UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA

Basketball Intelligent System

Proof of Concept

Angelo Mozzillo 160313, Brandon Willy Viglianisi 156662, Cristian Di Feo 160327

Project Goals

Introduction

Intelligent system capable of understanding the game of basketball and collect statistics through visual data

Test and analyze the pros and cons of different procedures to resolve the task of detection and identification of basketball players in videos

The Approaches

Classical Approach

Court Detection

Player Detection

Player Retrieval
Identification

Players Homography
to 2D

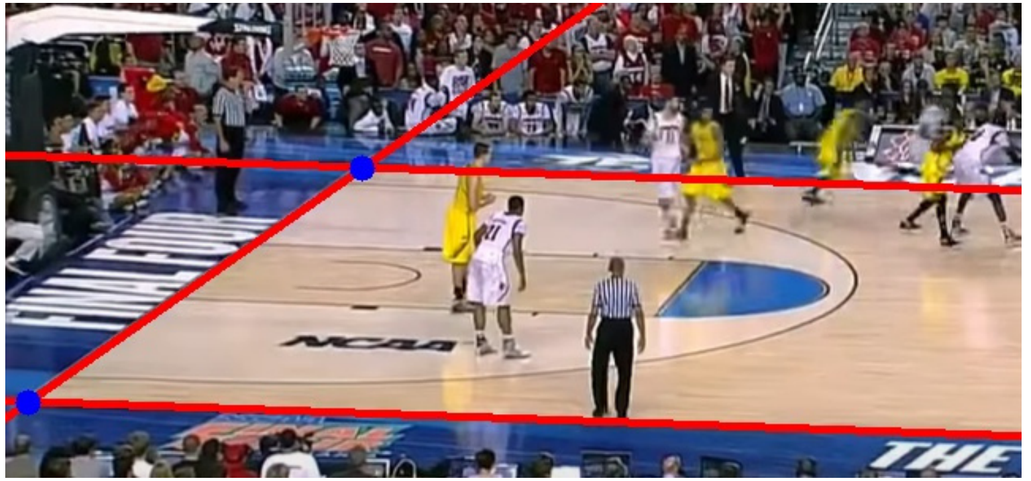
Deep Approach

YoloV5

01 Court Detection

Edge Based Hough lines

- To prevent errors due to noise some pre-processing steps were implemented including a gaussian blur
- Iterative approach changing the threshold of Canny and Hough in order to obtain only 3 lines



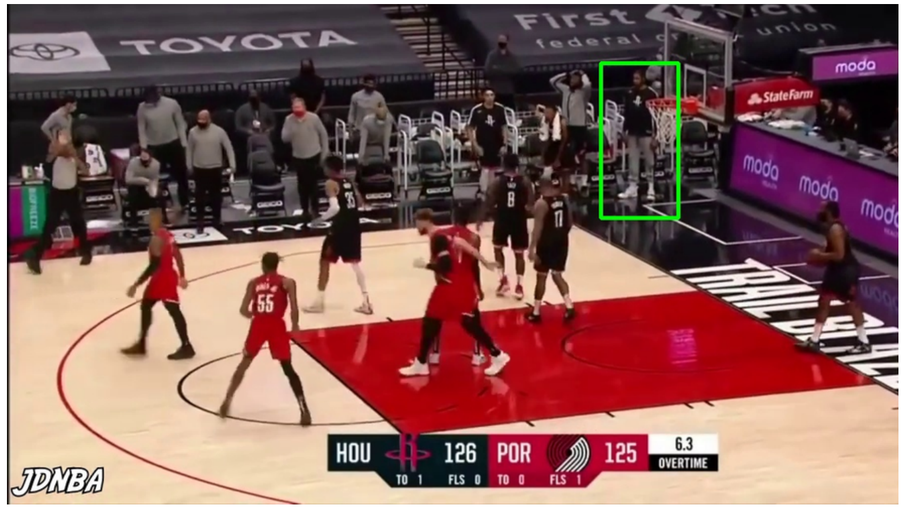
Lens distorsion

02 Player Detection

HOG Detector combined with pre-trained SVM

Standard Parameters:

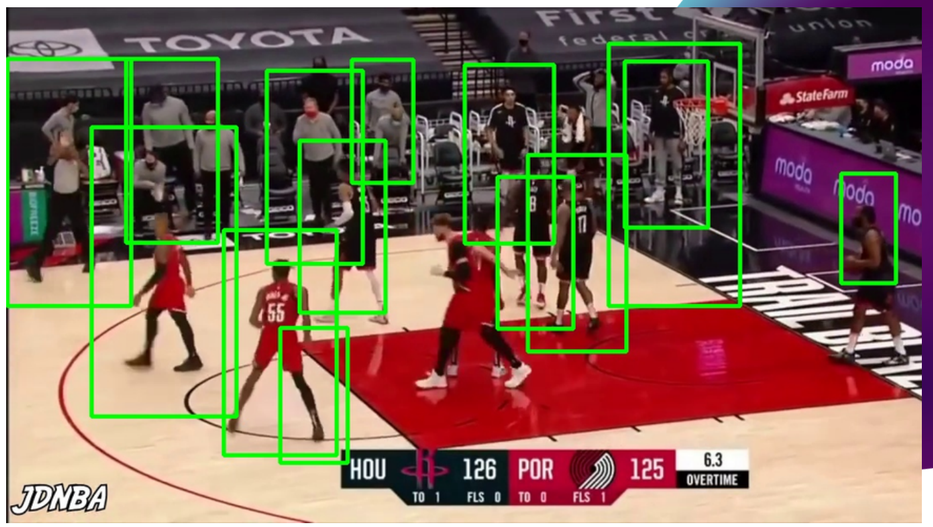
- **sliding_window: 8 * 8**
- **scales: 1.5**



HOG Detector combined with pre-trained SVM

Tuned Parameters

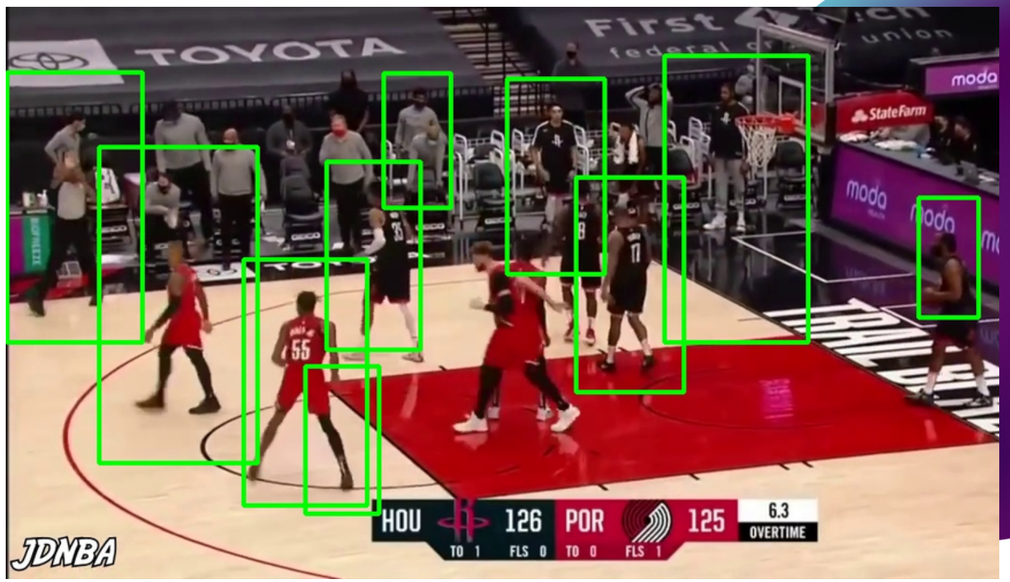
- **sliding_window: 4 * 4**
- **scales: 1.05**

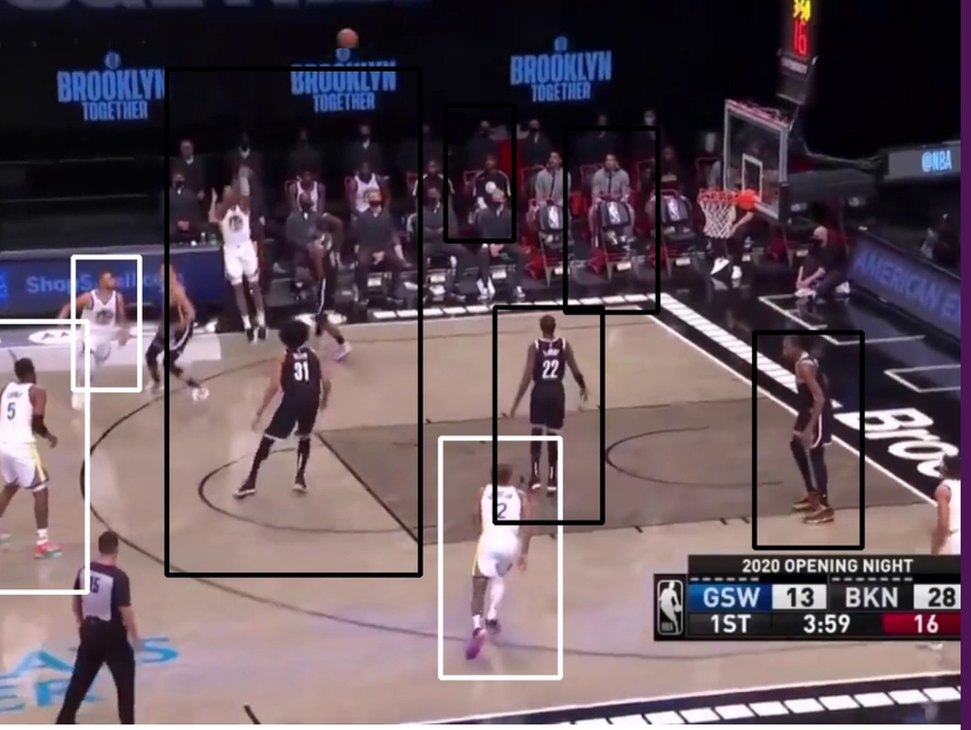


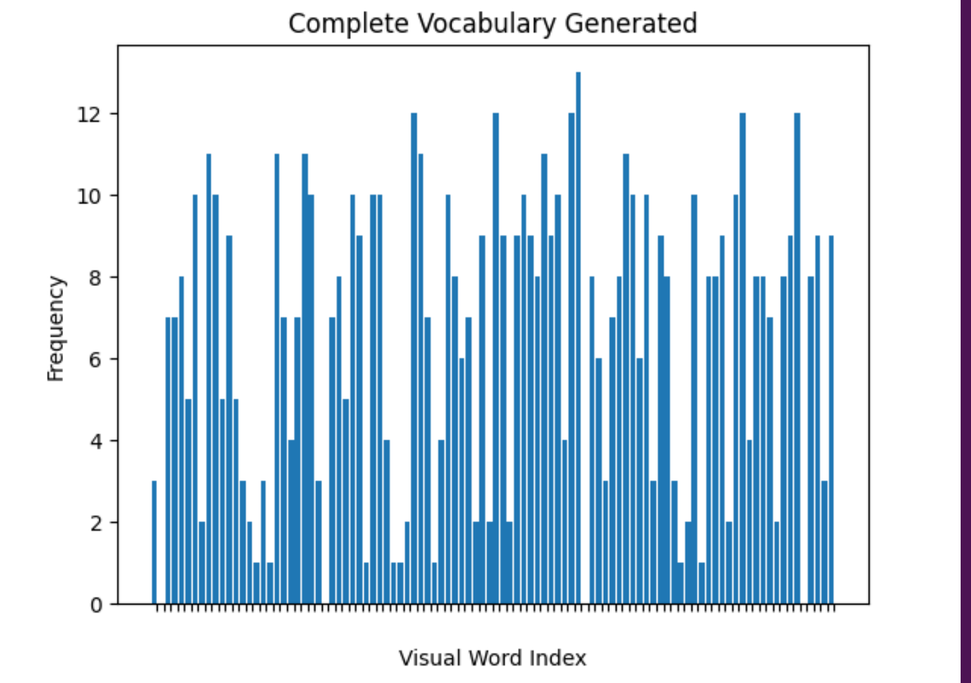
02 Player Detection

Non-Maximum Suppression (NMS)

Final Results







Team Identification

03

Color Based

Three predefined colors

Assignment through minimum difference between the bounding box main color and the predefined colors

BoW & SVM

Vocabulary definition using SIFT and Kmeans

Description of every image with the histogram of visual words occurrences

Computing SVM to classify players in teams

Query image

First retrieved image

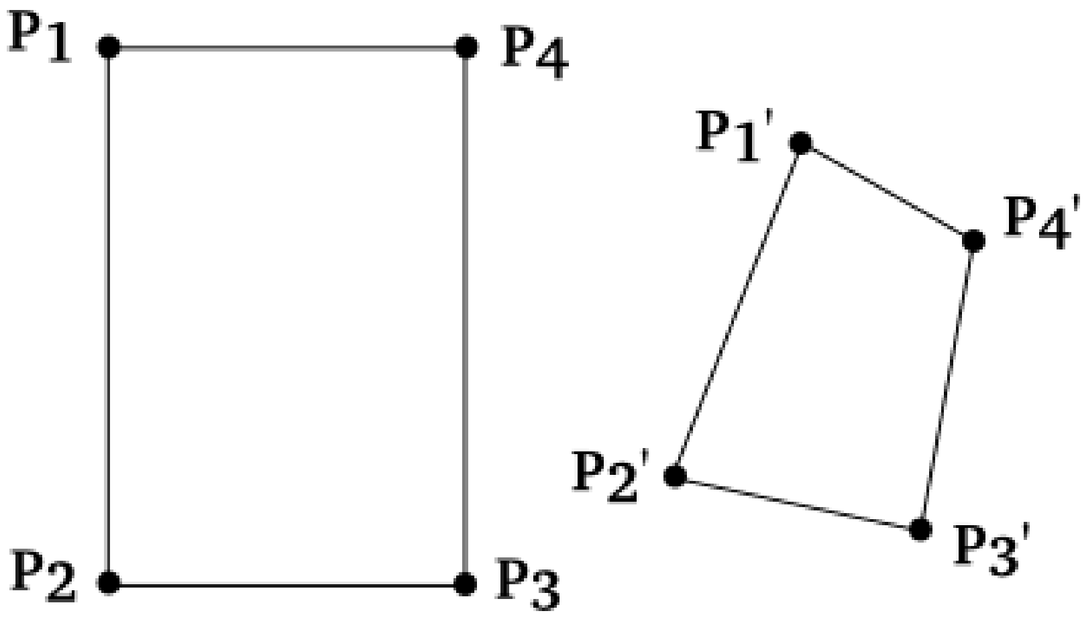
Player Retrieval Identification

04

- A feature is created for every image using BoW with SIFT
- The histograms of visual words describing each image of the dataset are computed and stored
- Given a query image we calculate using our vocabulary its features descriptor
- Earth Mover's Distance is used to compare our image with the one of the dataset
- The algorithm return the closest image

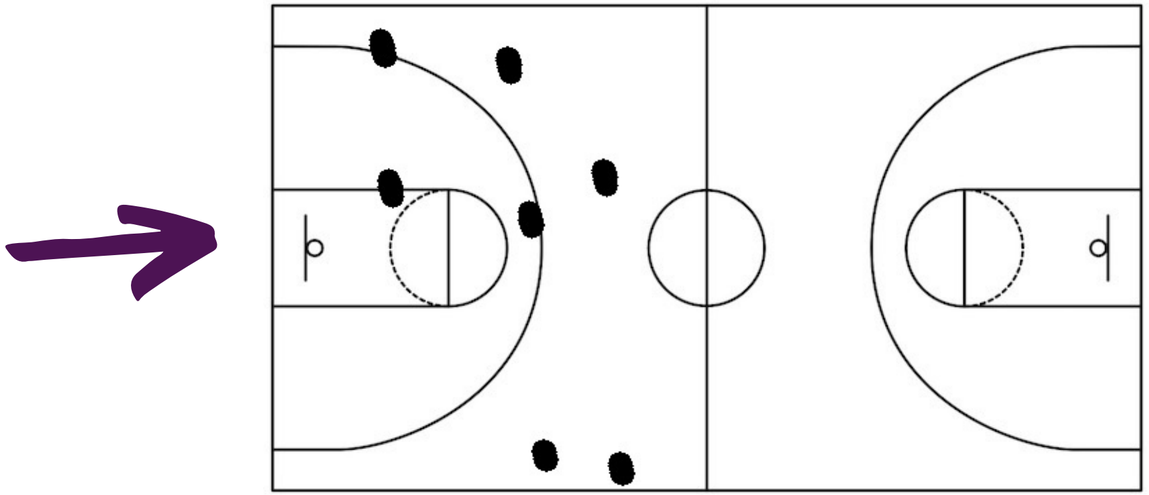
05

Players Homography to 2D

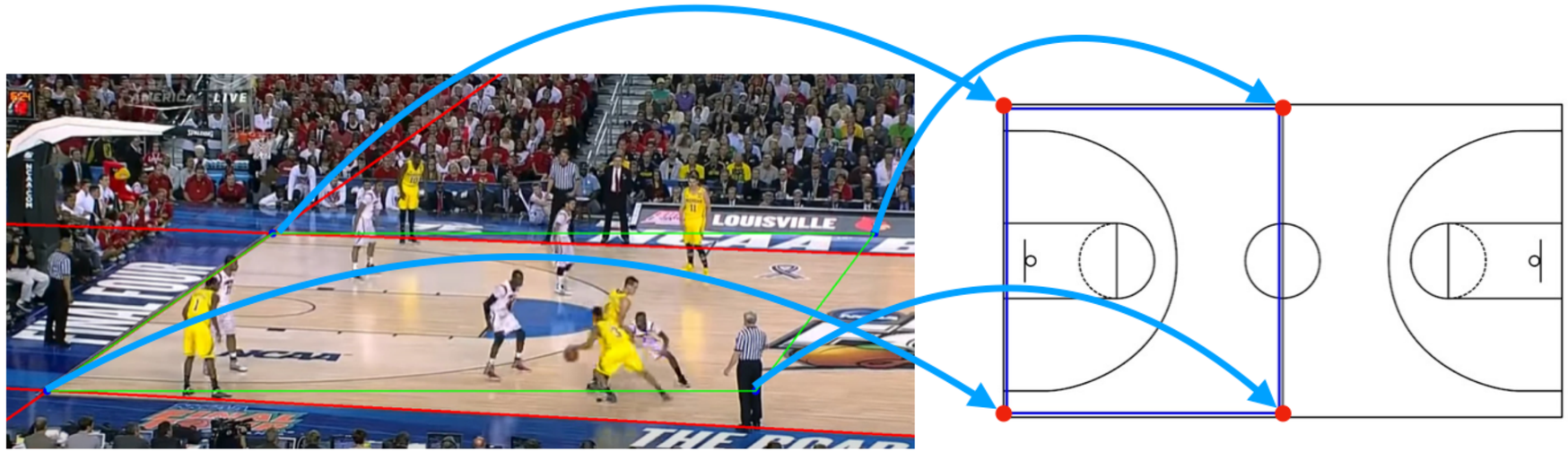


Goal

Find the bird view of the court in 3D, projecting the players coordinates in order to do tactic analysis



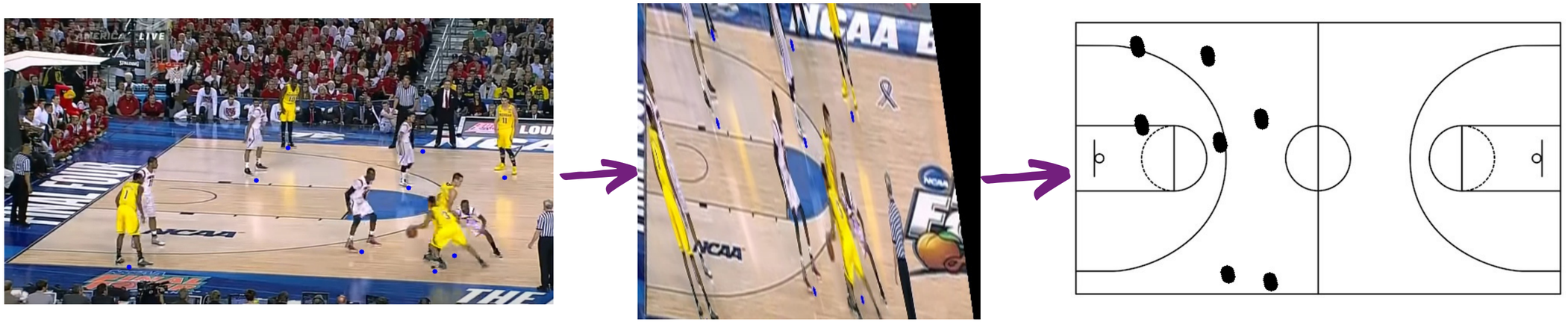
The Process 1/2



✅ findHomography()

❌ getPerspectiveTranformation()

The Process 2/2



- warpPerspective()
- warpAffine()

Future Improvements

Game Tactics Analysis

Through Homoghaphy Best Actions in a 2D court

Complete Pipeline

Direct link from video to analytics

Deep Approach

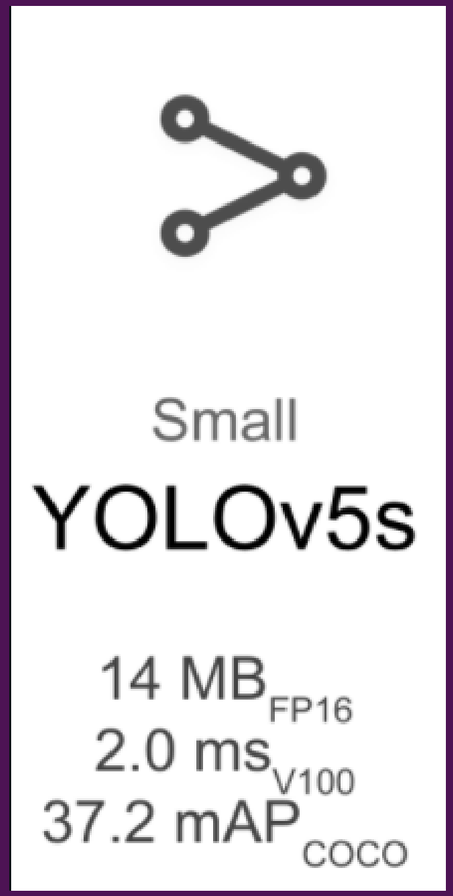
Why YoloV5?

- All of the others object detection algorithms use regions to localize the object within the image and don't look the complete image
- YOLO v5 has a clear advantage in terms of run speed
- Easy to train and run inference on
- Small dataset and low power of our machine

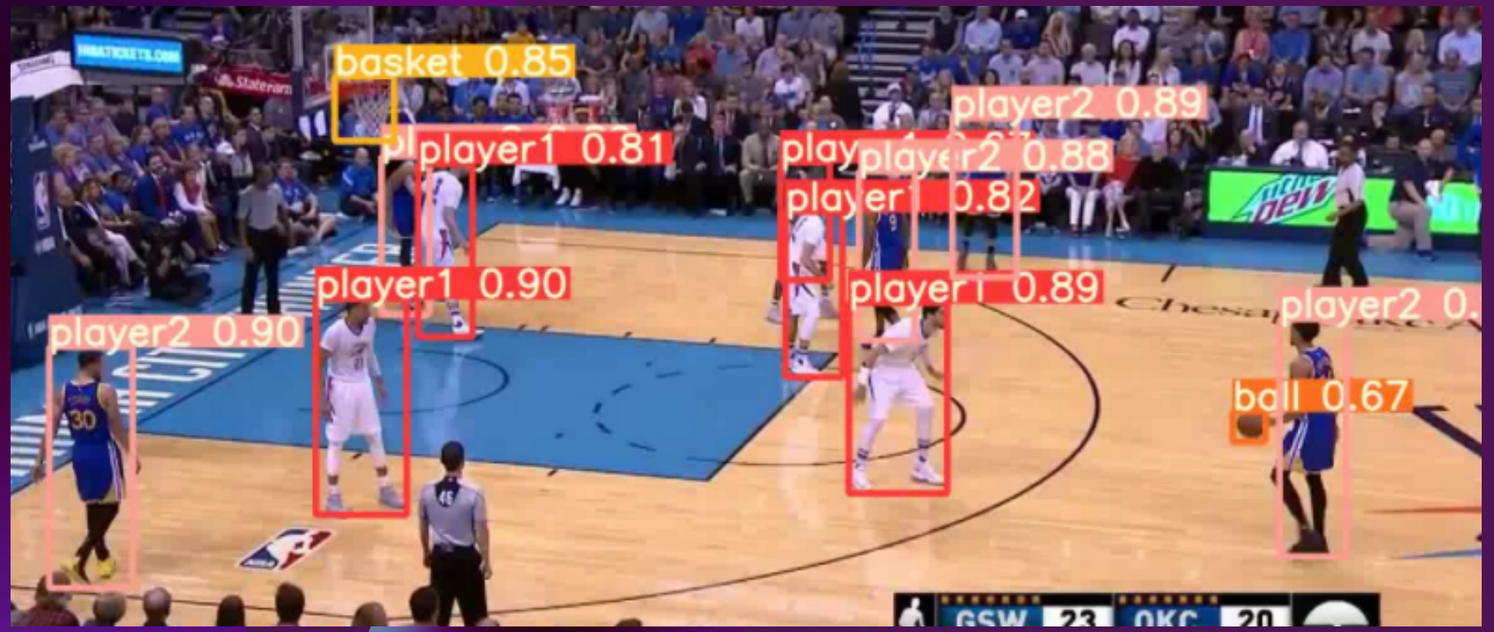
YoloV5 Training

Dataset closer to our needs, containing 500 images of basketball games, and label it by hand, using the LabelImg tool.

100 epochs
32 batch
640x640 images

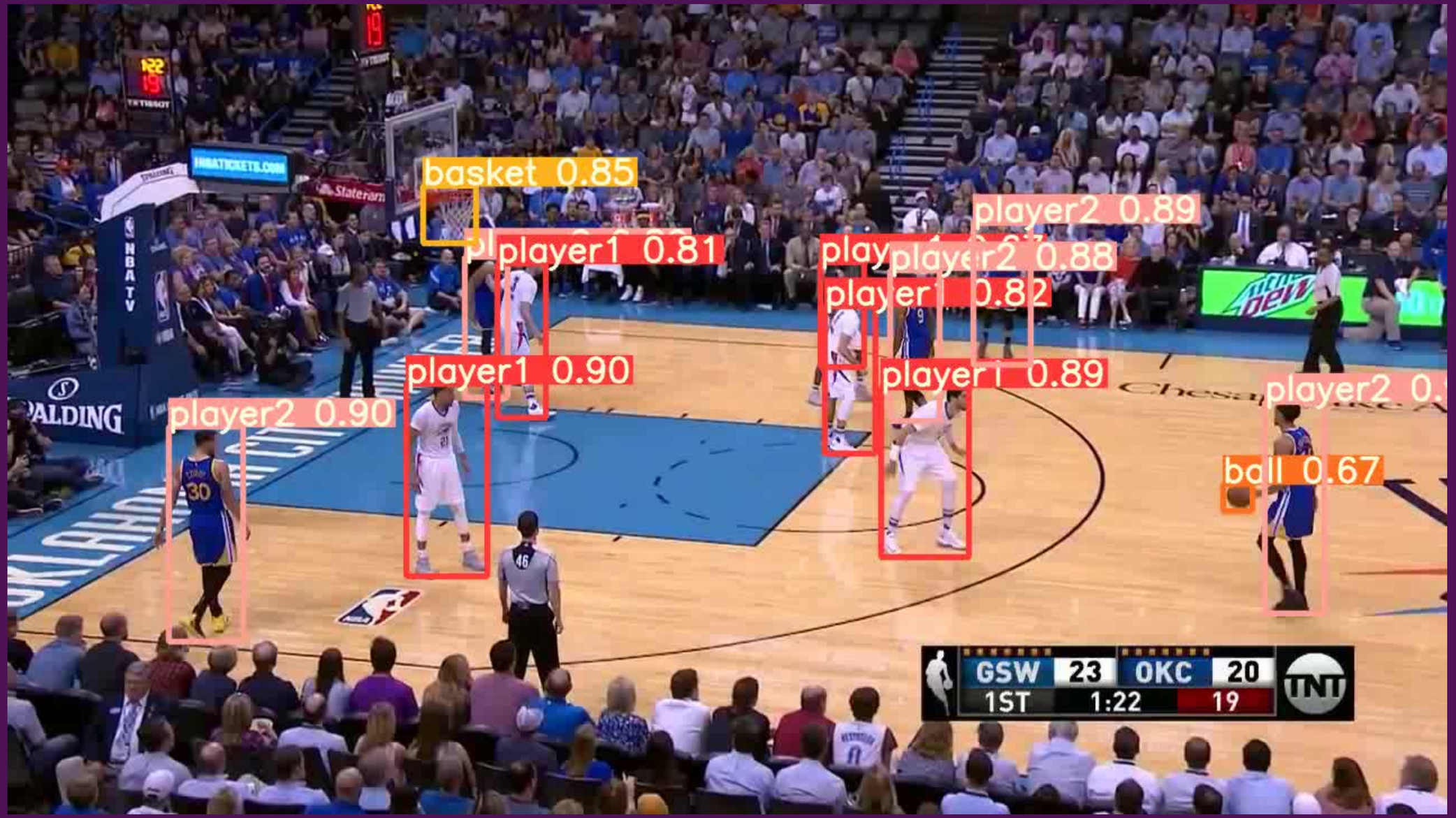


Player Detection & Team Identification

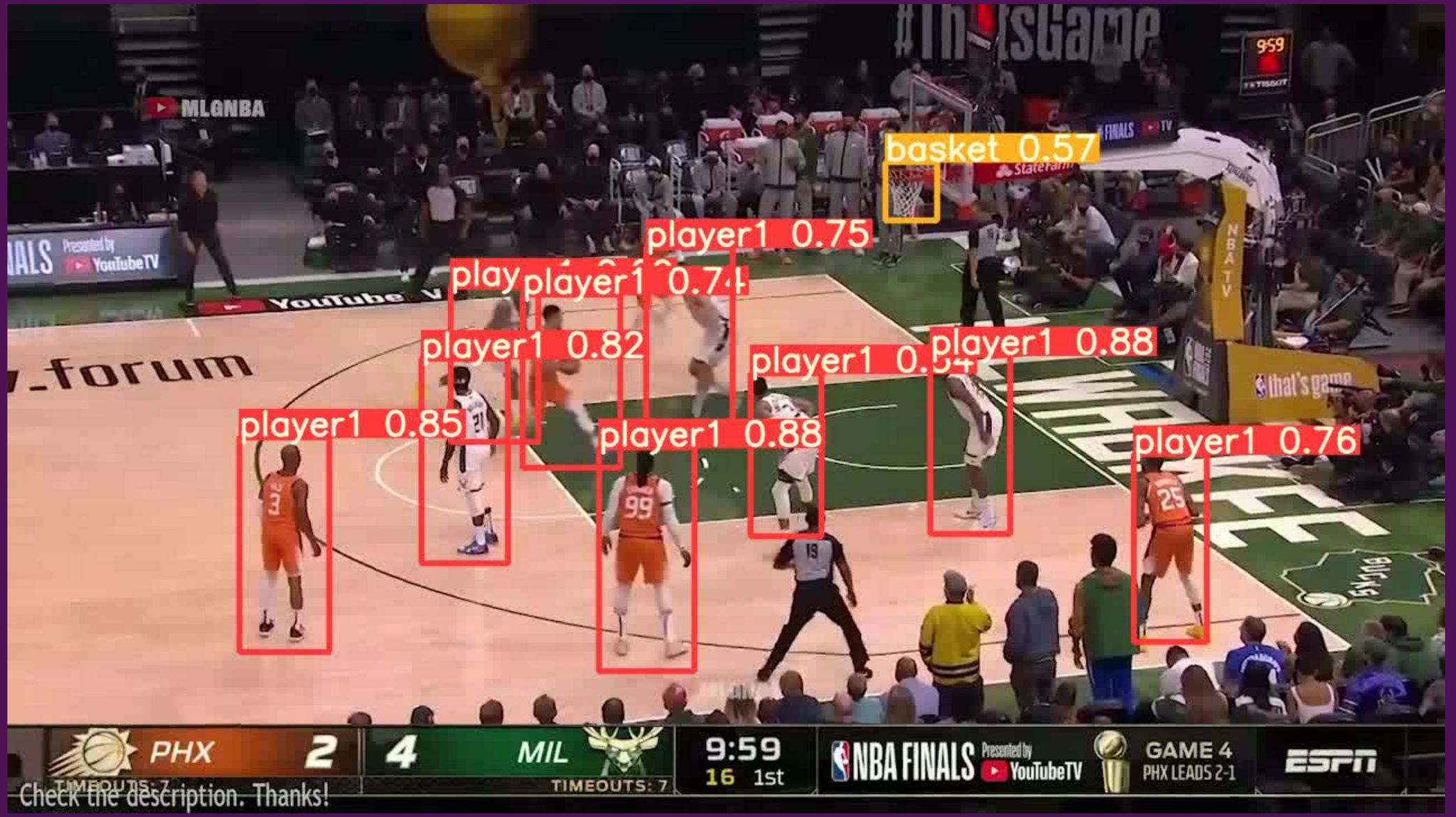


4 **Labels**:

- Player1
- Player2
- Ball
- Basket


basket 0.85
player1 0.81
player2 0.89
player2 0.88
player1 0.90
player1 0.89
player2 0.90
ball 0.67
GSW 23
OKC 20
1ST 1:22 19
TNT

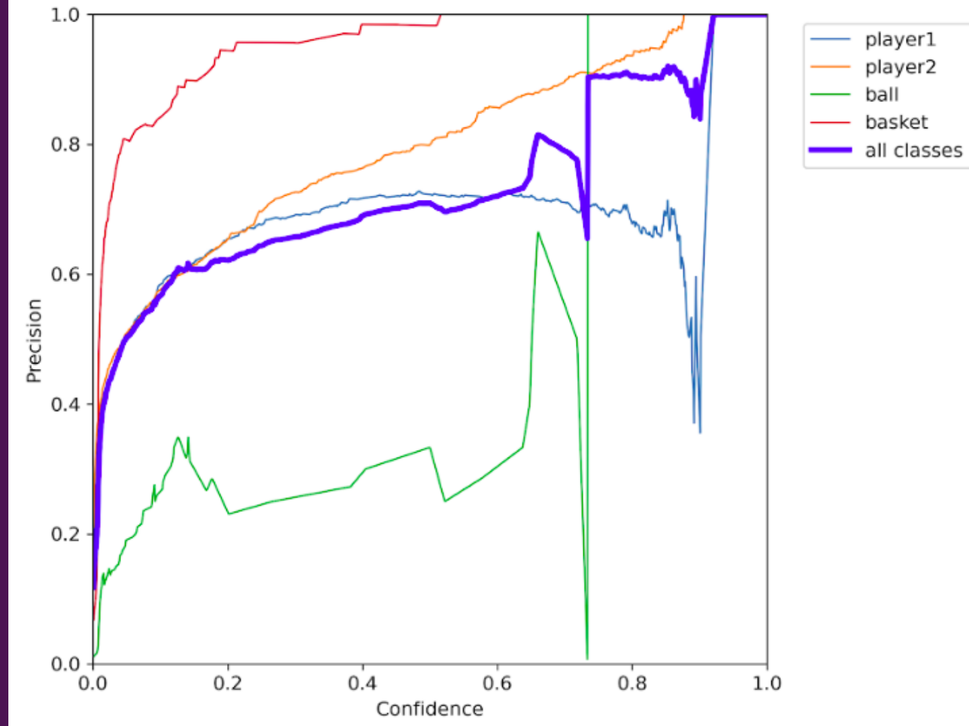

player2 0.69
ball 0.43

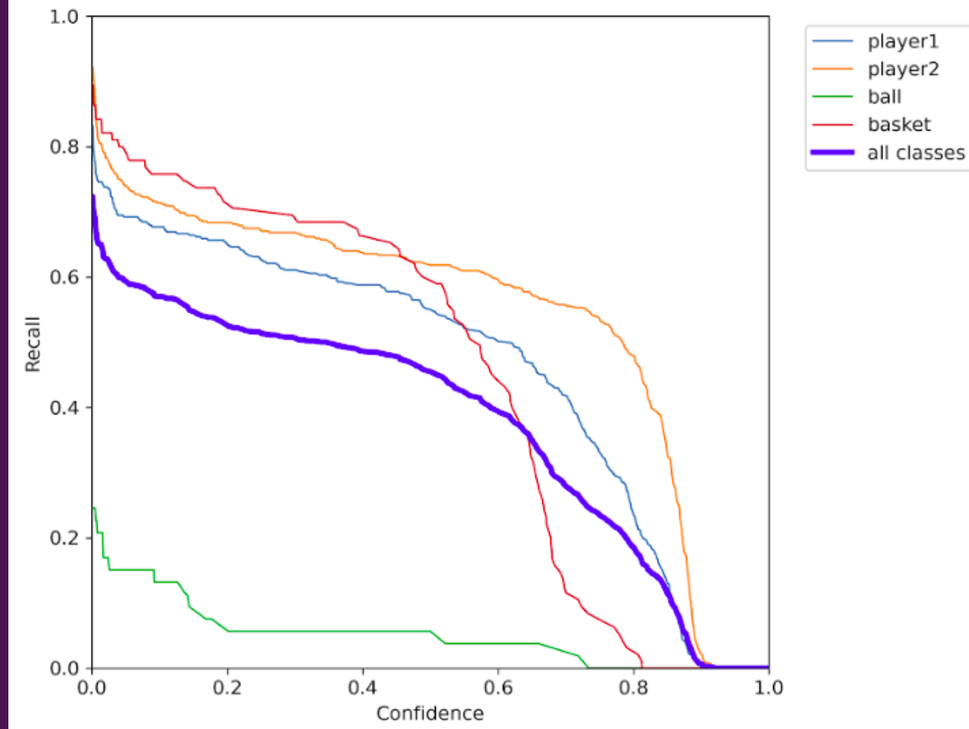

basket 0.57
player1 0.75
player1 0.74
player1 0.82
player1 0.88
player1 0.85
player1 0.88
player1 0.76
PHX 2
4 MIL
TIMEOUTS: 7
9:59
16 1st
NBA FINALS Presented by YouTubeTV
GAME 4
PHX LEADS 2-1
ESPN
Check the description. Thanks!

Results

Test dataset of 100 images

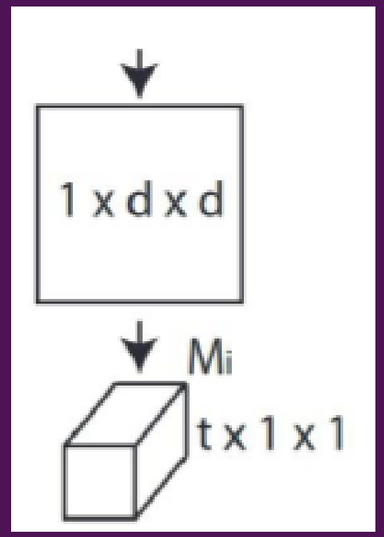
	Precision	Recall
• Overall	0.61	0.57
• Player1	0.6	0.67
• Player2	0.6	0.7
• Ball	0.35	0.13
• Basket	0.89	0.76



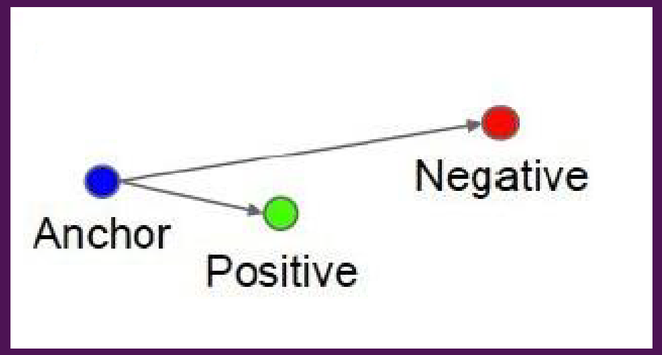


Future Improvements

Test a 2D + 1 neural network for action recognition



Deep network with triplet loss training for player identification



Improve dataset quality and size

Conclusions

Classical Approach

- It was interesting to understand how to change the parameters to obtain the best possible results
- Some tasks are still difficult to solve
- Tuning of algorithms of this type is extremely complex in the face of a chaotic environment such as that of a basketball court.

Deep Approach

- We are able to observe the power of these networks despite the poor quantity and quality of data
- The computational power required for the training of these networks is to be taken into consideration (28 hours in our case).
- It became clear in us the capability of elegantly and compactly translating complex semantics such as sports.

Thank you!

Angelo Mozzillo
Brandon Willy Viglianisi
Cristian Di Feo

Computer Vision and Cognitive System - Unimore 2021

UNIMORE
UNIVERSITÀ DEGLI STUDI DI
MODENA E REGGIO EMILIA