

Game Structure and Spatio-Temporal Action Detection in Soccer Using Graphs and 3D Convolutional Networks

Authors : Jérémie Ochin

Abstract : Soccer analytics are built on two data sources: the frame-by-frame position of each player on the terrain and the sequences of events, such as ball drive, pass, cross, shot, throw-in... With more than 2000 ball-events per soccer game, their precise and exhaustive annotation, based on a monocular video stream such as a TV broadcast, remains a tedious and costly manual task. State-of-the-art methods for spatio-temporal action detection from a monocular video stream, often based on 3D convolutional neural networks, are close to reach levels of performances in mean Average Precision (mAP) compatibles with the automation of such task. Nevertheless, to meet their expectation of exhaustiveness in the context of data analytics, such methods must be applied in a regime of high recall - low precision, using low confidence score thresholds. This setting unavoidably leads to the detection of false positives that are the product of the well documented overconfidence behaviour of neural networks and, in this case, their limited access to contextual information and understanding of the game: their predictions are highly unstructured. Based on the assumption that professional soccer players' behaviour, pose, positions and velocity are highly interrelated and locally driven by the player performing a ball-action, it is hypothesized that the addition of information regarding surrounding player's appearance, positions and velocity in the prediction methods can improve their metrics. Several methods are compared to build a proper representation of the game surrounding a player, from handcrafted features of the local graph, based on domain knowledge, to the use of Graph Neural Networks trained in an end-to-end fashion with existing state-of-the-art 3D convolutional neural networks. It is shown that the inclusion of information regarding surrounding players helps reaching higher metrics.

Keywords : fine-grained action recognition, human action recognition, convolutional neural networks, graph neural networks, spatio-temporal action recognition

Conference Title : ICSSPA 2024 : International Conference on Sport Statistics and Performance Analysis

Conference Location : Vienna, Austria

Conference Dates : December 30-31, 2024