

Project Proposal

헬린이의 득근득근 성장일기

- 팀명 : 러닝 머신
- 팀장 : 정창욱[201820713]
- 팀원 : 김문만[201820744], 서재명[201720724], 이영민[201820729], 이준수[201820704]

• Intro

2022 트렌드에 대해 찾아보면 바른생활 루틴, 헬스 플레저 등 자기 관리와 관련한 많은 단어들을 어렵지 않게 찾을 수 있다. 또한 최근 10년간 피트니스 센터가 54%나 증가하고, 코로나 시기에도 피트니스 센터가 늘었다는 기사도 어렵지 않게 확인 할 수 있었다. 그만큼 최근들어 사람들의 자기관리에 대한 관심도가 높아지고, 그 중에서도 눈에 띄게 달라진 것은 헬스장 이용 인원이 많아졌다는 것이다. 또한 SNS 나 광고판 및 각종 미디어에서 보아도 바디프로필을 준비 하는 사람들이 많아졌고 예전보다 헬스에 대한 이해도가 상당히 수준 높아졌다는 것을 실감할 수 있다.

많은 타 운동 기록 서비스들은 사용자가 직접 어떤 운동을 했는지 캘린더에 작성해야 한다. 때문에 내가 어떤 운동을 했는지 또 얼마나 수행했는지 직접 작성해야한다는 단점이 있다. 또 다른 운동 기록 서비스들에서는 추가 디바이스를 구매하여 운동 횟수 및 시간을 측정하고 기록하는 기능을 제공하는 경우도 있다. 추가 장비가 있다는 점과 현실적으로 센서의 위에서 할 수 있는 운동만이 측정 대상이기 때문에 운동 자세에 대한 제약사항이 발생하여 실용적이지 않다고 판단 했다.

이러한 기존 서비스들의 단점들을 보완하여 본 팀은 사용자가 본인들의 헬스 영상을 기록하며 운동 일지가 자동으로 기입되는 “득근득근 성장일기”를 기획하였다. “득근득근 성장일기”는 사용자들이 운동하는 동안 본인의 자세를 카메라로 촬영하여 본인의 운동 하는 모습을 동영상으로 만들고, 촬영된 영상에서 이용자의 운동하는 자세를 분석하여 ‘어느’ 운동을 몇 회 하는지 “득근득근 성장일기”에 자동으로 기록되는 기능을 제공한다.

“득근득근 성장일기”에 운동 영상을 기록하고, 자동으로 본인의 운동 일지가 기록하게 되므로 이후 자신에 대한 기록이 되는 것은 물론 운동에 대한 열정도 더할 수 있을 것이고, 운동의 루틴과 일정을 편하게 관리 할 수 있을 것으로 기대한다.

이후 추가적으로 자신이 기록한 영상들의 자세를 분석하여 자세가 정확한지, 올바른 자세는 어떤 자세인지에 대한 정보도 제공 할 수 있도록 확장 시킬 예정이다.

• Dataset

본 프로젝트에서는 AI허브(aihub.or.kr)에 있는 **피트니스 자세 이미지**를 데이터로 사용할 것이다. 좀 더 구체적으로 설명하자면 먼저 한국인 평균의 $\pm 20\%$ 의 체격을 가진 운동 경력 2~5년의 20~30대의 남녀 70여 명을 모델로 하여 40여 개의 운동 동작에 대해 200,000개의 영상(원천 데이터)을 얻었다. 다음으로 얻어진 영상으로부터 초당 1~3개의 영상 이미지를 추출하여, 총 300만 장 이상의 영상 이미지를 추출했고, 각 이미지에서의 24개의 키포인트(Keypoint)와, 운동 정보를 JSON 파일들에 레이블링 하였다. 아래 그림은 운동 형태, 운동 신체 부위, 운동 동작을 실행하는 모델의 성별과 나이에 따른 피트니스 동작 영상의 Clip 분포이다.

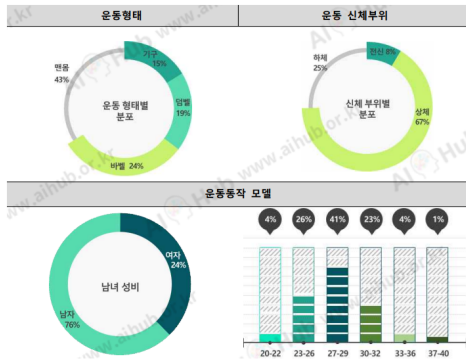


그림1. 피트니스 동작 영상 Clip의 분포

NO	항목명		길이	타입	필수 여부	비고
	한글명	영문명				
	데이터셋 정보	info			Y	JSON Object
1	데이터셋명	info.name	100	String	Y	
2	데이터셋상세설명	info.description	1000	String		
3	데이터셋URL	info.url	200	String		
4	데이터셋생성일자	info.date_created	100	String	Y	
	비디오 정보	(VIDEO_NAME) json		List	Y	각 비디오마다 하나의 JSON
5-1	비디오식별자	id	100	String	Y	
5-2	비디오시나리오식별자	s_id	100	String	Y	같은 시나리오를 공유하는 비디오
5-2	비디오너비	width	4	Number	Y	
5-3	비디오높이	height	4	Number	Y	
5-4	비디오명	name	100	String	Y	
5-5	비디오촬영정보	mat	64	Object		카메라 촬영정보
5-6	프레임 별 파일명	frames[]		List		
5-6-1	-	-	1000	String	Y	
5-7	프레임 별 어노테이션	infos[]		Object		
5-7-1	키포인트	keypoints		Object	Y	
5-7-2	운동 종류	category	1000	String	Y	
5-7-3	운동 상태	status	1000	String	Y	
5-8	비디오길이	length	4	Number		
5-9	비디오라이선스	license	100	String		
5-10	비디오촬영일자	date_captured	100	String		
	라이선스	licenses		List	Y	List of JSON Object
6-1	라이선스명	licenses[] name	100	String	Y	
6-2	라이선스URL	licenses[] url	200	String	Y	

그림2. 어노테이션 포맷

위 그림 2는 우리 조가 사용하는 데이터셋의 어노테이션 포맷이다. 크게 **데이터셋 정보**, **비디오 정보**, **프레임 별 어노테이션(운동 정보)**, **기타 정보** 4가지로 나눌 수 있다. 이중 우리가 구해야 할 것은 운동 정보이며, 먼저 키포인트와 운동 상태를 구해 이를 바탕으로 **최종적으로 운동 종류를 판별할 수 있는 모델을 만들 것이다.**

• Main approach

1. 인체 관절(keypoints) 추출 모델링

본 과정에서는 인체의 주요 움직임을 만들어내는 관절 포인트(이하 keypoints)를 추출해 내는 단계이다.

데이터 셋에 존재하는 24개의 keypoints 데이터를 기반으로 모델링을 진행하여 어떠한 프레임이 주어졌을 때 keypoints를 추출하는 모델을 훈련시킨다. 새로운 데이터에 훈련된 모델을 적용하여 사람의 몸에서 주요 관절을 나타내는 keypoints를 추출해 내는 과정이 완료되면, 탐지한 keypoint 별로 annotation(keypoints에 labeling 하는 작업)을 진행한다.

추후, 프레임에서 주요한 Feature(여기서는 keypoints)를 효율적으로 추출해 내는 Neural Network 알고리즘(ex. CNN)을 학습하게 된다면 그 알고리즘을 채택하여 모델링을 진행할 것이다.

2. 모션 분류 모델링

본 과정에서는 앞서 파악해둔 keypoints의 움직임을 통해 어떤 운동 동작을 하는지 추정하는 작업을 하며, AI 허브의 데이터 셋을 활용하여 지도 학습을 진행한다.

위의 모델에서 결과로 나온 24개의 keypoints의 좌표 정보(각 keypoints 별 움직인 거리, 방향성, 벡터가 이루는 각도 등)를 features로 하여 각 운동에 대한 samples에 대해 KNN, SVM, Random Forest, Logistic Regression, XGBoost 등의 알고리즘을 이용한 모델링을 진행할 수 있다. 모델링 과정에서, Cross validation을 통해 각 모델의 Optimal Hyper Parameter를 선정하고, 가장 좋은 성능을 내는 모델을 최종 모델로 채택하거나 또는 앙상블 기법을 활용할 수도 있을 것이다.

또한, Neural Network 알고리즘을 통한 모델링을 진행할 수도 있지만, 아직 배우지 못한 부분이 많아 해당 모델링에 대해서는 학습 후에 진행하도록 할 것이다.

Neural Network 모델링을 하게 된다면, 우리는 계속해서 움직이는 사람의 동작(연속적인 프레임)의 변화를 처리하도록 해야 하므로, 순차적 데이터의 처리에 좋은 LSTM(Long Short-Term Memory)을 이용할 수 있을 것으로 생각된다.

앞으로 배울 Neural Network 알고리즘을 사용한다면 전체적인 모델 흐름도는 다음과 같다.

1. 입력으로 들어온 이미지에서 keypoints를 추출하는 모델링(CNN 이용)을 한다.
2. LSTM을 이용하여 각 이미지 시퀀스의 연관성을 새로운 모델로 학습시킨다.

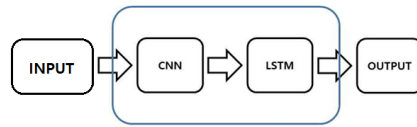


그림 3. 전체 모델 흐름도

- Performance Evaluation

모델링 부분에서 선택한 여러 알고리즘 후보군들 중 다음의 평가 방법을 적용하여 가장 좋은 성능을 내는 모델을 선택할 것이다.

1. One-vs.rest VS One-vs.one

class의 수가 30개나 되기 때문에 비교해야 할 classifier의 개수가 너무 많아지는 One-vs.one 방식보다는 One-vs.rest 방법을 활용하여 모델의 성능을 평가할 예정이다.

2. Confusion matrix 분석

Confusion matrix를 통해 혼동하기 쉬운 class를 파악하고, confusion matrix의 결과를 바탕으로 Accuracy, Precision, Recall, F1 Score를 계산하여 모델의 성능을 평가할 예정이다.

또한 Confusion matrix로부터 Class Prediction error와 ROC curve를 활용하여 시각적인 검증도 가능하게 할 예정이다.