Introduction

Translating image coordinates into real-world coordinates is a very interesting application into sports. Through camera calibration and the estimation of an homography matrix this process can be done, making it possible to measure distances with confidence. But, what about developing an API that can do so?

Methods

First of all, it is interesting to know the methods already existing that can make this operation.

After some research, I have found the Graph Neural Networks that are useful for the homography matrix estimation. I think this is a complex and interesting solution that might fulfill the objective.

More examples could be the manual calibration or maybe estimating camera points using known 2D-3D correspondencies.

Datasets

As far as I am concerned, some datasets of sports fields are available on the internet, but I would try to create my own dataset to get better accuracy at the end of the process.

Evaluation

To evaluate the system, I would use some metrics such as the distance error, since it’s a metric I am familiar with, to compare those distances with a ground truth.

Conclusions

In conclusion, the first approach I would try to do is a manual calibration, then train the model and evaluate it, to have a start point, and after that I would improve the system by using GNN’s and other methods.