜플이 들어간 형식으로 가공해주었습니다. 그리하여 매 frame마다 신체의 위치를 나타내는 좌표값들이 데이터로 생성되고, 그 값들을 활용하여 공부 자세를 추정하는 모델을 생성하고자 하였습니다. 하지만 그러기 위해서는 공부자세를 정의하고 기준에 맞는 임계치들을 정의하는 과정이 필요했습니다. 그 부분은 아래에서 설명하겠습니다.

**5. 조원 평가**

1. 김채은: 뛰어난 코딩 실력으로 항상 캐리해줬던 짝찌였습니다. 이번 인공지능 프로젝트에서도 Object detection과 어플리케이션, 기획 등등 다방면에서 활동해줘서 프로젝트가 더욱 원활하게 진행될 수 있었습니다. SSAFY에서도 많이 배워서 훌륭한 ICT 인재가 되길 바랍니다.

2. 김한빈: 프로젝트를 할 때마다 방향을 잘 잡아주었습니다. 임계치 설정하고 학습중인 자세에 대한 기준을 선정하는 데에 큰 공이 있었다고 생각습니다. 최종 발표 준비하느라 너무 고생 많았고 잠재력이 많이 보이는 교육생으로 더 깊이 학문에 몰두하려는 계획에 차질이 없길 기원합니다.

3. 이경원: 우리 팀의 PPT 마스터입니다. 인공지능 뿐만 아니라 빅데이터 프로젝트에서도 PPT를 만들어서 큰 힘이 되었습니다. 특히 인공지능 과정에서는 설정한 임계치로 Pose estimation을 구현하는데 크게 기여하였습니다. 다방면으로 유능하고 나날이 노력하며 성장하는 교육생이라 앞으로가 더욱 기대됩니다.

4. 이다연: Object detection 모델을 수립하는데 정말 지대한 공을 세웠습니다. 모델을 구현하느라 밤을 새워가며 노력했는데 그 결과가 잘 나타나서 다행입니다. 팀원들과의 커뮤니케이션 능력도 매우 좋으며 어떤 분야로 진출하던 역량을 십분 발휘하는 인재라 생각합니다.

5. 이상엽: 우리팀의 유일한 컴퓨터 공학과 출신으로써 어플리케이션을 만드는데 매우 큰 역할을 하였습니다. 자칫 모델 수립이나 아이디어 제안으로 끝날 수 있었던 프로젝트에 활력을 불어넣어준 교육생이었습니다. 또한 상엽이가 주재한 딥러닝 스터디는 프로젝트를 더 수월하게 진행시켜 주었습니다. 컴퓨터 관련 분야에서 지식도 많으며 컴퓨터활용 능력도 매우 높은 교육생입니다. IT 분야 어디를 가든 제 몫을 다할 인재입니다.