Predicting Volleyball Events using Graph Neural Networks

Alex Rasla, Rhys Tracy

Abstract

For our class project, we decided to explore the effect of using GNNs to predict the outcome of volleyball rallies. Graphs have an exceptional ability of being able to take into account various attributes into their nodes and link together sets of the attributes through their edges. Thus, in theory, a graph that has node attributes that consist of characteristics of a rally such as the locations and types of hits and sets with consecutive nodes linked together, should be able to accurately predict the outcome of a rally. In this project, we use a novel dataset and compare several GNN models to baseline approaches for predicting the outcome of a given volleyball rally. We found that for the most part, the graph approaches do improve on baseline approaches. Further, we were able to apply the same GNN models for rally predictions to the task of predicting where a setter will set the ball during a volleyball rally. This is a completely novel task that was able to generate some impressive results, as critiqued by volleyball experts.

1 Introduction

Millions of people everyday watch and play sports for entertainment, profession, or exercise. There are many different forms of sport, each of which are very competitive in their own nature. Some allow physical contact, some are played on dirt, some require pads, but all of them have a very strict rule set and strategies that have been established and fine-tuned over the lifespan of the sport itself. Oftentimes, these strategies take advantage of certain aspects of the game or the players such as the speed, strength, height, location, etc. Using this intuition, we believe an interesting research area to explore is being able to collectively analyze these aspects of the game, and create a model to predict outcomes and winners of parts of the game.

With this topic in mind, we decided to explore the aspects and predictable nature of volleyball. Volleyball rallies have very discrete outcomes in terms of which team wins a rally, how many touches a team can take per "possession", and what each touch represents in the processes of getting the ball over the net (pass/receive, set, hit). These different attributes can be organized and filtered in a variety of ways in order to capture the current states of the game. Another reason for exploring volleyball is that there has not been a substantial amount of research or exploration for it. Most of the research that has gone into the field of using neural networks for sports-related tasks has been for more popular sports such as soccer or basketball [1], [6], [10], and mainly focus on Computer Vision approaches. Thus, we decided to focus on being able to predict the outcome of rallies and set locations for volleyball.

In order to do this, we used a novel dataset collected, filtered, and processed by Rhys that contains detailed information for rallies from 10 NCAA Division 1 and professional volleyball games from 2020 to 2021—including Big West Conference matches and national team matches. For our models, we decided to test the effectiveness of being able to generate predictions using a few different types of Graph Neural Network (GNN) models including Graph Convolution Networks (GCN) [4] and Graph Transformers [11]. Graphs are able to very efficiently and effectively generate relationships between the different nodes and the attributes within them, and create weighted edges between different nodes within a themselves. Then GNNs are able to capture and understand the underlying relationships within a graph. Thus, if given enough relevant information regarding aspects of a rally and a well-structured graph format for that information, GNN models should theoretically be able to outperform other types of models and approaches such as CNNs or RNNs.

2 Related Work

Previously, there have been several attempts that use neural networks for sports-related applications. Some of these studies use GNNs as their neural networks, and others use CNNs or RNNs. The most similar study to ours attempts to use Graph Neural Networks to also predict sports outcomes [9]. This paper focuses on creating a unique way of representing game-states using graph approaches. Through this technique, they are able to capture inter-player relationships and local player relationships that can otherwise not be taken into account when training a model. Further, the paper tests its player-specific graph approach on American football and a popular esports game, Counter-Strike. Through their methods, they demonstrate a reduction in test set by 20% and 9% for football and Counter-Strike respectively.

A similar work to [9] uses GNNs for predicting future player locations and movements [10]. The work focuses on multi-agent sports such as basketball and soccer and takes advantage of both GNNs and Variational RNNs to generate these future locations. Through these techniques, they show that the statistical player distribution of their generative model predictions outperform previous works. Further, they also test their model using conditional prediction to answer questions such as: *How will the players move if A passes to B instead of C?*.

A comprehensive review of sports-related applications that utilize machine learning techniques can be found in [2], [3]. As we can see, there are a variety of avenues where neural networks can be used for sports-related applications; however, there are few existing applications for volleyball predictions, the most notable being [5], [7], [8]. None of these approaches for volleyball predictions take advantage of graph based structures or GNNs.

3 Proposed Methods

In our experiments, we analyze two different applications of GNNs on our dataset: predicting rally outcomes and predicting set locations. These two applications involve different graph structures, different target formats, and different losses and metrics; however, we will test the same GNN architectures for both. The two applications will both make use of the labeled dataset of 9 full NCAA Division 1 Men's Volleyball matches and one full professional match from the corresponding 2020-2021 seasons made using Volleyball Rally Expression Notation (VREN).

3.1 Underlying Data Representation

VREN defines a rally as a sequence of rounds with rally level encoding information (winning team, winning reason, losing reason). VREN defines a round as a set of 1-3 contacts (pass-set-hit) on one side of a net with a number of attributes for each contact; a round ends and a new round starts whenever the ball crosses the net or a block occurs. VREN includes all the most vital information of a volleyball rally according to volleyball experts.

3.2 Graph Representation for Rallies

In order to test the GNN approaches, we first processed the dataset into a set of graphs. For rally outcome predictions, we separate each round into an individual graph. Each of these graphs include 4 nodes: one for a pass (first contact of round), one for a set (second contact of round), one for a hit (third contact of round), and one for a block touch (potential fourth legal contact per round by blockers of the opposite team). Each of these nodes includes all the VREN attributes for that given contact. Our graph structure only consists of edges between consecutive contacts to offer a temporal encoding for the round. The target for each round's graph is 0 if team A (defined as the home team by VREN) eventually loses that rally or 1 if team A eventually wins that rally.

3.3 Graph Representation for Set Locations

For the set location predictions, the graphs and targets are slightly different. Each round again is its own graph, but the graph only contains information prior to the set location such as the hitting and blocking contacts of the previous rounds. Each of these graphs include 2 or 4 nodes depending on if there was a previous round of the current rally: one for the previous round's hit node (third contact of previous round if exists), one for the previous round's block node (potential fourth legal contact by blockers of the opposite team of previous round if exists), one for the current round's pass (first contact of current round), and one for the current round's set (second contact of current round). The set node for the current round includes only location encoding for where the setter will set the ball from (pass landing location). It includes no other attributes for that contact that would allow the model to see into the future. All other nodes include all of the VREN attributes for that given contact. Similar to the rally prediction graphs, the graphs involve edges connecting only consecutive contacts to give a type of temporal encoding to the round information. The target for each graph is an array of 9 elements corresponding to the 9 possible values for setting location defined by VREN. This target array will be one-hot encoded to the true value of the set location. To create these graphs and store these two graphs datasets, we use pytorch-geometric.

3.4 GNN Models

Once the graph datasets are set up, we can then train and test our models. For both applications, we use the same three general GNN structures, all of which were made using pytorch and pytorch-geometric. First, we used a simple GCN with one graph Conv1D Layer, one graph global pooling layer to get a graph level value, and one output Linear Layer. Second, we used a slightly more complex GCN model (LargeGCN) with one graph Conv1D Layer and one graph global pooling layer again, but with 2 Linear Layers with 64 and 128 hidden neurons before the output Linear Layer. Lastly, we created a Graph Transformer model with one graph TransformerConv Layer, a global pooling layer, and one output Linear Layer. The rally outcome models had a single float output with a sigmoid activation function on the final output layer to yield a probability value between 0 and 1; these models used MSE as their loss function for training, and MSE, binary accuracy, AUC, and Brier Score as metrics for validation and testing. The set location models had an array of 9 values as an output with a softmax activation function on the final output layer to yield probability predictions for each set location that sum to 1; these models used the cross-entropy loss function for training, and cross-entropy loss and categorical accuracy as metrics for validation and testing.

Table 1: *Rally Outcome Prediction Performance: Baseline Models vs New GNN Models*

Type of Approach	Model	Validation Game	AUC	Brier Score	Binary Accuracy
Baseline/NLP	CNN	NCAA Game	0.746	**0.210**	0.678
		Professional Game	0.748	0.203	0.699
	Transformer	NCAA Game	**0.830**	**0.170**	**0.769**
		Professional Game	0.851	0.155	0.755
GNN	Base GCN	NCAA Game	0.780	0.219	0.707
		Professional Game	0.800	**0.199**	0.713
	Large GCN	NCAA Game	**0.785**	0.225	**0.714**
		Professional Game	**0.807**	0.206	**0.740**
	Graph Transformer	NCAA Game	0.825	0.184	0.752
		Professional Game	**0.860**	**0.147**	**0.790**

4 Experimental Results

Overall, results for both experiments were positive. Rally outcome prediction results can be seen in Table 1 and set location prediction results can be seen in Table 2.

4.1 Rally Outcome Predictions

By using graph representations and GNNs, we were able to see improvements over baseline NLP approaches. Both GCN approaches seemed to have noticeably better AUC scores and accuracies with around the same or slightly worse Brier Scores as compared to baseline for both the NCAA and Professional validation games. Overall it would appear that the GCN approaches yielded better predictions than the baseline CNN approach. Further, Large GCN performed slightly better than the smaller GCN for AUC and accuracy, but slightly worse with BS for both the NCAA and pro validation games. Thus, it appears that adding a little more complexity to the model can make a GCN perform slightly better overall in this case. Lastly, the Graph Transformer model performed noticeably better than baseline for all metrics when looking at the professional validation game, but slightly worse in all metrics when looking at the NCAA validation game. Overall it would appear that encoding the volleyball rally information into graph structures yielded improvements for predicting rally outcomes. Lastly, all 3 GNN models saw better performance on the pro game than the NCAA game for all metrics, exactly like the baseline approaches.

4.2 Set Location Predictions

For predicting set locations, GNNs and graph structures yielded promising results, even though they may seem poor at first glance. For predicting set location, 43-49% (Table 2) accuracy is good since there are 9 possibilities for set location in VREN. More importantly, it is almost impossible to guess where a high level setter will set for a large majority of situations, even for top professional volleyball players. We see that smaller GCN did perform slightly better than Large GCN on the NCAA validation game and exactly the same on the pro validation game, so it appears that simpler is better when it comes to the GCN model structure. We can also see that the Graph Transformer performs noticeably better than both GCNs, as was expected.

Table 2: *Categorical Accuracy for setting location predictions*

Model	Validation Game	*Categorical accuracy%*
GCN	NCAA Game	44.2
	Professional Game	45.6
Large GCN	NCAA Game	43.3
	Professional Game	45.6
Graph Transformer	NCAA Game	**48.3**
	Professional Game	**49.1**

5 Discussion

As shown in the analysis above, using GNN approaches for volleyball rally prediction yielded clear overall improvements for both the NCAA and pro validation games. Additionally, the GNN approaches worked well for predicting setter locations and offer a baseline for future work.

5.1 Rally Outcome Predictions

The above results suggest that graph based approaches can better model a volleyball rally than a pure NLP approach at least when using VREN as an underlying data representation. This is likely because the graph structure we used could specify the consecutive nature of the different contacts in a round and add an encoding of grouped attributes for certain contacts that the baseline NLP approach could not. This allowed the GNN to capture the underlying distributions of a volleyball rally better. On the other hand, the Graph Transformer model had more situational improvements. The results showed that Graph Transformer yielded a slight but noticeable improvement on the pro validation game for all metrics over the baseline

Transformer model, but saw worse performance on all metrics for the NCAA game. This suggests that a graph-based encoding is able to analyze more information of the underlying relationships in professional play that a baseline Transformer model is unable to distinguish. Further, it shows that this information may not be as significant in NCAA play, or that it may be distinguished by the base NLP Transformer model easier. The final takeaway from the volleyball rally outcome prediction experiments is that similar to the baseline approaches, each model saw better performance for all metrics on the pro validation game as compared to the NCAA validation game. This makes sense because professional players are more disciplined and skilled, so outcomes should be more deterministic and predictable. Therefore, a well trained model should be better able to understand and predict the relationships in a pro game. This is also likely the reason behind Graph Transformer not yielding improvements on the NCAA validation game: more randomness in NCAA games could lead to more difficulty for complex models to learn. The complexity of the Graph Transformer's input encoding could actually be counterproductive when predicting the NCAA games, but with more training epochs, better optimization, and more data, the Graph Transformer can likely beat out the baseline Transformer model on NCAA game performance as well.

5.2 Set Location Predictions

For our second set of experiments, we saw all 3 GNN models have promising results. Base GCN performed slightly better than Large GCN on the NCAA validation game, but the same on the pro game. Further, Graph Transformer performed better on both types of games. Prior to running these experiments, we expected the models to struggle with predicting set locations; we expected these models would not be able to perform better than a random guess between the 5 most common values of set location (20% accuracy) or to primarily guess the most commonly occurring set (outside = 32% of all set location values). Even an expert volleyball analyst or high-level professional player would have trouble predicting the set location in most situations, aside from rare cases where the team's passes are very poor. Good professional middle blockers—the players that are the best at following sets and getting in front of the ball before a hitter hits—almost never guess the set location because of the difficult nature of it; instead, they react to the trajectory of the set. Most volleyball experts, analysts, and players would agree that being able to correctly predict 43-49% of all sets in a volleyball match is impressive, and certainly better than the results that were expected. And even if the performances are not as excellent as some might hope, this experiment still provides a useful baseline for future comparison and improvement.

5.3 Final Takeaways

Overall it would appear that graph-based approaches are well suited for volleyball sports predictions and can beat out NLP approaches. We would also like to mention that our GNN models did not include any form of hyperparameter tuning and do not make use of any optimization techniques, while the baseline NLP approaches do have hyperparameter tuning, optimization, and early stopping that boosted their performance significantly. With more time and experimentation, as well as fully optimized GNN models, we believe we will see even better results than we currently have.

6 Future Work

In the future, we would like to expand upon this project in several ways. First, we would like to optimize our models by adding hyperparameter tuning, early stopping, and more epochs of training. With these additions, we will likely see more improvements to the point where these GNN approaches would be clearly superior to baseline approaches. Second, we would like to improve the graph structures we have by including player locations or other useful attributes (ie. starter or not) to give the graphs more information to build its relationships with. Another interesting avenue to explore would be predicting hit types, predicting overall game outcomes, and creating and analyzing graph embeddings for events and rallies. Lastly, we could improve upon current work by gathering more data across different leagues and levels of play to further test the robustness of these models.

References

[1] Determining the phases of play using graph neural network embeddings. (Cited on page 1.)

[2] BUNKER, R. P., AND THABTAH, F. A machine learning framework for sport result prediction. *Applied Computing and Informatics 15*, 1 (2019), 27–33. (Cited on page 2.)

[3] HORVAT, T., AND JOB, J. The use of machine learning in sport outcome prediction: A review. *WIREs Data Mining and Knowledge Discovery 10*, 5 (2020), e1380. (Cited on page 2.)

[4] KIPF, T. N., AND WELLING, M. Semi-supervised classification with graph convolutional networks, 2016. (Cited on page 1.)

[5] LIU, Q., LIU, Q., AND BABAR, M. Prediction of volleyball competition using machine learning and edge intelligence. *Mob. Inf. Syst. 2021* (jan 2021). (Cited on page 2.)

[6] NAIK, B. T., HASHMI, M. F., AND BOKDE, N. D. A comprehensive review of computer vision in sports: Open issues, future trends and research directions. *Applied Sciences 12*, 9 (apr 2022), 4429. (Cited on page 1.)

[7] SANGHVI, D., DESHPANDE, P., SHANBHOGUE, S., AND SHAH, V. Analyzing and predicting ncaa volleyball match

outcome using machine learning techniques. (Cited on page 2.)

[8] WENNINGER, S., LINK, D., AND LAMES, M. Performance of machine learning models in application to beach volleyball data. *International Journal of Computer Science in Sport 19* (07 2020), 24–36. (Cited on page 2.)

[9] XENOPOULOS, P., AND SILVA, C. Graph neural networks to predict sports outcomes. In *2021 IEEE International Conference on Big Data (Big Data)* (2021), pp. 1757–1763. (Cited on page 1.)

[10] YEH, R. A., SCHWING, A. G., HUANG, J., AND MURPHY, K. Diverse generation for multi-agent sports games. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2019), pp. 4605–4614. (Cited on page 1.)

[11] YUN, S., JEONG, M., KIM, R., KANG, J., AND KIM, H. J. Graph transformer networks, 2019. (Cited on page 1.)