Signality Interview Challenge - Corner Detection
Ioannis Athanasiadis
3 March 2021

1. Problem Description:

In the context of this challenge, we were tasked to implement a system capable of recognizing the occurrence of a specific event in a football match, that of a corner. In order to do so, we were provided a sequence of data associated with how the relevant entities (e.g. players, ball) move during the football matches. Additionally, we were also provided with ground truth annotations indicating the occurrence of corners.

2. Approach:

2.1 Obtaining and visualizing the football games

At first we visualize a few games during the occurrence of corners as well as during the non occurrence of them, in order to get a more thorough understanding of the problem as well as the data that we were provided.

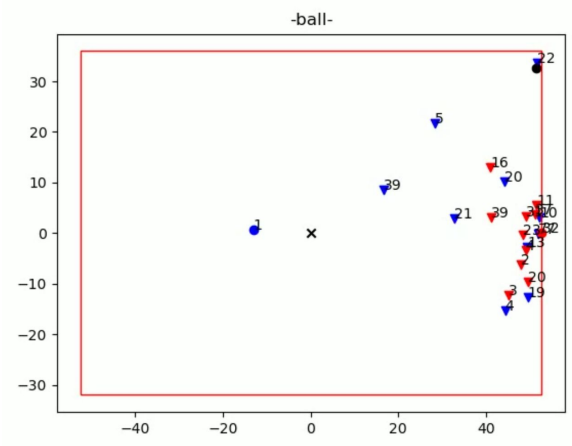


Figure 1. Visualizing corner occurrence

2.2 Observation - Identified Challenges

Although the data we were provided effectively capture the on-going action during the football match, there are instances of missing or inaccurate data. More specifically, there are cases where the position of some players or the ball are not provided. Additionally, there are also instances where the ball has been inaccurately detected. On top of that, the various states have variant milliseconds (eg. 39-41 ms) differences which does not allow for efficiently retrieving the indices referring to corners.

Another observation is that the ground truth annotations referring to corner occurrences are only provided at a specific time frame, meaning that there are identical consecutive frames, close to the annotated corner frame, that are not annotated explicitly as corners.

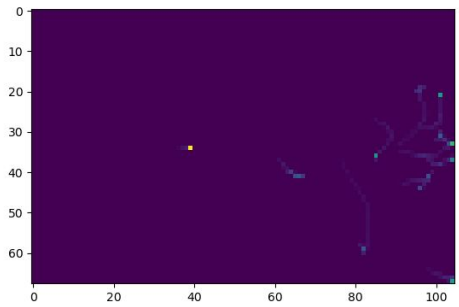
Finally, since during the corner occurrences, the way the teams are distributed across the terrain is quite distinctive, enough distrimative information is enclosed in the spatial cues for accurate corner recognition.

2.3 Dataset Construction

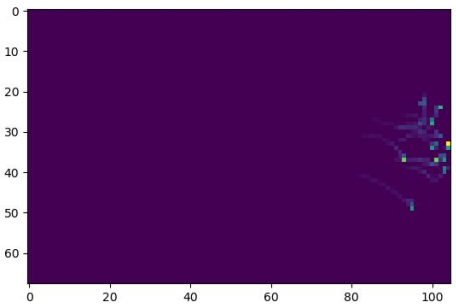
In order to turn the football games samples into training and validation datasets, we split them into 4 second-long windows that either enclose a corner or not. More specifically to generate sufficient numbers of corner samples, we employ a moving window that generates 4 second-long entities that go from 2.8 seconds before the annotated time of a corner up to 2.8 second after its occurrence.

When it comes to the non-corner occurrences, we make sure that they are sufficiently away (20 seconds) from the corner instances to avoid having ambiguities. Additionally, we aim at a balanced dataset in terms of corner to non-corner instances by generating the same number of non-corner instances to that of the corners ones.

We turn the multiple images into a single one by merging the spatial information across the obtained window frames based on how frequently a player occupied a pixel on a 68x105 grid. More specifically, we capitalize explicitly on the players' position information and create a two-channel image with each channel referring to the home and away team respectively, as shown in the figure below.



Home Team



Away Team

Figure 2. From 4-second time window to two channel image

Finally, we used the first-half of each game as a training dataset, while the second-half was used for validation purposes. In the table we report the numbers of the training and validation instances.

	Corner	**Non-Corner**
Training	228	228
Validation	327	327

Table 1. Training and validation samples for each event category

2.4 Implementation Details

We trained for 8 epochs using the VGG19 architecture utilizing the SGD optimizer with a learning rate of 10^{-3} and a weight decay parameter of 10^{-4}. Moreover, we used a batch size of 2 images per iteration while horizontally flipping the 2-channel inputs with a probability of 0.5. The intuition behind flipping is to avoid introducing any bias since both sides are equally possible to result in a corner.

2.4 Results

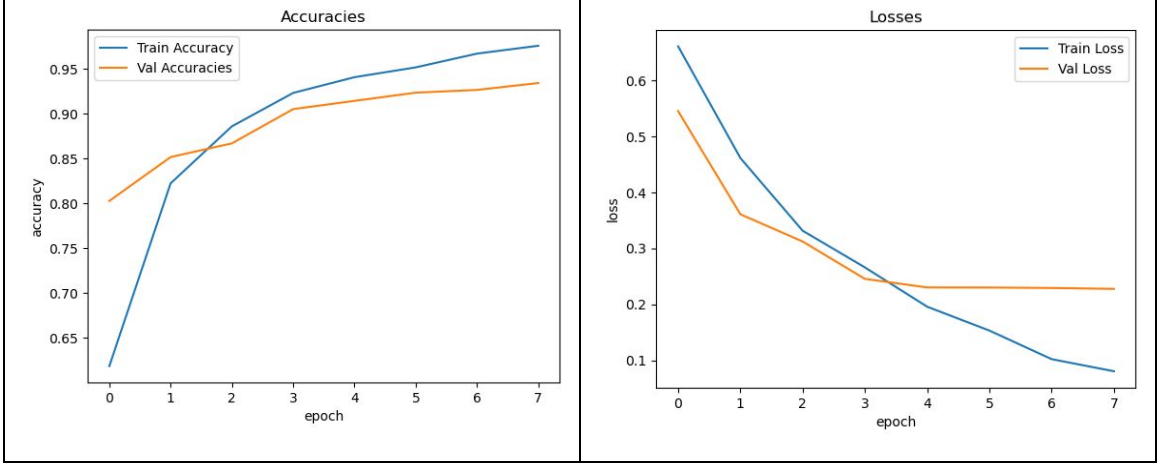


Figure 3. Loss and Accuracy graphs

The Loss and the Accuracy graphs on the balanced dataset indicate that the model is capable of grasping the information enclosed in the generated 4-second window frames and effectively classifying an event between enclosing a corner instance and not. In the table below we report the confusion matrix in the validation set.

Prediction \ Ground Truth	**Corner**	**Non-Corner**
Corner	295	11
Non-Corner	32	316

Table 2. Confusion Matrix - Validation Set