

Contrastive Learning for Sports Video: Unsupervised Player Classification

Maria Koshkina
York University
Toronto, Canada
koshkina@yorku.ca

Hemanth Pidaparthi
York University
Toronto, Canada
phemanth@yorku.ca

James H. Elder
York University
Toronto, Canada
jelder@yorku.ca

Abstract

We address the problem of unsupervised classification of players in a team sport according to their team affiliation, when jersey colours and design are not known a priori. We adopt a contrastive learning approach in which an embedding network learns to maximize the distance between representations of players on different teams relative to players on the same team, in a purely unsupervised fashion, without any labelled data. We evaluate the approach using a new hockey dataset and find that it outperforms prior unsupervised approaches by a substantial margin, particularly for real-time application when only a small number of frames are available for unsupervised learning before team assignments must be made. Remarkably, we show that our contrastive method achieves 94% accuracy after unsupervised training on only a single frame, with accuracy rising to 97% within 500 frames (17 seconds of game time). We further demonstrate how accurate team classification allows accurate team-conditional heat maps of player positioning to be computed.

1. Introduction

Team membership classification (i.e. labelling each person on a playing surface as a member of team A, team B or a referee) is a critical task in sports video analytics: most inferences and statistics depend upon knowing which player are on each team, including attempts on goal, offsides, and player configurations. Accurate team affiliation labels can also improve player tracking. The problem can be challenging due to the extreme variations in player pose, occlusions, motion blur and uneven illumination.

Prior work (e.g., [17, 13]) has framed the problem as a supervised learning task in which labelled data (e.g., bounding boxes with team identifiers) are used to learn a classifier. Early supervised methods employed hand-crafted colour-based features [18, 16], while more recent approaches train convolutional neural networks (CNNs) on labelled datasets to perform player segmentation [13] and classification [17].

Unfortunately, the supervised player classification approach [17] has limited application, since it requires fine-tuning on every new game for optimal classifier performance. The team segmentation approach [13] has been found to generalize better but does not provide player instance segmentation and requires expensive pixel-wise annotation to train the system. For all of these reasons, an unsupervised approach is preferred.

To date, unsupervised approaches [21, 14, 7, 4, 26] rely solely on colour-based features such as colour histograms. While these are simple and lightweight, typically many frames are needed from each new game in order to learn the colour distributions, and these methods fail when the two teams are wearing similar colours.

Our goal in this paper is to understand whether a more powerful representation, that may include both colour and configural information, can be learned in a fully unsupervised manner, and whether such a representation can reduce the number of frames needed for training and improve generalization to novel teams, jerseys, lighting and camera parameters.

To achieve this, we employ unsupervised contrastive learning to train a CNN to cluster players into two teams. We demonstrate our system’s performance on a new hockey dataset and compare it to previously proposed unsupervised team affiliation learning approaches. Figure 1 demonstrates overall system design. The dataset and code are available at <https://github.com/mkoshkina/teamId>.

Our main contributions are:

1. We introduce what is, to our knowledge, the first unsupervised deep learning approach for team classification. This novel contrastive learning approach allows us to generalize to novel games, teams and jerseys without labelled data.
2. We introduce a new annotated hockey dataset that can be used to evaluate player detection and team classification algorithms.
3. We show that our novel unsupervised algorithm outperforms prior unsupervised approaches by a large

