Player Movement Predictions Using Team and Opponent Dynamics for Doubles Badminton

No Author Given

No Institute Given

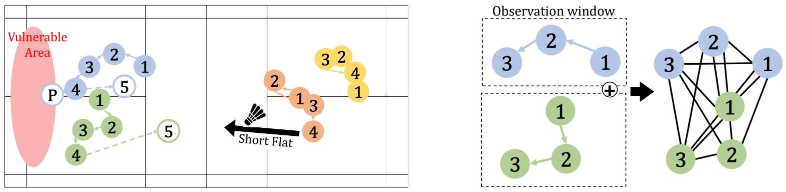
Abstract. With the advancement of AI, strategic analysis in team sports has become increasingly valuable. In doubles badminton, the challenge lies in accurately predicting player movements in a fast-paced, dynamic environment. To tackle this, we introduce the novel TS-CVAE specifically designed for doubles badminton. TS-CVAE incorporates Team GAT, which leverages team influence graph over a few strokes to capture rapidly changing team strategies, and Opponent GAT, which holistically analyzes interactions between opposing players. Additionally, a new data augmentation module, SOSA, enhances the understanding of player positioning and strategy by incorporating singles data. Experimental results on real doubles badminton data show that TS-CVAE outperforms state-of-the-art sport forecasting models across multiple evaluation metrics. Visualized results also confirm that TS-CVAE's predictions closely align with the ground truth.

1 Introduction

In the recent decade, predicting movements with AI has been a game changer that provides coaches and players valuable insights to refine techniques and tactics [5, 11]. Team sports can be categorized into net sports (e.g., badminton and tennis) and full-court games (e.g., basketball and football), with moving patterns and rules differentiate them. In full-court sports, every player can move freely across the whole field, and there are specific zones to score. In net sports, teams are divided by a net, and the scoring zones are anywhere on the opponent's side. A line of research [6, 10] analyzed multi-player dynamics in full-court games. Yet, net sports, including badminton, have not drawn enough attention.

This paper focuses on doubles badminton. Although there is emerging literature on singles badminton [13, 14, 4, 12], generalizing it to doubles scenario is not trivial since the concept of *teammates* does not exist in singles. Applying singles models to doubles scenario may overlook the importance of team chemistry and strategy during learning and forecasting. Few studies were specifically proposed for doubles badminton, and those that did [11] actually failed to finely formulate team strategy and opponent competitions, which are crucial in doubles games.

Figure 1(a) presents player locations during a rally. Colors distinguish each player. A sequence marks the movement of each player stoke-by-stroke with numbers. The figure reveals that different strategies result in distinct player movements and positions. Defensive strategies position players on the left court



(a) Player locations. (b) Team influence graph.

Fig. 1. Examples of doubles badminton scenarios.

(blue and green) in parallel to cover a wider area, while offensive strategies position players on the right court (yellow and orange) vertically to the net to maximize their effectiveness in attacking. In contrast, singles players do not have such tactical considerations and positioning with teammates.

Given the input of four-player movements in doubles, the goal of this paper is to forecast the movements of all four players, shot types, and players executing each shot. The problem also comprehensively incorporates team strategies, competitive reactions against opponents, and sequential relationships involving both teammates and opponents. Two key factors need to be addressed:

1. **Team Strategy.** Since the pace of badminton is fast, doubles players must rapidly pursue the shuttle while simultaneously covering their teammates, leading to significant changes in player locations and team strategies from stroke to stroke. Failing to consider positional impacts from teammates may result in inaccurate predictions of both players' locations.
2. **Opponent Influence.** Opponent reactions strongly affect team decision-making. For instance, teams may capitalize on taking front-and-back positioning to execute many smashes and drives when their opponent makes a mistake by raising the shuttle high. However, analyzing the tactical influence among all strokes against each other, which may form a complete influence graph, could be challenging due to scalability issues.

To address these factors, we propose a novel model, namely *Team Strategy-aware CVAE* (*TS-CVAE*), built with Conditional Variational Autoencoder (CVAE) due to its effectiveness in generating additional samples to improve sport trajectory predictions [6, 10, 11]. For the team strategy, we introduce a new structure, named *team influence graph*, to capture the rapid and dynamic interactions between teammates over a few strokes. Using these team influence graphs, *Team GAT* quantifies and interprets teammate interactions via graph attentions.

For the opponent influence, the novel *Opponent GAT* combines Temporal Convolution Network (TCN) [3] and Graph Attention Network (GAT) to derive the socio-temporal impacts among all players. By modeling teammate and opponent dynamics, TS-CVAE can potentially be generalized to other net sports. Finally, we propose a new data augmentation module, *Singles Location and Shot Type Auto-encoder* (*SOSA*), to extract similar knowledge from abundant singles data to remedy data scarcity.

Figure 1(a) also illustrates a scenario of how coaches and players can exploit TS-CVAE. The orange player is hitting a short flat to the left as the black arrow.

The green player intercepts the shuttle and the blue player moves toward the middle-front to cover, as the fifth steps with hollow nodes. Node P is TS-CVAE's prediction suggesting player B should moved back instead to cover a huge red vulnerable area at the back court by considering teammate movements. Therefore, TS-CVAE helps identify discrepancies, allowing the left team to review and adjust their strategies. For the right team, recognizing the opponents' potential misalignment provides insights for designing counter-strategies.

Experimental results on real doubles badminton data show that TS-CVAE outperforms state-of-the-art sports forecasting models. Our codes can be found on the anonymous github [1] for reproducibility. Our contributions are as follows:

- This paper identifies fine-grained teammate and opponent dynamics in doubles badminton, and proposes a new model, TS-CVAE, to address them.
- Novel Team GAT and Opponent GAT analyze team strategies and opponent influences. SOSA alleviates data scarcity with novel data augmentation.
- Experimental results on real datasets manifest that TS-CVAE effectively predicts player locations, shot types, and hitting players, outperforming baselines in terms of ADE, FDE, and cross-entropy. Visualized results confirm that TS-CVAE's predictions better align with real data.

2 Related Works

Ball Sport AI. AI models have been playing a crucial role in ball sports, including net sports and full-court sports [15, 5], recently. In regards of badminton games, most literature [13, 14, 4, 12] focused on singles badminton. However, generalizing these models to doubles badminton is not trivial since there exists no concept of teammates in singles. Little research studied doubles badminton. MoCVAE [11] presented a CVAE-based model to analyze multiple objectives jointly, but failed to formulate teammate and opponent dynamics. Hence, there is a need for learning team strategies inherent in doubles badminton.

Generative Models. CVAE models generating diverse possible movements are commonly used to improve trajectory forecasting [6, 10] and sport dynamics [11]. dNRI [6] inferred interpretable entity relations with deep encoder and decoder. Grin [10] employed additional graph neural networks to formulate agent movements. Graph-TERN [2] used control point prediction and multi-relational weighted graphs to improve trajectory prediction. Nevertheless, these models did not incorporate the complex dynamics in doubles badminton.

3 TS-CVAE

The proposed Team Strategy-aware CVAE (TS-CVAE) for movement predictions in doubles badminton is illustrated in Figure 2. The left grey shaded area presents SOSA, introducing extra features as a data augmentation module. The right shows a new CVAE structure that incorporates Team GAT and Opponent GAT to finely analyze complex dynamics within and across teams, respectively.

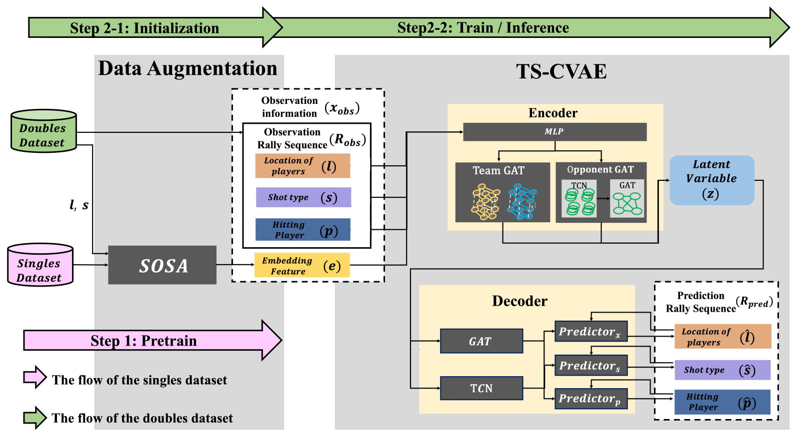


Fig. 2. Illustration of TS-CVAE.

3.1 Preliminaries and Notations

Let a rally be defined as a sequence $R = \langle (l^t, s^t, p^t) \rangle_{t=1}^{T}$. The numbering of strokes and the maximum number are denoted as t and T, respectively. Triplet (l^t, s^t, p^t) records the condition at stroke t. $l^t \in \mathbb{R}^{2\times 4}$ denotes the two-dimensional locations of the four players. $s^t \in \mathbb{R}^{L_s}$ is the one-hot encoding of the shot type occurring at stroke t, where L_s denotes the number of shot types, which is 16 in this paper. $p^t \in \mathbb{R}^4$ is the one-hot encoding of the player hitting the ball. A training instance divides rallies into observation rally sequences $R_{obs} = \langle (l^t, s^t, p^t) \rangle_{t=1}^{T_{obs}}$ and prediction rally sequences $R_{pred} = \langle (l^t, s^t, p^t) \rangle_{t=T_{obs}+1}^{T}$, where T_{obs} denotes the number of strokes for observations. Given the observation rally sequence R_{obs}, TS-CVAE aims to forecast the predicted rally sequence R_{pred}.

3.2 Encoder and Prior

The movement of players in doubles badminton includes various twists and turns, making it challenging to predict. Inspired by [11], TS-CVAE employs a CVAE structure to generate diverse player movements, and further introduces the novel *Team GAT* and *Opponent GAT* to finely learn team and opponent dynamics.

As shown by the green arrows at the top of Figure 2, the input includes two parts: original observation rally sequence R_{obs} in doubles and its processed version (e), enriched by SOSA with singles knowledge. An encoder derives the latent variable z with Team GAT and Opponent GAT. A decoder further predicts all player locations, shot types, and hitting players based on z.

Encoder. The goal of the encoder is to extract the posterior distribution of a latent variable z^E by breaking down team strategies and opponent competitions with Team GAT and Opponent GAT in doubles badminton. The input includes both observation and prediction rallies, which is a common setting in CVAE-based models [6, 10], and is further processed via an MLP layer:

$$Comb^E = MLP(\chi_{obs}),$$

where χ_{obs} denotes observed information by concatenating R_{obs} with additional features e (detail in Section 3.4). $Comb^E$ represents the output of the MLP layer.

Team GAT is designed to capture temporal relations and interactive dynamics between teammates. Specifically, teammates are supposed to cover each other or set up attacks via their positioning. Hence, team strategies significantly influence their movements from stroke to stroke. Team GAT first constructs *team influence graphs* for each team to observe moving patterns.

Figure 1(b) illustrates an example of building a team influence graph with the blue and green player locations on the left court in Figure 1(a). Black dashed boxes represent sliding observation windows of length three, focusing on the previous stroke, the current stroke, and the next stroke. The window length is set to three because it is a sweet spot of capturing fast changing doubles game (examined in experiments). The first three stokes of both players are then merged into a complete graph. Each edge is weighted by a stroke-wide influence, reflecting how each stroke impacts each other, thereby interpreting the team strategy. The influence is derived by graph attentions with GAT and $Comb^E$:

$$TeamGATout^E = GAT^E_{team}(Comb^E),$$

where $TeamGATout^E$ denotes the output of Team GAT. Hence, stronger influences results in greater attention.

On the other hand, it is essential for players to consider their opponents' movements and positions while playing to maximize scoring effectiveness. Accordingly, the Opponent GAT is introduced to learn counteracting behaviors. It first applies a TCN to each player to capture stroke-by-stroke dynamics:

$$OppoTCNout^E = TCN^E_{oppo}(Comb^E),$$

where $OppTCNout^E$ denotes the output of TCN, which polishes $Comb^E$ with temporal influence of each individual. TCN is chosen since its non-recursive structure is less likely to encounter vanishing or exploding gradients [3].

Then, to capture the collective and competitive dynamics among all players, a GAT layer is used after the TCN layer:

$$OppoGATout^E = GAT^E_{oppo}(OppoTCNout^E),$$

where $OppoGATout^E$ denotes the output of the Opponent GAT. This GAT layer attends the sequential embedding $OppoTCNout^E$ based on the interactions between players, modeled as a complete graph. An alternative approach would be using a complete bipartite graph, where teammates are disconnected but opponents remain connected. However, since not only individual players but also team may be affected by opponents, we instead connect the teammates, making it a complete graph.

Note that Team GAT and Opponent GAT conduct different structures for specific reasons. Team GAT focuses on capturing rapidly changing tactical decisions within a team, using complete graphs with fewer strokes to emphasize short-term impacts. Opponent GAT aims to understand the broader influences disrupting interactions between players, requiring analysis across all strokes. Attending complete graphs with all strokes and players could be overwhelming, so Opponent GAT uses a variation to avoid significant computational overhead.

Finally, assume that the latent variable z^E follows a multivariate normal distribution. The posterior is designed using the reparameterization trick [9], ensuring differentiability during optimization:

$$\mu_{z^E}, \sigma_{z^E} = FC(TeamGATout^E, OppoGATout^E),$$

where FC denotes a fully connected layer. μ_{z^E} and σ_{z^E} are the mean and the standard deviation of the posterior distribution $\mathcal{N}(z^E|\mu_{z^E}, \sigma_{z^E})$, respectively. The posterior distribution is derived from the outputs of Team GAT and Opponent GAT, which depict the interactive dynamics.

Prior. The prior aims to approximate z^E, derived by the encoder, by generating another latent variable, denoted by z^P, based solely on the observation rally sequences, excluding the prediction rally sequences. Let $\mathcal{N}(z^P|\mu_{z^P}, \sigma_{z^P})$ denote the prior distribution, where μ_{z^P} and σ_{z^P} are the mean and the standard deviation. To obtain μ_{z^P} and σ_{z^P}, following CVAE-based models, the prior conducts the same structure as the encoder to maintain implementation consistency.

3.3 Decoder and Loss Function

Decoder. The decoder aims at forecasting future strokes by decoding latent variable z^P with TCN and GAT for time domain and interactive dynamics features, respectively. Existing literature of doubles badminton conducted a vanilla decoder [11], failing to retrieve these essential factors. The bottom of Figure 2 illustrates our novel design of the decoder, which is formulated as follows:

$$\begin{aligned}
DecTCNout &= TCN^D(FC^D(z^P)) \\
DecGATout &= GAT^D(FC^D(z^P)) \\
Dout &= Concat(DecTCNout, DecGATout),
\end{aligned}$$

where FC^D denotes a fully connected layer. $DecTCNout$ and $DecGATout$ denote the decoded results of TCN and GAT, respectively. TS-CVAE employs TCN and GAT separately rather than in a cascade to improve efficiency while reducing dependency on processing order. Altogether, $Dout$ concatenates both decoded features for further predictions.

Afterwards, the decoder predicts all four player locations $\hat{l}^{t+1}$, the shot type $\hat{s}^{t+1}$, and the hitting players $\hat{p}^{t+1}$ of next stroke $t+1$ based on $Dout$ as follows:

$$\begin{aligned}
\hat{l}^{t+1} &= FC_l(Concat(Dout^d, \hat{l}^t)) \\
\hat{s}^{t+1} &= FC_s(Concat(Dout^d, \hat{s}^t)) \\
\hat{p}^{t+1} &= FC_p(Concat(Dout^d, \hat{p}^t)),
\end{aligned}$$

where FC are fully connected layers and subscripts l, s, and p denote player locations, shot types, and hitting players respectively. In addition to $Dout$ containing information from observation rally sequences, these predictors also consider the preceding prediction results (i.e., $\hat{l}^t$, $\hat{s}^t$, and $\hat{p}^t$) to align with the sequential nature inherent in doubles badminton.

Loss Function. The loss function includes two parts. First, the Evidence Lower Bound (ELBO) [9] maximizes the negative log-likelihood while minimizing the Kullback-Leibler (KL) divergence between the posterior and prior distribution:

$$ELBO = E[log(p(R_{pred}|z^E, \chi_{obs})] - KL[q(z^E|\chi_{obs}, R_{pred})||p(z^P|\chi_{obs})].$$

Maximizing the negative log-likelihood ensures that the predictions derived from the latent variable (z^E) closely match the original input data. Minimizing the KL divergence aligns the prior distribution with the posterior distribution.

For the second part, cross-entropy $\mathcal{L}_{shot\_type}$ and $\mathcal{L}_{hitting\_player}$ respectively evaluate the shot type and the hitting player predictions:

$$\mathcal{L}_{shot\_type} = -s \cdot log(\hat{s}) \text{ and } \mathcal{L}_{hitting\_player} = -p \cdot log(\hat{p}).$$

Altogether, our goal is to minimize the total loss $\mathcal{L}_{total}$:

$$\mathcal{L}_{total} = min(-ELBO + \mathcal{L}_{shot\_type} + \mathcal{L}_{hitting\_player}).$$

Therefore, minimizing $\mathcal{L}_{total}$ enables TS-CVAE to learn the team strategies and opponent competitions in doubles badminton through a CVAE-based structure.

The inference time of TS-CVAE is about 2 seconds on average, which is sufficient in real-world usage. In doubles badminton, coaches are not permitted to communicate with players during games until technical breaks at every 11 points or the end of a set, making real-time predictions unnecessary. Instead, TS-CVAE prioritizes refining prediction by using the available time effectively.

3.4 Singles Location and Shot Type Auto-encoder

In concern of data scarcity in doubles badminton, borrowing similar knowledge from relatively abundant singles badminton datasets [4, 14, 7] could be a feasible solution. We propose *SOSA* as a pretrained data augmentation module that integrates shot type information with player locations due to their strong correlations [11]. For instance, a player positioned at the back court is more likely to opt for a smash than a short flat.

Figure 3 illustrates the auto-encoder structure of SOSA. The encoder takes player locations and shot types from singles data as input, and further exploits GAT and TCN to derive positional influence and temporal features, respectively:

$$Cout^{SGL} = Concat(GAT^{SGL}(l), TCN^{SGL}(s)).$$

Specifically, GAT attends a complete graph constructed by connecting the location of every stroke of a player within a rally. TCN analyses the temporal relations in shot types within a rally. The output of the encoder, $Cout^{SGL}$, concatenates the results of GAT and TCN and is then transformed into a low-dimensional vector with a fully connected layer.

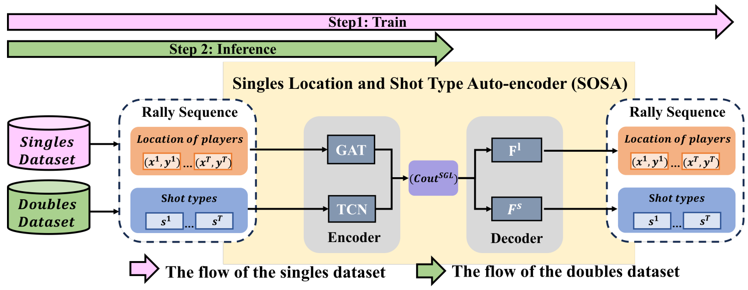


Fig. 3. Illustration of SOSA.

The goal of the decoder is to rebuild the initial input from the encoded low-dimensional features. Two fully connected layers, F^l and F^s, are employed to reconstruct player locations and shot types, respectively:

$$\hat{l} = F^l(Cout^{SGL}) \text{ and } \hat{s} = F^s(Cout^{SGL}),$$

where the predicted player locations and shot types are denoted by $\hat{l}$ and $\hat{s}$, respectively. The auto-encoder loss function, AE_{loss}, is designed as follows:

$$AE_{loss} = MSE(l, \hat{l}) + \lambda \cdot CE(s, \hat{s}),$$

where the first term conducts mean square error (MSE) to evaluate the predicted player locations $\hat{l}$, and the second term uses cross-entropy (CE) to evaluate the predicted shot types $\hat{s}$. λ balances the two losses.

After training, only the encoder is retained for further use. As shown in the middle of Figure 2, when doubles data is used as input, $e = Cout^{SGL}$ offers extra knowledge to TS-CVAE. As e is appended (rather than replacing) to R_{obs}, the original doubles features are preserved and remain visible during training.

4 Experiment

4.1 Experiment Setting

To evaluate TS-CVAE, we compare with state-of-the-art CVAE-based models for doubles badminton and full-court sports (e.g., basketball) models:

1. *MoCVAE* [11] utilized shot type and the hitter to predict the movement of players in doubles badminton without separating teams and opponents.
2. *dNRI* [6] uncovered the dynamic relations between players with RNN.
3. *Grin* [10] utilized graph neural networks to incorporate interactive dynamics and agent intention.

To understand each baseline's behavior, we consider three observation and prediction stroke combinations, which are (11,1), (10,2), and (8,4), respectively. The default combination is (8,4) if not specified. Note that original dNRI and Grin did not predict shot types and hitters. To maintain a fair setting, we concatenate this two information as the input for them. Furthermore, they all use an MLP and two individual softmax layers to predict the shot type and the hitter.

To pretrain SOSA, 170 singles matches including 448 sets and 11,284 rallies are used [7]. For the main task, 25 doubles matches including 58 sets and 2230 rallies are used. Shot types, positions of each player, and hitting players are labeled by [8]. The number of shot type classes is 16. The player locations adopt a min-max normalization. The dataset is separated based on rallies, where 80% and 20% are for training and testing, respectively.

In TS-CVAE, the dimensions of the hidden states and the latent variables are 8 and 16, respectively. The learning rate is 10^{-3} and the batch size is 32. λ is set to 1. For SOSA, the learning rate is set to 5×10^{-4} and the batch size is 1024. All baselines are trained on a machine of AMD Ryzen 7 5700X 8-Core CPU and an Nvidia RTX 3060 Ti GPU.

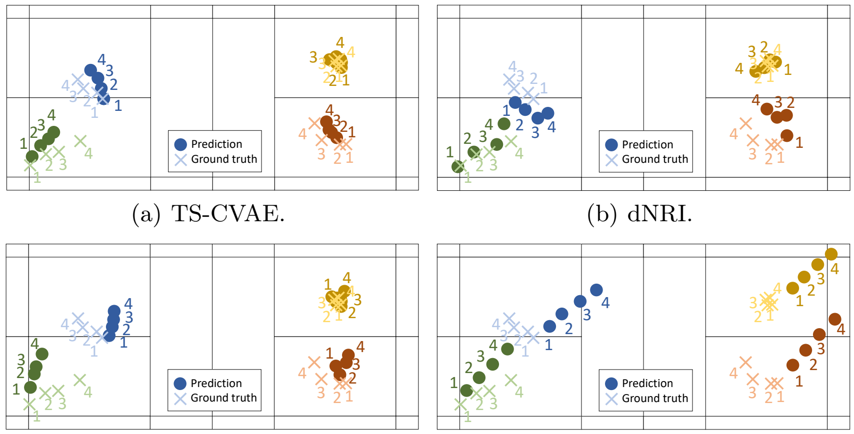
Following baselines [6, 10, 11], same evaluation metrics of player locations are used, specifically final displacement error (FDE) and average displacement error (ADE). ADE provides an overall assessment of movement performance by calculating the average squared error between the actual and predicted locations in x-y coordinates across all strokes. Conversely, FDE evaluates final movement performance by measuring the average squared error at the endpoint of each movement. The original court dimensions are 13.4 meters in length and 6.1 meters in width. We equally divided and normalized the court into x-y coordinates, so each unit of ADE and FDE after preprocessing being approximately 1.34 meters. Cross-entropy evaluates shot types and player predictions.

4.2 Performance Evaluation

Table 1. Prediction performance of each baseline.

Model	11 strokes/ 1 stroke				10 strokes/ 2 strokes				8 strokes/ 4 strokes			
	ADE	FDE	Shot Type	Hitting Player	ADE	FDE	Shot Type	Hitting Player	ADE	FDE	Shot Type	Hitting Player
dNRI	0.110	0.110	3.175	2.546	0.128	0.158	3.593	2.887	0.160	0.208	3.653	2.978
Grin	**0.108**	**0.108**	3.899	3.955	0.124	0.152	4.339	3.807	0.153	0.176	3.907	3.567
MoCVAE	0.136	0.136	2.586	1.433	0.147	0.181	2.746	1.401	0.289	0.42	3.204	1.448
TS-CVAE	0.111	0.111	**2.437**	**1.348**	**0.117**	**0.140**	**2.676**	**1.389**	**0.142**	**0.172**	**2.733**	**1.383**

Table 1 compares each baseline in predicting player position, shot type, and hitting player. The results demonstrate that TS-CVAE outperforms all other models across most scenarios. Specifically, for an observation window of 8 strokes and a prediction of 4 strokes, TS-CVAE surpasses all baselines by 7.19% to 50.86% in ADE and 2.27% to 59.05% in FDE. Its ADE and FDE are about 0.19 and 0.23 meters in practice and both are much smaller than an adult step length (0.7 meter). It also exceeds by 14.7% and 4.49% in terms of CE for shot types and hitting players, respectively. These results manifest that TS-CVAE effectively identifies player locations by introducing Team GAT and Opponent GAT to simultaneously analyze teammate and opponent dynamics. Note that the improvement is significant since TS-CVAE optimizes multiple objectives jointly, unlike Grin and dNRI. Moreover, we achieve similar improvement level to state-of-the-art singles [14] and doubles [11] badminton movement prediction papers.



(a) TS-CVAE. (b) dNRI.

(c) Grin. (d) MoCVAE.

Fig. 4. Player movement predictions of each baseline. Colors distinguish each player.

While MoCVAE focuses on doubles badminton as well, it does not separate teammates and opponents, resulting in less accurate predictions. On the other hand, in the setting of 11 and 1 strokes, TS-CVAE only achieves comparable ADE and FDE to Grin because predicting one single stroke is relatively easy for full court models. However, TS-CVAE is more comprehensive in all crucial tasks of doubles badminton, as evident by at least 37.5% improvement on predicting the hitting players and shot types

Figure 4 visualizes four consecutive predicted player locations. Different colors indicate different players, and the numbers represent the order of strokes. The visualization result highlights that TS-CVAE in Figure 4(a) predicts player locations more accurately. Specifically, the team on the right (brown and yellow) adopts a defensive strategy, so they remain parallel on their respective halves and hardly move. In contrast, Figure 4(b) shows that dNRI predicts the two players progressively moving closer, which contradicts typical defensive team strategies by reducing their coverage area. Grin predicts that the green player aggressively moves toward the middle, as shown in Figure 4(c), which exposes larger vulnerable areas on both upper and lower courts. These observations demonstrate that TS-CVAE, by incorporating Team GAT, better reflects team strategies in practice. Figure 4(d) shows that all four players are predicted to move upper-right, indicating that MoCVAE using a single MLP layer for decoding fails to correctly decode temporal and interactive features. This result underscores the necessity of employing both TCN and GAT in the decoder of TS-CVAE.

4.3 Sensitivity Tests

Different prediction strokes. Figure 5 shows the ADE of different prediction strokes of all baselines while fixing the observation setting at 10 strokes. The solid line presents the average results, and the shaded area shows the range between the maximum and the minimum. For TS-CVAE, the slope is less steep and

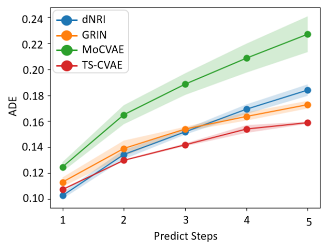
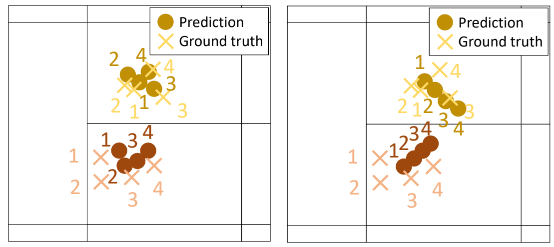


Fig. 5. Performance of different prediction strokes.



(a) Using 3 strokes. (b) Using 7 strokes.

Fig. 6. Player location predictions using various observation window lengths in team influence graphs.

Table 2. Performance on observation strokes in Team GAT.

Strokes	2	3	5	7
ADE	0.150	**0.143**	0.148	0.149
FDE	0.182	**0.174**	0.180	0.182

has lower ADEs in most of the prediction strokes. Besides, the shaded area is relatively smaller. All these results manifest that TS-CVAE is robust in extending prediction strokes by accurately interpreting and forecasting game situations via the comprehensive knowledge extracted by Team GAT and Opponent GAT.
Team influence graphs in Team GAT. Table 2 examines the sensitivity of the observation strokes for constructing team influence graphs in Team GAT with ADE and FDE. The results manifest that 3-stroke setting provides the best overall performance. The ADE increased by 4.9% for 2-stroke setting, suggesting that it is insufficient to capture team strategies. Although the 5 and 7-stroke settings are supposed to bring in more context, their results show at least a 3.5% degradation in ADE. It is because the content of doubles badminton change rapidly, requiring team strategies to react quickly. Including too many strokes actually brings noise instead of useful information.

Figure 6 further visualizes the outcomes of 3 and 7-stroke settings for the same rally. Predictions of both players in Figure 6(a) show more twists and turns, aligning with the ground truth. Figure 6(b) shows that incorporating too many strokes may eliminate twists and turns, causing overly smooth predictions.

4.4 Ablation Tests

Table 3 presents the prediction performance of each variant. Specifically, the removal of TCN from the decoder (*w/o TCN in Decoder*) results in the worst prediction outcomes for player movement, with decreases of 5.33% and 5.49% in terms of ADE and FDE, respectively. The drop in performance indicates that temporal features are crucial in retrieving dynamics on courts. Removing Team GAT from the encoder (*w/o Team GAT*) leads to declines in player movement predictions by 1.41% in ADE and 1.16% in FDE, highlighting the importance of the Team GAT in capturing team strategies to better understand movement

Table 3. Ablation study on leave-one-out variants.

Task	Movement Forecasting		Shot Type	Hitting Player
	ADE	FDE	Cross Entropy	
w/o TCN in Decoder	0.150	0.182	2.742	1.412
w/o Team GAT	0.144	0.174	2.731	1.387
w/o Opponent GAT	0.143	0.176	2.751	1.386
w/o SOSA	0.144	0.176	2.760	1.391
TS-CVAE	0.142	0.172	2.733	1.383

between teammates. *w/o Opponent GAT* increases CE for both shot types and hitters due to losing insight into opponents' positioning, which is supposed to identify vulnerable areas to attack. *w/o SOSA* results in the worst performance for shot type prediction, showing the effectiveness of SOSA in transferring shot type knowledge from singles to doubles. These results manifest each component contributes to improve player movement predictions for doubles badminton.

5 Conclusion

This paper recognizes the crucial team strategies and opponent competitions in doubles badminton. Accordingly, we propose the novel TS-CVAE to predict player locations, shot types, and hitting players from stroke to stroke, respectively. TS-CVAE includes two innovative modules: Team GAT copes with rapidly varying team strategies based on novel team influence graphs and Opponent GAT holistically analyzes rival dynamics. Moreover, SOSA leverages data augmentation from singles data to improve predictions. Evaluations on real datasets manifest that TS-CVAE outperforms state-of-the-art sport forecasting models.

References

1. The full version and codes on anynomous github, `https://tinyurl.com/drj7pkmu`
2. Bae, I., Jeon, H.G.: A set of control points conditioned pedestrian trajectory prediction. In: AAAI (2023)
3. Bai, S., et al.: An empirical evaluation of generic convolutional and recurrent networks for sequence modeling (2018)
4. Chang, K.S., et al.: Where will players move next? dynamic graphs and hierarchical fusion for movement forecasting in badminton. In: AAAI (2023)
5. Chen, X., et al.: Professional basketball player behavior synthesis via planning with diffusion (2023), `https://arxiv.org/abs/2306.04090`
6. Graber, C., Schwing, A.G.: Dynamic neural relational inference. In: CVPR (2020)
7. Huang, Y.H., et al.: S2-labeling: Shot-by-shot microscopic badminton singles tactical dataset. In: APNOMS (2022)
8. Huang, Y.H., et al.: Shot-by-shot technical data collection for badminton doubles games. In: APNOMS (2023)
9. Kingma, D.P., Welling, M.: Auto-encoding variational bayes. In: ICLR (2014)
10. Li, L., et al.: Grin: Generative relation and intention network for multi-agent trajectory prediction. In: NeurIPS (2021)

11. Sung, P.C., et al.: Mocvae: Movement prediction by a conditional variational autoencoder for doubles badminton. In: IEEE BigComp (2024)
12. Wang, K.D., et al.: Offline imitation of badminton player behavior via experiential contexts and brownian motion (2024), https://arxiv.org/abs/2403.12406
13. Wang, W.Y., et al.: Exploring the long short-term dependencies to infer shot influence in badminton matches. In: IEEE ICDM (2021)
14. Wang, W.Y., et al.: Shuttlenet: Position-aware fusion of rally progress and player styles for stroke forecasting in badminton. In: AAAI (2022)
15. Yanai, C., et al.: Q-ball: Modeling basketball games using deep reinforcement learning. In: AAAI (2022)