Journal of Soft Computing Exploration

Homepage: https://shmpublisher.com/index.php/joscex

p-ISSN: 2746-7686
e-ISSN: 2746-0991

.

Enhancing soccer pass receiver prediction in broadcast images through advanced deep learning techniques: A comprehensive study on model optimization and performance evaluation

Biplov Paneru[1*], Bishwash Paneru[2], Ramhari Poudyal[3], Khem Narayan Poudyal[4]
[1]Departement of Electronics and communication, Pokhara University, Nepal
[2,4]Departement of Applied science and Chemical Engineering, Tribhuvan University, Nepal
[3]Departement of Electrical Engineering, Purbanchal University, Nepal

Article Info

Article history:

Received March 29, 2024
Revised April 2, 2024
Accepted April 29, 2024

Keywords:

Graph neural network
Soccer pass receiver prediction
Computer vision
Sports analytics
Class imbalance

ABSTRACT

In this study, we present a graph neural network (GNN) model specifically designed for football pass receiver prediction in Broadcast Images is presented in this study. Important node properties, including ball possession indicators, hot-encoded team values, and normalized ground placements, are incorporated into the model along with a careful weighting of edges to account for player distances. With weighted BCE loss used to overcome class imbalance, its architecture consists of a linear layer, numerous GNN Message Passing layers, a SoftMax activation, and binary cross-entropy (BCE) loss for training. Of 206 examples, 101 valid predictions were made, indicating a predictive accuracy of 0.50 according to the evaluation data. Comparative analyzes show that GAT-V2 (0.85) and GAT (0.63) perform better in terms of optimization and accuracy, respectively. The effectiveness in recognizing football pass receivers is demonstrated in this paper, highlighting developments in computer vision applications for sports analytics.

Corresponding Author:

Biplov Paneru
Department of Electronics and Communication Engineering,
Pokhara University, Nepal
Email: biplovp019402@nec.edu.np
https://doi.org/10.52465/joscex.v5i2.301

1. INTRODUCTION

Recent advances in computer vision and sports video analysis have produced notable gains that make a number of vital functions possible [1]. To offer more sophisticated data, such intricate and extensive analyses in sports such as badminton, cricket, basketball, and soccer, research has mostly concentrated on computer vision algorithms utilized for various applications [2]. Within the dynamic sphere of contemporary football, gaining comprehensive insight into the nuanced technical and physical demands placed on players is crucial for performance optimization [3]. Sports have become more data-driven in recent years. In competitive and professional sports, all athletes are monitored in almost every game and, if possible, also during training [2]. The convergence of sports analytics and computer vision has allowed analysis and comprehend athletic performance at a deeper level in recent years. Their location and motion data are essential for sports digitalization-related applications, such as from monitoring player skill level to even media presentation of sports [4]. Annotation and analysis of sports videos is a time-consuming task that, once automated, will provide benefits to coaches, players, and spectators [5]. One of the most well-liked and active sports in the world, soccer offers a wealth of research opportunities [6]. Football (soccer) stands as the world's most popular sport, enthralling millions of fans globally [3]. In this paper we present a novel method that includes detailed node features such as ball possession indicators, one-hot encoded team values, and normalized ground positions. The model accommodates complex spatial relationships because edges are constructed in a directed manner and their weights are inversely proportional to player distances. In addition, the model architecture consists of softmax activation, GNN message passing layers, and a linear layer, which together produce a binary cross-entropy (BCE) loss for training.

Literature Review

Professional sports events are being revolutionized by data analytics, with teams investing in data departments to get previously unattainable game knowledge [7]. Tracking the moving target is a fundamental aspect of computer vision. Its main idea is to capture moving targets fast and precisely utilizing technical techniques such as image processing and video analysis. Movement tracking is to recognize which human actions are performed, in which sequence of frames, in which time interval, and where a person acting is located in the scene [8]. In the context of sports action analysis, the use of positional data allows new developments and opportunities by taking into account players' positions over time [9].

Pi rating was used as the features in a CatBoost model, which was found to be the best option for determining the win/draw/loss probabilities at first. In particular, in this specific task, deep learning models have often been neglected. Thus, the purpose of this study authors is to evaluate the performance of a deep learning model and identify the ideal feature set for a gradient-boosted tree model. The most recent five years of data were used to train the model, and a hyperparameter grid search was performed using three training and validation. When it comes to the prediction of wins, draws and losses, the results of the validation sets demonstrate that our model outperformed previously published models of the 2017 Soccer Prediction Challenge in terms of performance and stability [10]. Scientifically evaluating soccer players represents a challenging machine learning problem. Unfortunately, most existing answers have very opaque algorithm training procedures; relevant data are scarcely accessible and almost impossible to generate [6]. The authors derive spatial relations between players and some critical locations on the field from each soccer scenario. These relations are then hierarchically aggregated within the neural architecture, which is intended to derive potentially complex game patterns from these basic relations. Convolutions are then used to effectively capture the different regularities that are present in the game. They demonstrate a very promising performance of the method in the experiments [11]. It uses a two-stage network architecture to pinpoint high-accuracy jersey number information and highlight areas of interest. Firstly, they address the player detection problem in a congested context by using an object detection network, a detection transformer. Second, they use a secondary convolutional neural network to recognize player jersey numbers in order to identify them, and they synchronize this with a game clock subsystem. To index plays, the system outputs a complete log into a database [12]. The authors describe a "role-based" representation that dynamically modifies the relative roles of each player at every frame. They show how this captures the immediate context to facilitate analysis of individual players and team members. Using a minimum entropy data partitioning technique, we identify the role directly from the data and demonstrate how this can be applied to precisely detect and visualize formations, as well as analyze the behavior on an individual basis [13]. In the experiment, a completely automated pass device that can release passes upon command. An audio and video sequence of the particular catch attempt was recorded after the pass was released. An audio-visual recording system that was integrated into the passing machine was created specifically for this purpose. An audio and video data set containing 2276 recorded catch attempts is created using this system. For feature extraction, a convolutional neural network (CNN) is employed, and the video data is classified using a downstream long- and short-term memory (LSTM). CNN in one dimension is used for audio data classification [2].

Although deep learning and computer vision have been used recently to increase soccer statistics, there is still a significant research gap to precisely predicting pass recipients during gameplay. While previous research has examined a variety of approaches, including object identification and graph-based modeling, very few have fully integrated these methods to represent the complex dynamics of soccer interactions. Our suggested methodology closes this gap by offering a comprehensive system that incorporates perspective modification, object identification, team-level understanding, and pass receiver prediction into one framework. Our research aims to offer a more comprehensive understanding of soccer gameplay by addressing these interwoven components. This will provide valuable information for players, coaches, and analysts alike. The proposed model utilizes two different models along with YOLO to detect the models along with the GNN architecture mainly focused on predicting the passing of soccer by footballer.

2. METHOD

There are four main steps in our multistage system that we have proposed for Football/Soccer Pass Receiver Prediction. To start, the YOLO algorithm is used to detect players and balls simultaneously, establishing the foundation for real-time object localization. The YOLO model makes predictions with a single network evaluation, making this method extremely fast, running in real time with a capable GPU [14]. After that, team identification is achieved by clustering, which improves the system's capacity to identify individual players and their teams. The detected entities are transformed into an aerial view in the third stage of the perspective transformation, creating a comprehensive and contextually rich representation of the playing field.

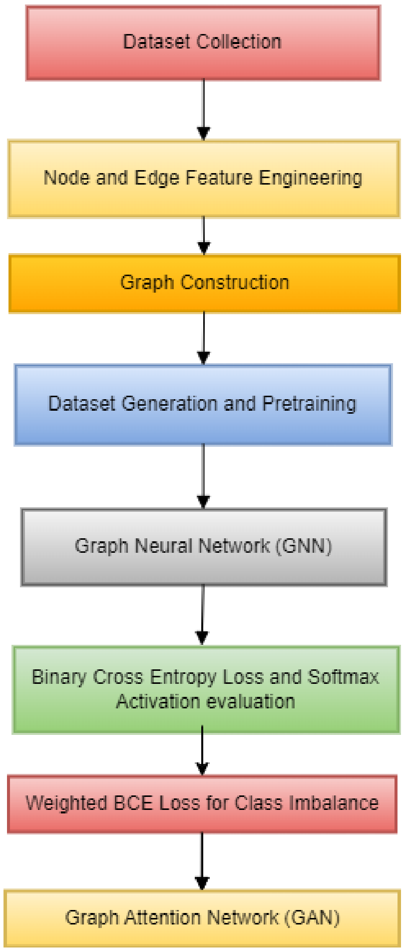

Figure 1. Proposed workflow

As seen in Figure 1, we suggest a multi-phase football/soccer pass receiver prediction system. First, we gather about 200 photos from different soccer leagues to create a varied dataset. The process of carefully designing features such as possession indications and team values is known as Node and Edge Feature Engineering. By employing methods like viewpoint transformation and clustering, graph construction generates a directed graph that records player interactions. Pre-training and dataset generation provide noise to the dataset to increase its durability. By embedding node features into a high-dimensional space, the graph neural network stage improves comprehension of player interactions. Probability distributions are produced via binary cross-entropy loss and Softmax Activation, while class imbalance is addressed by weighted BCE Loss. Pass Receiver Prediction is handled by the Graph Attention Network, which provides a detailed representation of the dynamics of soccer games.

Data Collection

Figure 2. Sample dataset image

Data is the primary goal of a proposed work [15]. We carried out a thorough process of data collection, compiling a varied collection of about 200 photos from different league videos, including but not restricted to well-known contests like the Europa League. With great care, this effort sought to create a manual data set that was comprehensive and specifically designed for the research task at hand, namely prediction of the football pass receiver.

Node and Edge Feature Engineering

Our method is based on meticulous engineering of node features. We include important components like a binary possession indicator, one-hot encoded team values, and normalized ground positions. Additionally, edge weights are meticulously assigned, being inversely proportional to player distances, thereby capturing the spatial relationships crucial for pass receiver prediction.

Graph Neural Network

Graphs are a kind of data structure which models a set of objects (nodes) and their relationships (edges) [16]. A graph neural network, a potent framework for discovering patterns in graph-structured data, is integrated into our model architecture. A linear layer is used in the first step to embed the node features into a high-dimensional space. GNN message passing makes up the following layers, which allow the model to repeatedly compile data from nearby nodes, improving its comprehension of intricate player interactions.

Binary Cross-Entropy Loss and Softmax Activation

To produce probability distributions, a softmax activation is applied to every node after the GNN layers. Because the task is binary (i.e., it determines whether a player is a pass receiver or not), the objective function used is the loss of binary cross entropy (BCE). This decision makes sense given the nature of our prediction task and makes model training more effective.

Weighted BCE Loss for Class Imbalance

Acknowledging the intrinsic class imbalance present in the dataset, we incorporate weighted BCE loss into the training process. By making this adjustment, biases resulting from an unequal distribution of classes are mitigated, and the model is guaranteed to be equally sensitive to positive and negative instances.

Graph Construction

The process uses several methods, such as distance matrix computation, clustering, and perspective transformation. The teams in the image are identified by clustering, and the true ground X and Y coordinates are obtained through the perspective transformation. The ball possession is then determined using the distance matrix. After that, a directed graph like the one below is created using these values and fed into the graph neural network.

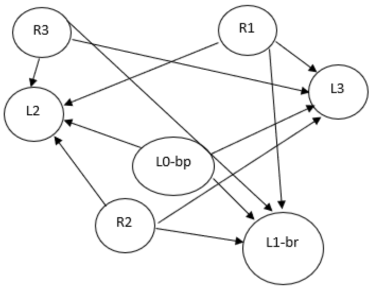

Figure 3. Graph arcitecture

Dataset Generation and Pretraining

The dataset is an essential part of our process. Approximately 200 manually created samples make up the initial set, which we then add noise to to reach approximately 1500 samples for pretraining. After being pre-trained, this model is refined using a separately created dataset of about 3000 samples, adding more noise to boost robustness.

Table 1. Dataset division

Dataset	Distribution
Manually created images	200
Noised images for pretraining	1500
Dataset images	3000

Lastly, a graph attention network is used to build a player graph and perform Pass Receiver Prediction (GATv2). With the help of sophisticated graph-based modeling, this multifaceted approach seeks to capture the complex dynamics of soccer gameplay, starting with object detection and progressing to team-level comprehension, perspective transformation, and pass receiver prediction.

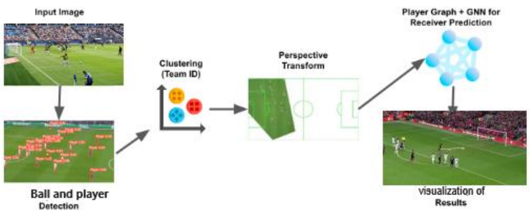

Figure 4. Proposed architecture

3. RESULTS AND DISCUSSIONS

Strong predictive abilities were displayed by the model architecture, which included a Linear Layer for embedding five node features into N dimensions, followed by K layers of Graph Neural Network (GNN) Message Passing Layers. When combined with binary cross-entropy (BCE) loss, softmax activation on all nodes worked well for the binary classification task of identifying pass receivers.

Figure 5. Broadcast image

Our model applied Weighted BCE loss to rectify the imbalance in classes, demonstrating its flexibility on datasets with non-uniform class distributions. The overall performance was enhanced by the introduction of this technique, which was crucial in guaranteeing the model's sensitivity to both positive and negative instances.

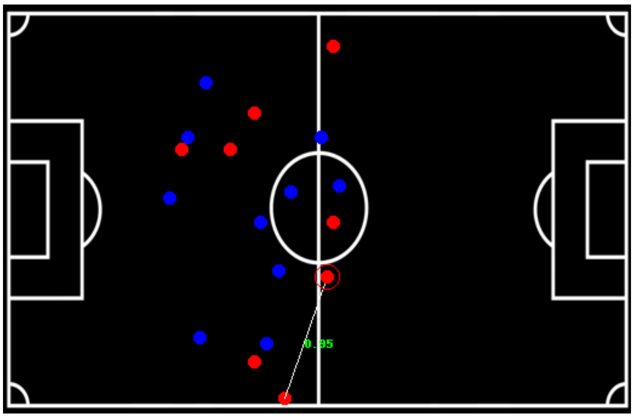

Figure 6. Graph v2 prediction

Quantitatively, our model, called graph-v2, demonstrated an impressive 86% accuracy rate in player movement and pass detection prediction. The model predicted pass outcomes correctly in 256 out of 300 images. This high accuracy highlights how well the selected architecture and training approach capture the subtle dynamics of playing soccer.

The performance of the graph-v2 model was superior to that of other GNN designs. Approximately 50% of the correct predictions were made using the GCN model, compared to 63% using the GAT model. Although these results are acceptable, the graph-v2 model's 86% accuracy in soccer pass receiver prediction stands out as a noteworthy development.

The predictive performance, as demonstrated by the model prediction in Figure graph-v2, is evidence of the efficacy of the selected architecture and training plan. The visual representation of model predictions provides valuable insight into the model's ability to capture complex player movements and predict successful passes accurately.

4. CONCLUSION

To sum up, the remarkable success of our graph-v2 model - achieving an impressive 86% accuracy rate - in predicting player movements and pass outcomes in soccer images highlights its importance as a reliable tool for sports analytics applications. This high degree of accuracy shows how well the model can represent the complex dynamics of player interactions in real-time gameplay scenarios. Outstanding performance has important implications for coaching and player development decision-making in addition to improving our knowledge of soccer tactics.

Because the model can pinpoint pass receivers so precisely, coaches will be able to examine and optimize player positioning, passing patterns, and team formations. This makes the model a promising tool for strategic planning. Furthermore, the model's insights aid in a deeper understanding of team dynamics, enabling customized training plans and tactical modifications.

The suggested football/soccer pass receiver prediction system has wide and bright future possibilities. First, investigating more sophisticated model architectures than the Graph Attention Network (GATv2) could reveal more possibilities for improving prediction efficiency and accuracy. Moving the system toward real-time implementation is an important development path. Hardware accelerators and optimization strategies are needed to guarantee quick analysis during live soccer matches.

An important improvement is the integration with player tracking data, which offers deeper insight into player movements and interactions. Furthermore, the system's comprehension of the soccer field environment may be enhanced by the addition of semantic segmentation techniques, leading to more complex predictions. The ongoing growth and diversification will continue to be essential to guaranteeing the model's flexibility across a variety of playing situations and styles. All things considered, the continued development has the potential to improve soccer analytics and give analysts and coaches insightful information in real time.

CREDIT AUTHORSHIP CONTRIBUTION STATEMENT

Author1: Conceptualization, Methodology, Software, Project administration. **Author2**: Software–model validation, original draft. **Author3**: Writing – review & editing. **Author4**: Supervision

DECLARATION OF COMPETING INTERESTS

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

DATA AVAILABILITY

Data will be made available on request.

REFERENCES

[1] B. T. Naik, M. F. Hashmi, and N. D. Bokde, "A Comprehensive Review of Computer Vision in Sports: Open Issues, Future Trends and Research Directions," Mar. 2022, doi: 10.3390/app12094429.

[2] B. Hollaus, B. Reiter, and J. C. Volmer, "Catch Recognition in Automated American Football Training Using Machine Learning.," *Sensors (Basel).*, vol. 23, no. 2, Jan. 2023, doi: 10.3390/s23020840.

[3] X. Li and R. Ullah, "An image classification algorithm for football players' activities using deep neural network," *Soft Comput.*, vol. 27, no. 24, pp. 19317–19337, Dec. 2023, doi: 10.1007/s00500-023-09321-3.

[4] J. Liu, G. Huang, J. Hyyppä, J. Li, X. Gong, and X. Jiang, "A survey on location and motion tracking technologies, methodologies and applications in precision sports," *Expert Syst. Appl.*, vol. 229, p. 120492, Nov. 2023, doi: 10.1016/j.eswa.2023.120492.

[5] J. Newman, A. Sumsion, S. Torrie, and D.-J. Lee, "Automated Pre-Play Analysis of American Football Formations Using Deep Learning," *Electronics*, vol. 12, no. 3, p. 726, Feb. 2023, doi: 10.3390/electronics12030726.

[6] P. Garnier and T. Gregoir, "Evaluating Soccer Player: from Live Camera to Deep Reinforcement Learning," Jan. 2021, [Online]. Available: http://arxiv.org/abs/2101.05388

[7] C. Wisdom and A. Javed, "Machine Learning for Data Analytics in Football: Quantifying Performance and Enhancing Strategic Decision-Making," 2023, doi: http://dx.doi.org/10.2139/ssrn.4558733.

[8] K. Host and M. Ivašić-Kos, "An overview of Human Action Recognition in sports based on Computer Vision," *Heliyon*, vol. 8, no. 6, p. e09633, Jun. 2022, doi: 10.1016/j.heliyon.2022.e09633.

[9] S. Barbon Junior *et al.*, "Sport action mining: Dribbling recognition in soccer," *Multimed. Tools Appl.*, vol. 81, no. 3, pp. 4341–4364, Jan. 2022, doi: 10.1007/s11042-021-11784-1.

[10] C. Yeung, R. Bunker, R. Umemoto, and K. Fujii, "Evaluating Soccer Match Prediction Models: A Deep Learning Approach and Feature Optimization for Gradient-Boosted Trees," *arXiv e-prints*, p. arXiv:2309.14807, Sep. 2023, doi: https://doi.org/10.48550/arXiv.2309.14807.

[11] O. Hubáček, G. Šourek, and F. Železný, "Deep Learning from Spatial Relations for Soccer Pass Prediction," in *Machine Learning and Data Mining for Sports Analytics*, 2019, pp. 159–166. doi: 10.1007/978-3-030-17274-9_14.

[12] H. Liu *et al.*, "Automated player identification and indexing using two-stage deep learning network," *Sci. Rep.*, vol. 13, no. 1, p. 10036, Jun. 2023, doi: 10.1038/s41598-023-36657-5.

[13] A. Bialkowski, P. Lucey, P. Carr, Y. Yue, S. Sridharan, and I. Matthews, *Large-Scale Analysis of Soccer Matches Using Spatiotemporal Tracking Data*. 2014.

[14] Z. Dahirou and M. Zheng, "Motion Detection and Object Detection: Yolo (You Only Look Once)," in *2021 7th Annual International Conference on Network and Information Systems for Computers (ICNISC)*, IEEE, Jul. 2021, pp. 250–257. doi: 10.1109/ICNISC54316.2021.00053.

[15] M. Rabeya, M. R. Tuly, M. S. Mahmud, and A. Sattar, "Sentiment Analysis of Bengali Textual Comments in Field of Sports Using Deep Learning Approach," in *2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, Oct. 2022, pp. 1–8. doi: 10.1109/ICCCNT54827.2022.9984326.

[16] J. Zhou *et al.*, "Graph neural networks: A review of methods and applications," *AI Open*, vol. 1, pp. 57–81, 2020, doi: 10.1016/j.aiopen.2021.01.001.