Supplementary Materials
Contrastive Learning for Sports Video: Unsupervised Player Classification

1. Effect of Noise in Pseudo-Labels

To evaluate effects of noise in initial pseudo-labels on embedding network performance we consider different team assignment confidence scores p_{ij} thresholds. Setting higher threshold results in selecting fewer samples in the initial training set, while using lower threshold results in having a larger but noisier sample set. Table 1 shows that selecting higher confidence threshold leads to better clustering performance.

Threshold	Error Rate	Train Set Size
$p_{ij} > 0.5$	0.134	77%
$p_{ij} > 0.7$	0.037	72%
$p_{ij} > 0.9$	0.031	60%

Table 1. Team classification error as a function of initial pseudo-labels confidence threshold. We indicate pecentage of training samples that satisfied the threshold and were included in the initial training set.

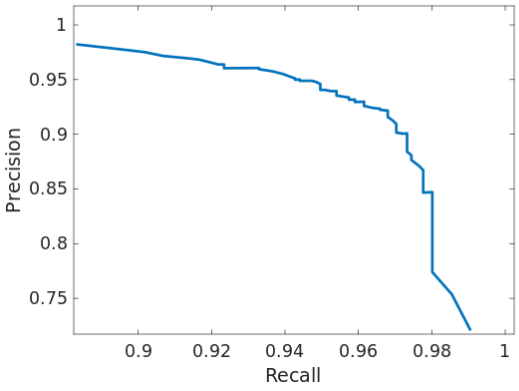

Figure 2. Precision-Recall curve of referee classifier.

2. Homography

In order to generate team position heatmaps, we use a learned homography to transfer the image coordinates of each detected player (midpoint of the bottom of each bounding box) to the corresponding point on a model of the playing surface. The homography was computed from 19 corresponding pairs of points in one video frame and in a template model of the ice rink Figure 1 shows keypoints and backprojected player positions.

3. Referee Classifier

Figure 2 shows precision-recall curve for referee classifier.

(a) Rink key points

(b) Player positions

Figure 1. Homography mapping video pixels to ice coordinates. (a) Keypoint pairs. (b) Backprojected player positions.