

# Facenet을 이용한 Liverpool 선수 검출기



신경회로망특론 Term Project  
120200190 이인권

# Contents

1. 프로젝트 주제
2. 데이터 수집
3. 데이터 전처리
4. 모델 및 학습
5. 결과 및 분석

# Subject

Facenet을 이용하여 축구팀 리버풀 선수들의 얼굴을 검출하고 어떤 선수인지 분류하는 검출기 구현



M. Salah



T. Alexander-Arnold



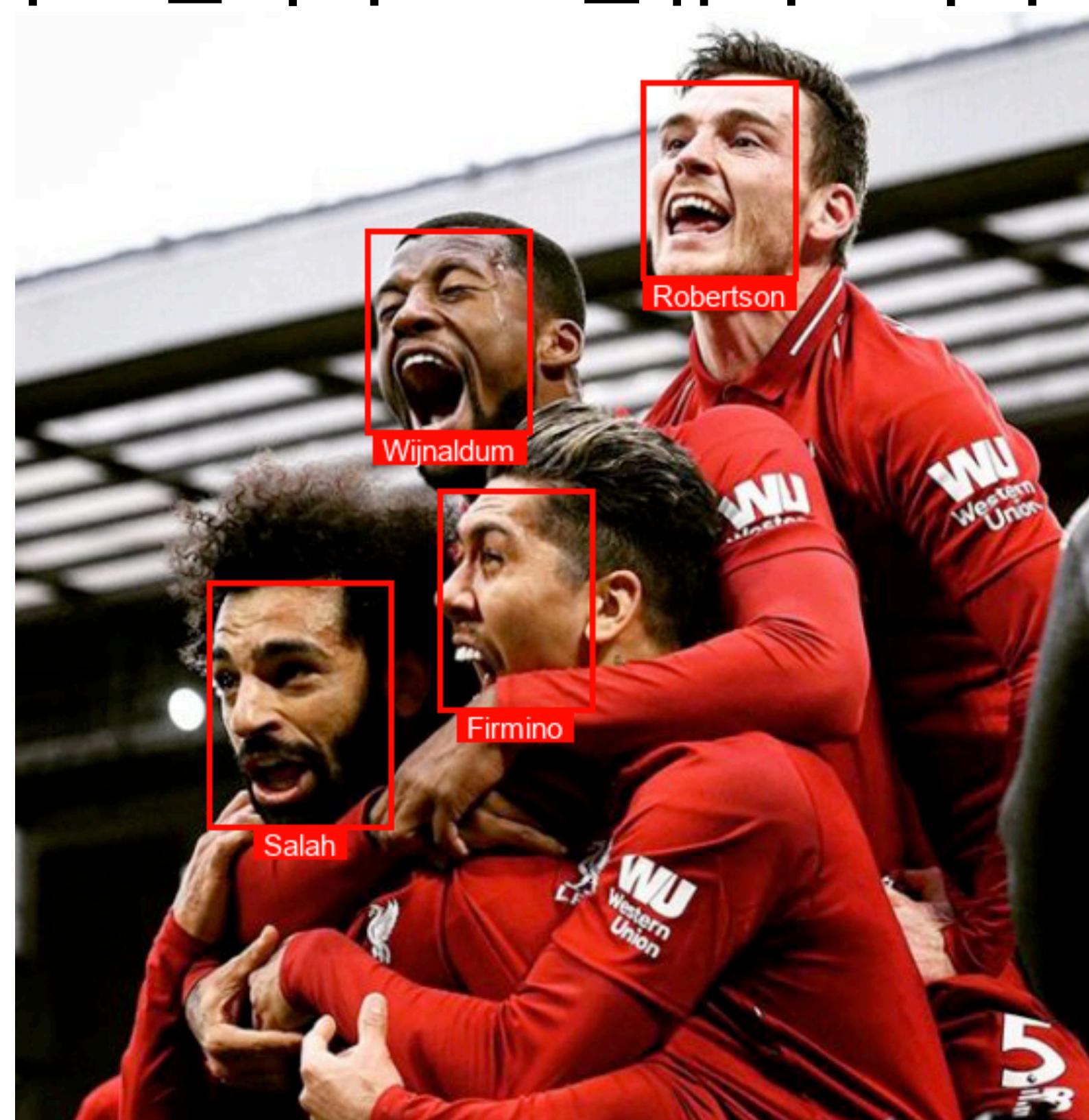
V. Van Dijk



→ J. Henderson

# Subject

- 한 명의 선수가 포함된 사진 뿐만 아니라, 여러 명의 선수가 포함된 사진에서도 선수들의 얼굴을 정확하게 검출하고 분류하도록 구현
- 더 나아가 동영상에서도 연속적으로 분류하도록 구현



# Domain

- 17-18 시즌 이후 리버풀 선수들을 도메인으로 선정
- 검출하고자 하는 **객체 클래스 개수는 총 23개 (23명의 선수)**

Alisson



V. van Dijk



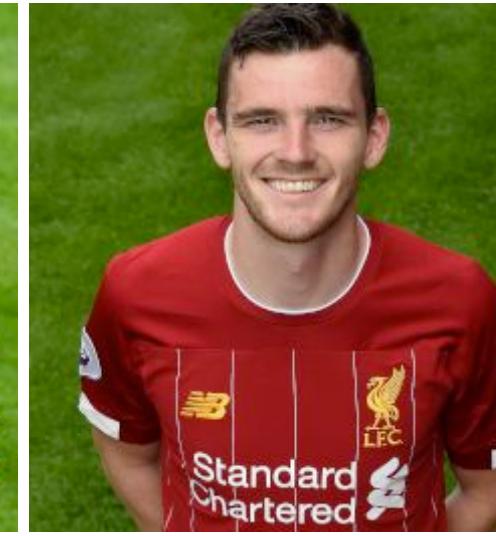
D. Lovren



J. Gomez



A. Robertson



J. Matip



T. Alexander-Arnold



Fabinho



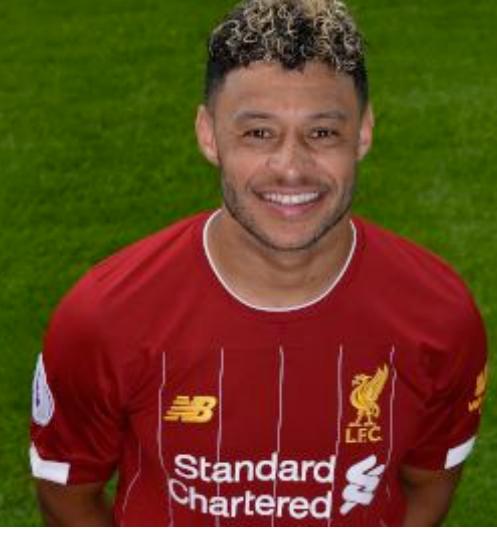
G. Wijnaldum



J. Milner



Oxlade-Chamberlain N. Keita



N. Keita



J. Henderson



A. Lallana



X. Shaqiri



R. Firmino



S. Mane



M. Salah



D. Origgi



S. Mignolet



A. Moreno



D. Sturridge

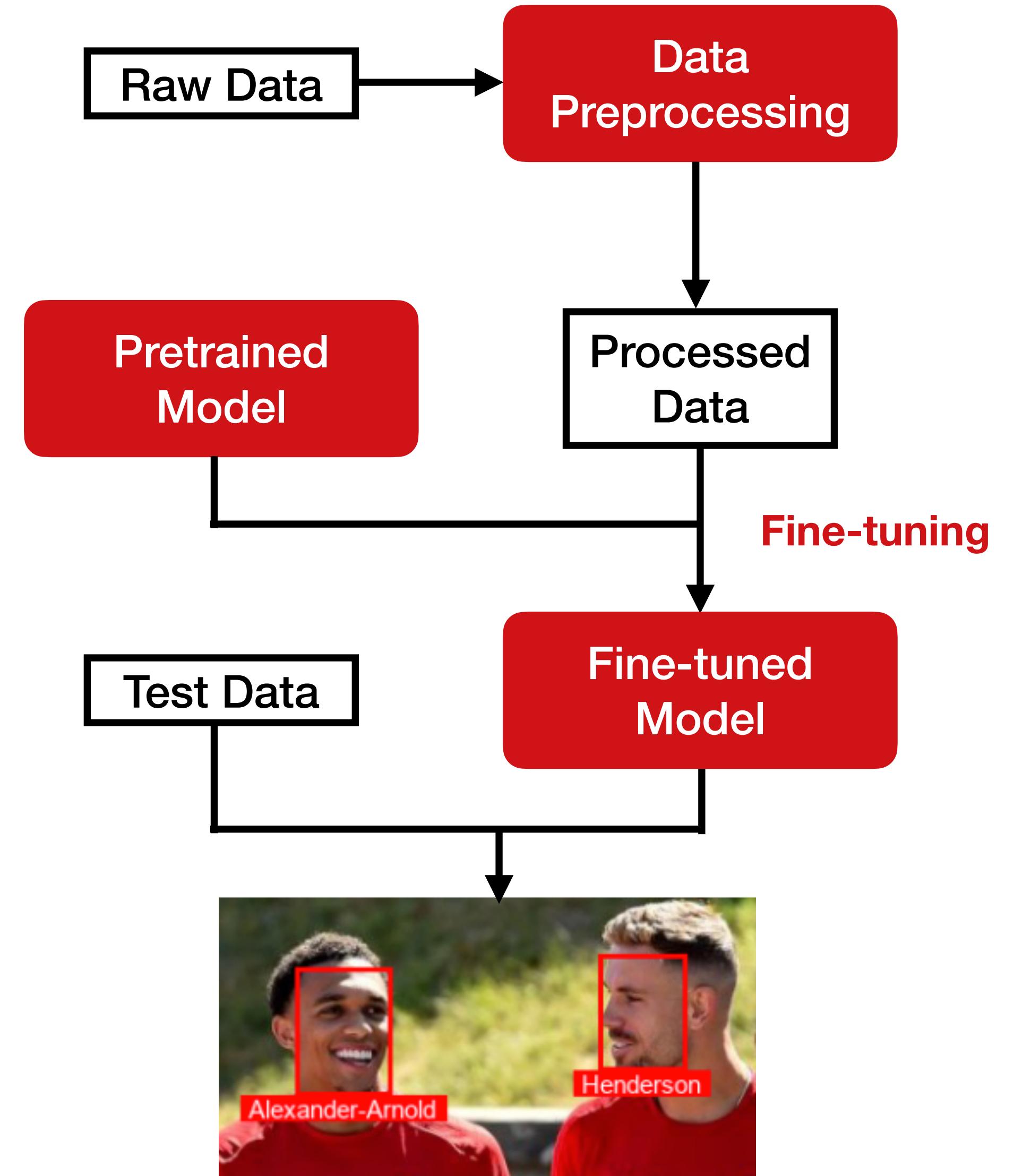


J. Klopp



# Model Overview

- 수집된 데이터를 전처리하여 입력 데이터를 생성
- 생성한 데이터를 사용해, 기존에 pretrain 되어있는 model을 **fine-tuning**
- Fine-tuning 된 모델에 테스트하기 위한 사진 혹은 동영상을 넣어주면 선수 검출 및 분류를 한 뒤 최종 결과를 생성



# Contents

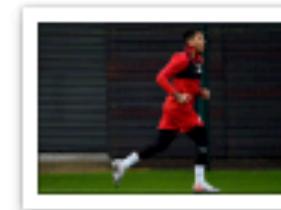
1. 프로젝트 주제
2. 데이터 수집
3. 데이터 전처리
4. 모델 및 학습
5. 결과 및 분석

# Data

- 네이버의 블로그에 모여있는 훈련 사진을 크롤링하여 수집

(출처 : <https://blog.naver.com/htpjk>  
블로그 주인에게 허가를 받은 뒤 사용)

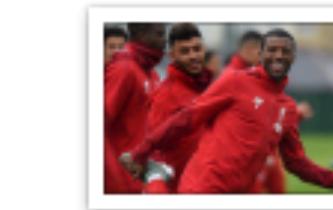
- 총 1022장의 사진을 수집
- 여러 선수들의 얼굴이 포함되어 있는 사진도 존재
- 반대로 얼굴이 하나도 포함되어 있지 않은 사진도 존재
- 검출하고자 하는 domain은 데이터를 조사하여 빈도수가 높은 선수들로 직접 선정



61111510.png



61111511.png



61111512.png



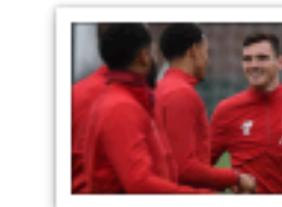
61111513.png



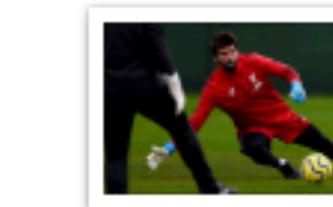
61111514.png



61111515.png



61111516.png



61111517.png



61111518.png



61111519.png



125571550.png



125571551.png



125571552.png



125571553.png



125571554.png



125571555.png



125571556.png



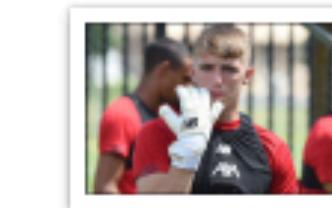
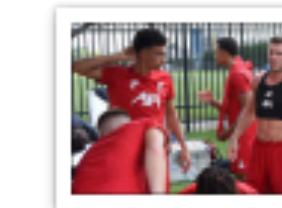
125571557.png



125571558.png



125571559.png

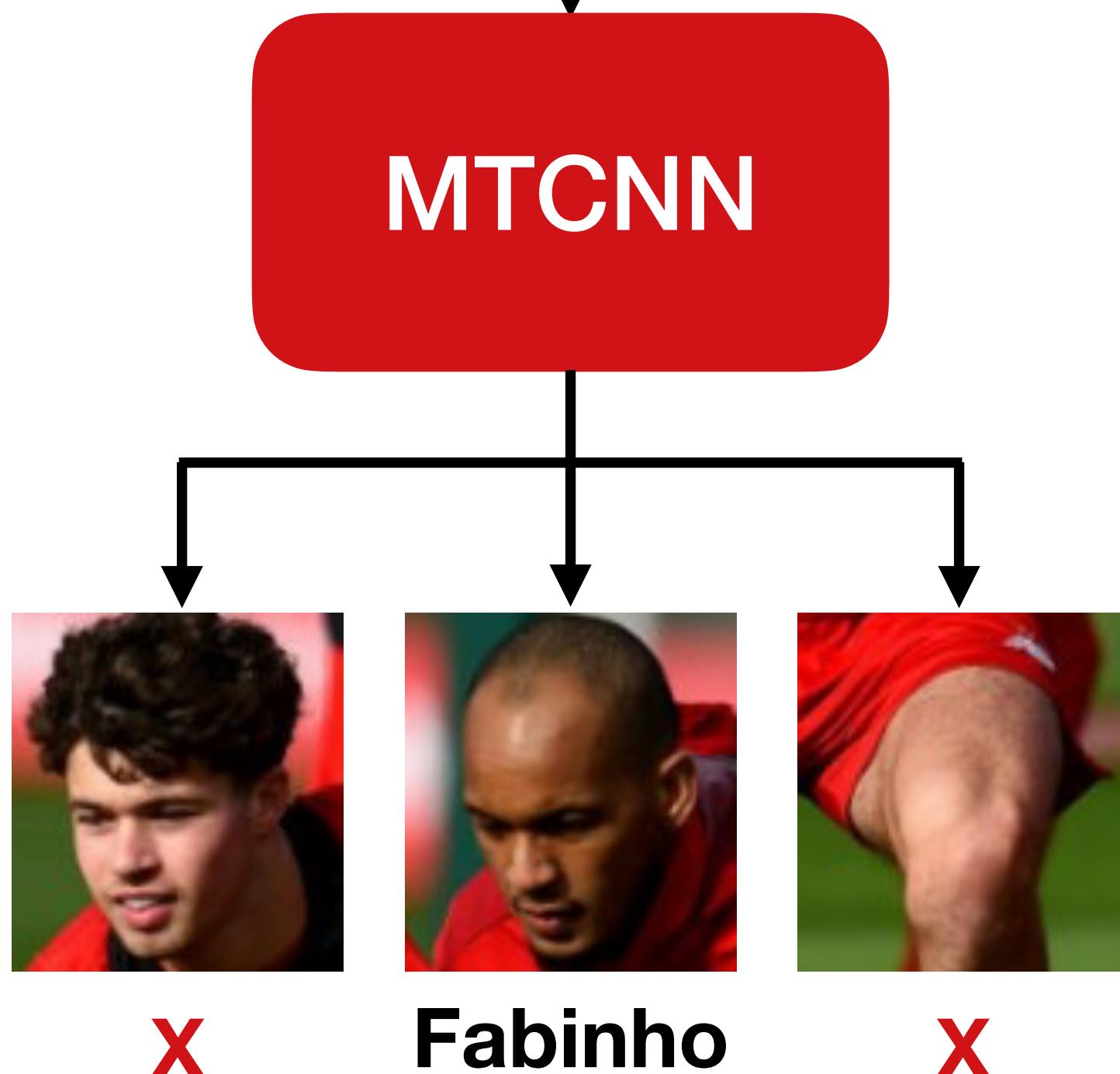


# Contents

1. 프로젝트 주제
2. 데이터 수집
3. 데이터 전처리
4. 모델 및 학습
5. 결과 및 분석

# Data Processing

- Pretrained model MTCNN을 이용해 이미지에서 사람 얼굴을 검출
- 총 1022장의 이미지로부터 MTCNN을 통해 2331개의 얼굴을 검출
- 검출된 후보들은 직접 labeling을 통해 얼굴이 아닌 후보들과 domain에 포함되지 않은 후보들을 제거한 뒤, 후보들을 분류
- 최종적으로 fine-tuning에 사용한 데이터는 총 1225개의 얼굴 이미지



# Data Processing

- 선수별 데이터의 개수를 나타내는 표
- 최대 103개, 최소 10개

|              |     |              |    |                       |    |            |    |              |    |
|--------------|-----|--------------|----|-----------------------|----|------------|----|--------------|----|
| M. Salah     | 103 | Fabinho      | 66 | J. Milner             | 57 | J. Matip   | 40 | S. Mignolet  | 17 |
| V. van Dijk  | 100 | G. Wijnaldum | 65 | A. Oxlade-Chamberlain | 54 | X. Shaqiri | 36 | A. Moreno    | 11 |
| J. Klopp     | 89  | S. Mane      | 65 | A. Robertson          | 53 | N. Keita   | 31 | D. Sturridge | 10 |
| J. Henderson | 80  | J. Gomez     | 58 | T. Alexander-Arnold   | 50 | A. Lallana | 29 |              |    |
| R. Firmino   | 78  | D. Lovren    | 58 | D. Origic             | 48 | Alisson    | 27 |              |    |

# Contents

1. 프로젝트 주제
2. 데이터 수집
3. 데이터 전처리
4. 모델 및 학습
5. 결과 및 분석

# Pretrained Models

- 사용하는 pretrained model은 총 2개
- Face detection을 위한 MTCNN과,  
face recognition & classification을 위한 InceptionResnetV1

**MTCNN**

**InceptionResnetV1**

# **MTCNN**

- CNN을 활용하여 face detection 정확도와 성능을 끌어올린 모델
- Face detection 분야에서 높은 성능과 빠른 속도를 보여주어 아직도 많은 논문들과 프로젝트들에서 사용
- P-Net, R-Net, O-Net 세 가지 CNN을 차례로 통과
- P-Net : 이미지에서 얼굴로 추정되는 박스들의 리스트 생성
- R-Net : 생성된 박스들 중에서 진짜 얼굴에 해당하는 영역을 추림
- O-Net: R-Net이 추려낸 박스들을 Conv layer와 FC layer를 통과시키며 최종적으로 face detection, face alignment output을 생성
- Multi Task Loss를 통해 세 가지 CNN을 학습

# InceptionResnetV1

- MTCNN 모델로 분류한 얼굴을 classification 하는 모델
- 네트워크는 망이 깊고(deep), 레이어가 넓을수록(wide) 좋지만 parameter의 크기와 연산의 한계 그리고 vanishing gradient 등의 문제가 존재
- Inception Module을 이용해 깊은 망과 넓은 레이어를 가지는 네트워크를 설계한 모델
- Vanishing gradient 문제를 해결하기 위해 중간 layer에서 보조적인 softmax layer를 추가하여 loss를 계산

# Pretrained Models

- 사용하는 pretrained model은 총 2개
- Face detection을 위한 MTCNN과,  
face recognition & classification을 위한 InceptionResnetV1

## MTCNN

## InceptionResnetV1

- 두 모델 중 fine-tuning 하는 모델은 InceptionResnetV1
- 데이터 중 80%를 training set, 20%를 validation set으로 나누어 학습

# Processing Outputs

- Test하기 위한 리버풀 선수들이 포함된 사진 혹은 동영상을 input으로 사용
- MTCNN을 통해 데이터에 존재하는 사람 얼굴을 검출
- Fine-tuning 된 InceptionResnetV1으로 각 얼굴들의 logit을 계산
- 계산된 logit을 softmax취해 가장 높은 값을 가지는 index로 classify
- 이때 domain에 포함되지 않은 얼굴과 얼굴이 아닌 것들을 걸러내기 위해,  
max logit 값이 일정 threshold 값을 넘기지 못하거나,  
두 번째로 큰 logit 값과의 차이가 크지 않을 경우 결과에서 제외
- threshold 값과 top 2 logit 값의 차이 간격은 여러 데이터 결과를 통해 결정

# Contents

1. 프로젝트 주제
2. 데이터 수집
3. 데이터 전처리
4. 모델 및 학습
5. 결과 및 분석

# Result

- Fine-tuning 결과 **82.42%**의 validation accuracy가 나옴

```
Epoch 8/8
-----
Train | 62/62 | loss: 0.1665 | fps: 160.3688 | acc: 0.9587
Valid | 16/16 | loss: 0.5392 | fps: 542.7217 | acc: 0.8242
```

- 학습데이터가 충분한 class들은 좋은 결과를 보임
- 하지만 학습데이터가 충분하지 않은 class들은 좋지 않은 결과를 보임
- 학습 데이터가 최소 30개는 넘어야 선수들을 안정적으로 분류

# Result

- 학습데이터가 **많은** class들의 분류 결과



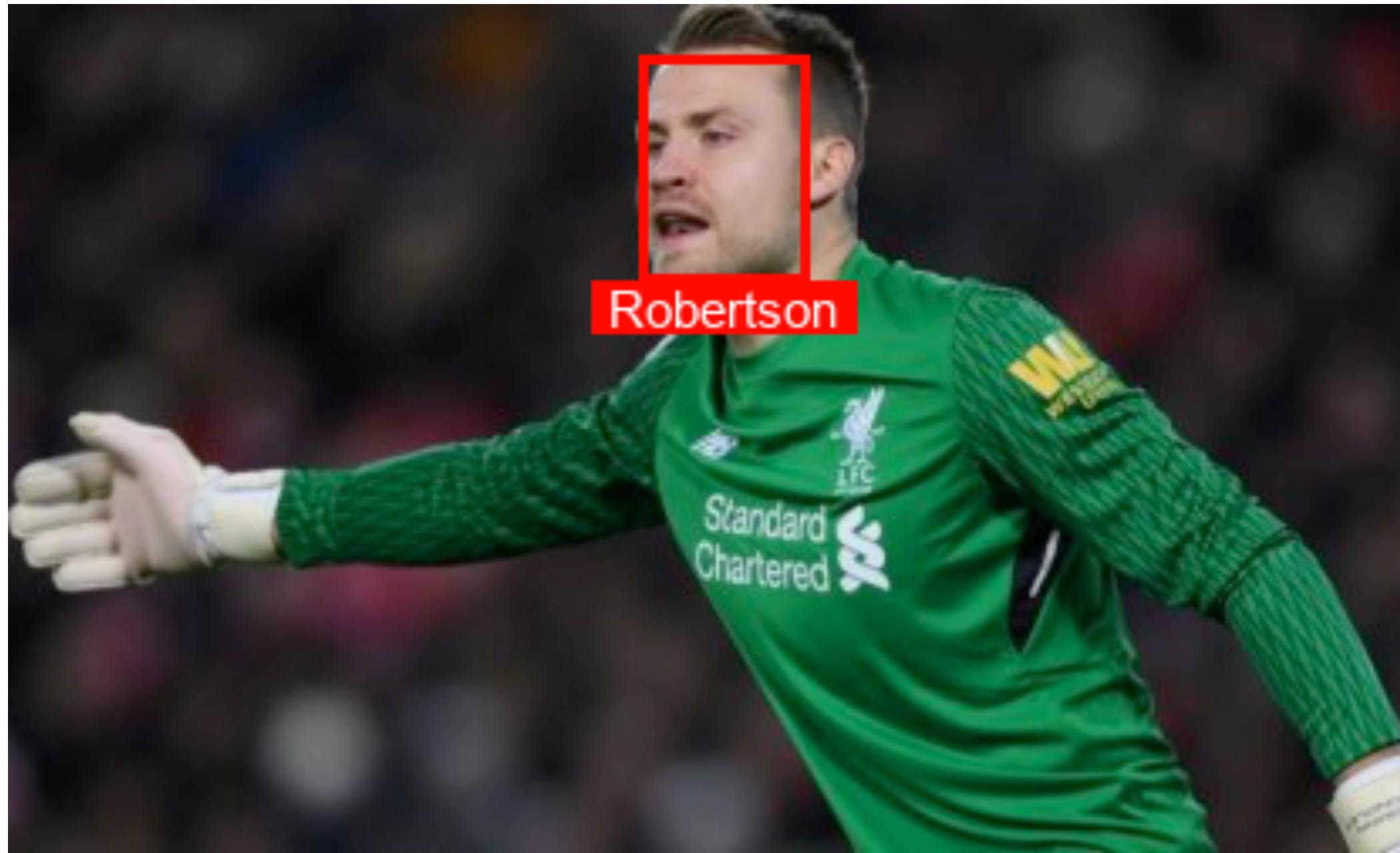
# Result

- 학습데이터가 **많은** class들의 분류 결과
- Domain에 존재하는 얼굴과 존재하지 않는 얼굴을 정확히 분류

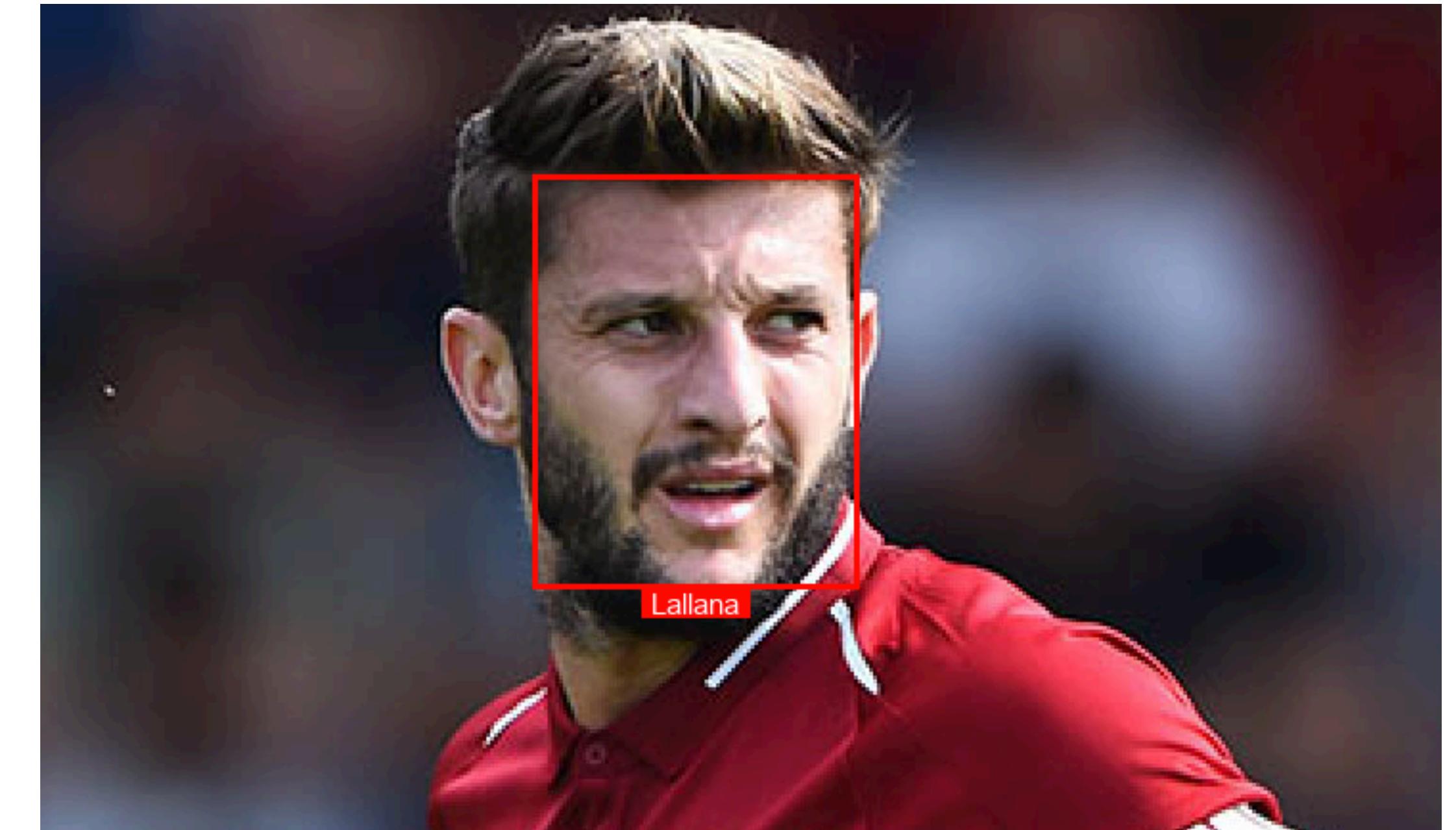


# Result

- 학습데이터가 **부족한** class들의 분류 결과



**오답**  
(정답 : Mignolet)  
학습 데이터 17개



**정답**  
학습 데이터 29개

# Conclusion

- Pretrained Model를 직접 구축한 데이터를 통해 Fine-tuning하여 축구 선수 얼굴 검출기를 제작
- 학습 데이터 구축부터 CNN모델 Fine-tuning 후 데모 생성까지 전체 프로세스에 대한 재미있고 좋은 경험
- 특히 학습 데이터 수집, 전처리, 라벨링 등 데이터를 처리하는 과정에 대한 많은 경험
- 부족한 데이터에 대한 처리를 하지못해 아쉬움

# 감사합니다

