Football Tactical Scenario Generation and Multi-Strategy Confrontation Simulation Guided by Diffusion Models

Juhai Wang
College of Physical Education, Weinan Normal University, Weinan 714099, Shaanxi, China
Corresponding author: Juhai Wang (e-mail: Wjh13992333818@sina.com)

ABSTRACT This paper proposes a collaborative framework integrating diffusion models and multi-agent reinforcement learning to address unrealistic trajectory generation and tactical incoherence in high-dimensional football tactical simulation. A conditional diffusion model is trained on 200 professional matches from Europe's top five leagues during the 2021–2023 seasons to learn tactical distributions and generate spatiotemporal trajectories through denoising. Tactical labels and a graph attention mechanism are incorporated to improve structural rationality and semantic controllability. The generated trajectories are then used directly as the initial state of the simulation environment. Hierarchical action design and behavioral cloning pre-training initialize multi-agent policies so that decision-making begins from tactically plausible states. Finally, the multi-agent proximal policy optimization algorithm is applied in a red-blue confrontation environment, with ball possession, spatial utilization, and tactical consistency rewards used to support dynamic deduction. Experiments confirm high trajectory fidelity, with average absolute trajectory error of 0.82–1.03 m, tactical consistency up to 83.4% role-behavior fit, and dynamic evolution stability of $3.1\ \times 10^{-2}$ strategy volatility. The framework provides a generative simulation method for multi-agent spatiotemporal systems and may inform scenario synthesis in wireless communication environments.

INDEX TERMS multi-agent reinforcement learning, diffusion model, trajectory modeling, adversarial simulation, wireless communication environments

I. INTRODUCTION

THE essence of football tactics is the coordination and interplay of multiple intelligent agents under spatiotemporal constraints, where its dynamic nature stems from the real-time coupling of individual player decisions and collective intentions. Analogously, the dynamic nature of a smart factory is defined by the real-time coupling of individual machine decisions and the collective production goals of the entire manufacturing system under material flow and energy consumption constraints [1,2]. The high-dimensional trajectory data generated by modern professional matches contains implicit laws of tactical structure. Transforming this data into a generative and deducible model expression becomes a key transition point connecting data analysis and actual combat decision-making [3]. Building a generative-adversarial simulation system with tactical semantic understanding capabilities and restoring tactical distribution through deep generative models [4–6] can promote the advancement of football intelligent simulation from passive restoration to active rehearsal.

This research aims to build a football tactical simulation system that integrates generative modeling and dynamic game theory, overcoming the dual limitations of existing methods in scene realism and strategic interactivity. This integrated approach offers a blueprint for the industry to create sophisticated smart-factory simulations that overcome current limitations in modeling the realism of material-flow scenarios and the interactive decision-making of autonomous production agents. The core goal is to learn tactical distributions from professional match data and generate initial configurations with clear tactical intent, thereby driving autonomous multi-agent strategy evolution in competitive environments. The innovations are embodied in four aspects: First, a conditional trajectory generation model based on a diffusion mechanism is designed. Using noise scheduling and denoising inversion, tactical patterns are extracted from high-dimensional spatiotemporal data, generating player trajectories that conform to physical constraints and teamwork laws. Second, by introducing tactical labels as generation guidance signals and combining classifier-free guidance with a graph attention network, the generation process is made semantically controllable and structurally rational, avoiding meaningless or unbalanced formations. Third, the generated results are converted into the initial state of the simulation environment. The strategy space is divided according to player roles, and hierarchical action design and behavior cloning pre-training are used to improve the quality of strategy initialization and shorten the exploration cycle. Fourth, a red-blue game framework is constructed. Based on the multi-agent proximal policy optimization (MAPPO) algorithm, offensive and defensive strategies are optimized. A composite reward function is designed that integrates ball possession efficiency, spatial utilization, and tactical consistency to promote the evolution of strategies toward a high level of equilibrium. This framework not only supports the automated construction of diverse tactical scenarios but can also be used to deduce the evolution of specific formations under different opponent strategies, providing an interpretable adversarial analysis tool. Overall, this research provides a new methodological approach for tactical simulation, enabling a leap in capabilities from replaying history to previewing the future.

II. RELATED WORKS

In recent years, researchers have attempted to model player movement patterns through data-driven approaches. Some work uses Bayesian optimization [7–9] to learn the distribution of football trajectories or combines it with variational autoencoders [10] to capture physical motion. Although these methods can capture basic movement trends, they are prone to structural collapse when generating long sequences, making it difficult to maintain team formation stability. Another approach introduces generative adversarial networks (GANs) [11,12] to improve trajectory realism during adversarial training, but these suffer from training instability and pattern collapse, and the generated results lack tactical consistency. To enhance semantic control, some researchers have introduced conditional tags to guide the direction of human motion generation through category embedding [13]. However, this is mostly limited to the single-agent framework and ignores the strategic interaction between the offensive and defensive sides. Existing simulation systems generally separate generation and deduction. The generation phase lacks executable verification, while the deduction phase lacks support for diverse initial conditions, limiting the system's generalization capabilities and the depth of tactical exploration.

Addressing the gap between generation quality and strategic rationality, diffusion models have recently demonstrated excellent performance in image and sequence generation [14,15]. They employ a stepwise denoising inverse process to stably model complex distributions, avoiding the non-convergence problem of GANs. Kong L. proposed a spatiotemporal relationship modeling method that combines an adaptive graph convolutional network and an attentive temporal convolutional network to effectively capture player interactions and team dynamics, achieving a comprehensive improvement in tactical recognition in sports videos [16]. In terms of multi-agent decision-making, MAPPO has become a mainstream algorithm for processing partially observable environments due to its advantages of decentralized execution and centralized training. It has been applied in multi-machine coordination [17,18], but few studies have applied it to football tactical scenarios. Some studies have used behavioral cloning (BC) as a pre-training method to significantly accelerate strategy convergence; Coimbra F. V. proposed a data-driven imitation learning method that uses deep neural networks to model team behavior in football games. The trained model achieved an accuracy of 84.5% in action selection and performed stably when playing against unknown opponents, with negligible performance degradation [19].

III. CONSTRUCTION OF A TACTICAL CONFIGURATION GENERATION AND ADVERSARIAL REASONING FRAMEWORK

Figure 1 illustrates a diffusion model-driven tactical configuration generation and adversarial reasoning framework. Conditioned on tactical labels, the diffusion model, combined with a graph attention denoising network, generates semantically structured player trajectories, which are mapped to the initial state of the simulation. The red and blue agents divide the strategy space by role, and behavioral cloning pre-training is used to improve initialization quality. Each team optimizes its strategy using the MAPPO framework, sharing a value network and receiving compound rewards. The system evolves tactical behavior through dynamic interaction, combining Voronoi field analysis with behavioral pattern monitoring to achieve a closed-loop deduction from generation to game play.

A. PLAYER TRAJECTORY DISTRIBUTION MODELING BASED ON THE DIFFUSION MODEL

1) Conditional Diffusion Model Construction and Training Process

The modeling foundation is based on high-dimensional spatiotemporal player trajectory data collected from professional leagues. Each sample contains the two-dimensional coordinates and instantaneous velocity vectors of 11 players on the field in continuous time frames. These dynamic states are organized into time series tensors, which serve as the original distribution sampling source for the diffusion process. The model sets the forward diffusion phase as a T-step Markov chain. At each step, a small Gaussian noise is added to the original trajectory data, gradually destroying its structural information until it approaches a pure noise distribution. This process is described by the parameterized equation:

$$q(x_t \mid x_{t-1}) = \mathcal{N}\left(x_t; \sqrt{1-\beta_t}\, x_{t-1}, \beta_t I\right) \tag{1}$$

x_t represents the noisy trajectory state at step t, and β_t is a preset noise scheduling coefficient that increases with each time step to control the noise injection rate. This scheduling strategy uses a cosine variant to ensure that macroscopic structure is preserved in the early stages and that local motion trends are fine-tuned in the later stages. The entire forward process gradually transforms the

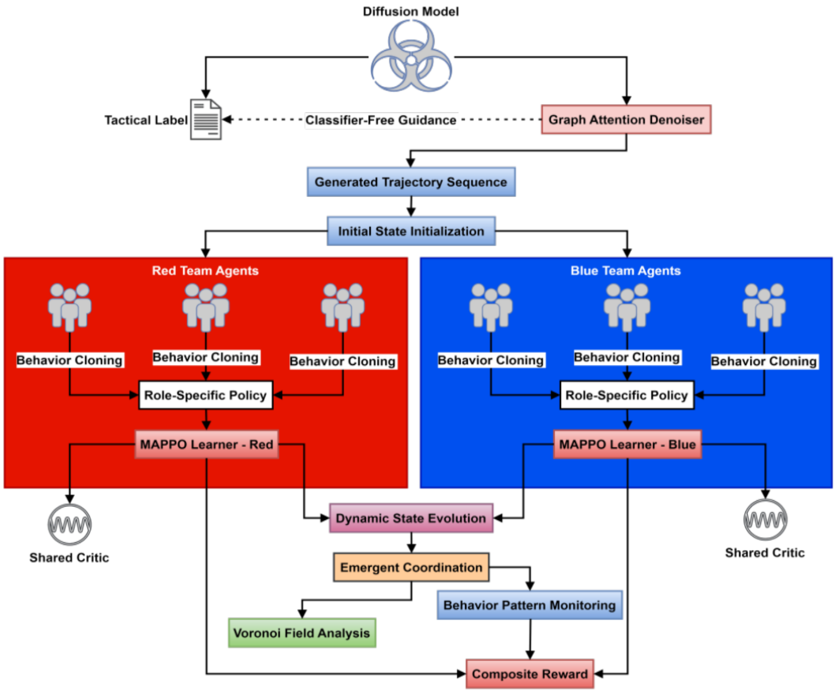

FIGURE 1. Diffusion Model-Driven Tactical Configuration Generation and Countermeasures Simulation Framework

true tactical trajectory distribution into a standard normal distribution, forming a reversible learning path.

The inverse denoising network is responsible for reconstructing tactical trajectories from noise. Its core is a time-aware graph neural network architecture. The network inputs are the embedded representations of the noisy trajectory xt and time step t, and outputs the predicted noise residual. The training objective is to minimize the mean squared error between the predicted noise and the actual added noise:

$$L_{\text{diff}} = \mathbb{E}_{x_0,t,\epsilon}\left[\|\epsilon - \hat{\epsilon}_\theta(x_t,t)\|^2\right] \tag{2}$$

x_0 represents the original noise-free trajectory, ϵ represents the sampled noise, and $\hat{\epsilon}_\theta$ represents the estimated function for the model parameterization. The network weights are iteratively updated on a large-scale match dataset, enabling the model to recover trajectories that conform to kinematic constraints and team coordination rules from any noise level. Each denoising step considers the relative positions of players, utilizing a learnable spatial adjacency matrix to capture potential interaction patterns and prevent individual movements from deviating from the overall tactical framework.

2) Tactical Condition-Guided Trajectory Generation Mechanism

The generation phase incorporates adjustable tactical prior information to achieve semantic control over the output trajectory. Condition variables include possession, initial formation encoding (e.g., 4-3-3 or 3-5-2), and attack/defense status labels. These discrete categories are transformed into continuous vectors by an embedding layer, concatenated with time-step information, and then fed into the multi-layer transformation module of the denoising network. This design allows the generation process to be carried out within a specific tactical context, ensuring that the output trajectory is not only physically feasible but also clearly tactically targeted. For example, in a high-pressure scenario, forwards and midfielders start closer to the opponent's half, with their movements exhibiting a forward-pressing, convergent pattern. In a quick counterattack scenario, a few attacking players exhibit high-speed vertical penetration, while the rest lag behind. The denoising process starts with a pure noise tensor and performs T iterations in reverse chronological order. Each iteration calls the trained network to estimate the current noise component and remove it. The update rule follows the reparameterized path:

$$x_{t-1} = \frac{1}{\sqrt{1-\beta_t}}\left(x_t - \frac{\beta_t}{\sqrt{1-\bar{\alpha}_t}}\hat{\epsilon}_\theta(x_t,t,c)\right) + \sigma_t z \tag{3}$$

c is the conditional input, $\bar{\alpha}_t$ is the cumulative noise coefficient, z is the standard normal sampling term, and σ_t controls the randomness intensity. The final output is a complete sequence of player trajectories that conform to the specified tactical settings. This generation mechanism reliably reproduces common spatial distribution and movement coordination patterns in real matches while maintaining diversity, providing a high-fidelity initial configuration for subsequent competitive simulations.

B. CONDITIONAL GENERATION MECHANISM GUIDED BY TACTICAL SEMANTICS

1) Tactical Semantic Control Under Classifier-Free Guidance

Tactical labels serve as the core semantic anchor in the generation process and directly determine the strategic attributes of the output trajectories. The model uses a classifier-free bootstrapping mechanism to achieve refined control of tactical inclinations. This approach relies on a dual-path conditional modeling architecture. During the training phase, two denoising branches are constructed in parallel: one accepts the complete tactical label as conditional input, while the other ignores the label information under a random mask, forcing the network to learn unconditional generation capabilities. The two branches share core parameters and are separated only at the conditional injection layer, ensuring the co-evolution of semantic perception and general denoising capabilities. This design avoids reliance on explicit gradient guidance functions and improves generation stability.

During the derivation process, the final noise prediction result is generated using a weighted difference mechanism:

$$\hat{\epsilon}_{\text{guided}} = \hat{\epsilon}_\theta(x_t, t, c) + \omega\left(\hat{\epsilon}_\theta(x_t, t, c) - \hat{\epsilon}_\theta(x_t, t, \emptyset)\right) \quad (4)$$

$\hat{\epsilon}_\theta(x_t, t, c)$ is the conditional noise estimate, $\hat{\epsilon}_\theta(x_t, t, \emptyset)$ is the unconditional estimate, and ω is the guidance strength hyperparameter, controlling the balance between semantic fidelity and diversity. When ω is set to a high value, the generated trajectories strictly adhere to the specified tactical pattern. For example, the “wing penetration” label can significantly enhance the lateral stretching and deep penetration of wingers and fullbacks. Lower ω allows for more random variation, preserving individual differences in tactical execution. The label space covers key tactical categories on both offense and defense. Each label is one-hot encoded and converted into a semantic vector using a learnable embedding matrix. These vectors are aligned with the denoising steps in time and fused with the player states in space, driving the network to maintain tactical consistency throughout each denoising step. Table 1 lists the core configurations of the tactical semantics-guided generation mechanism.

TABLE 1. Core Parameter Configuration of the Tactical Semantics-Guided Generation Model

Parameter	Value	Description
Guidance Scale	2.2	Controls semantic fidelity of generated trajectories toward specified tactics
Diffusion Steps	1,000	Number of denoising iterations; determines temporal resolution
Condition Embedding Dim	64	Dimension of encoded tactical label vectors
Graph Attention Networks (GAT) Heads	4	Number of attention heads in graph layers for multi-perspective interaction modeling
Hidden State Dim	128	Feature dimension of player trajectory representation
Neighborhood Threshold	25.0 m	Maximum distance for establishing player interaction edges
Condition Dropout Rate	0.2	Probability of masking tactical label during training for robustness
GAT Layers	3	Number of stacked graph attention layers for hierarchical spatial modeling

2) Graph Attention-Driven Structural Constraint Modeling

The denoising network integrates a graph attention mechanism to capture the dynamic interactions between players, ensuring spatial rationality and team coordination in the generated trajectories. The graph structure is defined as follows: nodes represent players (including their kinematic states), and edges represent spatial interactions between players, with weights computed based on Euclidean distance using a negative exponential function. The graph attention layer then calculates attention coefficients between connected nodes, enabling the model to prioritize interactions with tactically relevant teammates. This ensures that generated trajectories maintain structural rationality by dynamically adjusting player movements based on collective context:

$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}\left(a^\top [Wh_i \| Wh_j]\right)\right)}{\sum_{k\in N_i} \exp\left(\text{LeakyReLU}\left(a^\top [Wh_i \| Wh_k]\right)\right)} \quad (5)$$

h_i represents the player’s hidden state, W is the learnable projection matrix, $a^\top$ is the attention vector, $\|$ represents vector concatenation, and N_i represents its neighborhood set. A multi-head mechanism performs attention computations in parallel, concatenating the results and applying a linear transformation to output updated node representations, enhancing the ability to model complex spatial configurations. Figure 2 shows the average multi-head graph attention weight. Both the horizontal and vertical axes represent player numbers (1–11). Player numbers are defined in a standard 4-3-3 formation: 1 – goalkeeper, 2 – right back, 3 – left back, 4 – defensive midfielder, 5 – central midfielder, 6 – attacking midfielder, 7 – right wing, 8 – central midfielder, 9 – center forward, 10 – shadow forward, 11 – left wing. The color of the matrix elements reflects the attention weight of player i on player j. The data show that significant high-value regions are concentrated in the blocks numbered 4 (defensive midfielder), 5 (central midfielder), 6 (attacking midfielder), 7 (right winger), 9 (center forward), and 10 (shadow forward). These positions form the core network for central attack, with attention weights all exceeding 0.26.

C. MULTI-AGENT POLICY SPACE INITIALIZATION AND ROLE ASSIGNMENT

1) Policy Space Decoupling and Hierarchical Action Modeling

The action space adopts a hierarchical design structure, with the top layer for behavior mode selection and the bottom layer for continuous action execution. The top layer defines five high-level behavior modes: dribbling, short pass reception, long pass transfer, assist defense, and on-site observation. Each mode corresponds to a restricted set of low-level actions. The agent outputs a probability distribution of behavior categories in each decision cycle and activates the corresponding action subspace based on the maximum probability. This mechanism effectively compresses the exploration dimension and avoids blindly trying invalid actions that do not conform to role responsibilities, such as defenders frequently shooting or forwards staying in their own half for a long time. Behavior selection is achieved by normalization of the softmax function:

$$p_m = \frac{\exp(s_m/\tau)}{\sum_{k=1}^{5}\exp(s_k/\tau)} \tag{6}$$

p_m represents the activation probability of the mth behavior pattern, s_m represents its raw score, and τ is a temperature parameter that controls the randomness of the selection. The score is derived from the current observed state through a forward calculation of a character-specific network, encoding features such as position, relative distance, ball possession, and teammate distribution. Once a high-level decision is made, the underlying network outputs a specific action vector within the corresponding action domain, including movement direction, pass force, and target coordinates, forming a closed intention-execution loop.

2) Strategy Pre-Training and Initialization Optimization

The BC pre-training uses state-action pairs derived from the same professional data that fuels the diffusion model, ensuring alignment between generated scenarios and expert behavior. Specifically, the diffusion model's output trajectories are mapped to initial states, and BC pre-training initializes policies to mimic expert actions under these states, reducing semantic gap and preventing early exploration divergence. The hierarchical action design constrains agents to role-appropriate behaviors, which synergizes with BC pre-training to maintain tactical consistency. This prevents MAPPO agents from deviating into erratic actions, as the action space is structurally bounded by tactical roles.

The cloning training objective is to minimize the composite loss between the policy network output and the expert's actions:

$$L_{\mathrm{bc}} = \lambda_{\mathrm{disc}} L_{\mathrm{ce}} + \lambda_{\mathrm{cont}} L_{\mathrm{mse}} \tag{7}$$

L_{ce} is the cross-entropy loss for behavior categories, L_{mse} is the mean squared error for continuous actions, and λ_{disc} and λ_{cont} are balance coefficients. During training, network parameters are iteratively updated on a large batch of expert trajectories, learning the mapping from local observations to appropriate actions. After pre-training, each character-specific network has acquired basic tactical responsiveness and can react to typical scenarios with near-realistic results, such as a forward automatically advancing into an open area and a midfielder prioritizing lateral passing routes after receiving the ball. Figure 3 shows a comparison of strategy selection probability and training convergence performance based on hierarchical role behavior modeling. Figure 3(a) shows the probability distribution of selection based on high-level behavior patterns. The vertical axis represents five high-level behavior patterns: dribble, short pass, long pass, defensive cover, and hold position, while the horizontal axis represents the probability of behavior pattern selection (between 0 and 1). The corresponding three roles include forward, midfielder, and defender. In the "dribble" behavior, the forward has the highest probability (0.45), while the defender has a probability of approximately 0.05. In the "defensive cover" behavior, the defender probability is as high as 0.50, while the forward and midfielder probabilities are significantly lower. This distribution clearly demonstrates the effective differentiation of different role functions achieved by hierarchical action design. Figure 3(b) shows a comparison of the policy convergence speed and stability between pre-training with behavioral cloning and without using behavioral cloning. The behavioral cloning method neared convergence around round 10, with the reward fluctuating around 0.8. In contrast, the method without behavioral cloning converged in more rounds, with the reward level below 0.6.

Figure 4 shows a visualization of a simulated wing-break tactic in soccer. The green background represents the turf, and white lines mark field features such as the perimeter, penalty area, and center circle. Player positions are represented by colored dots: red for goalkeeper, blue for defender, green for midfielder, and orange for forward. The players' movements are clearly marked with black dotted lines, highlighting the midfielder's rapid forward run. The ball is marked with a black square, indicating the forward penetration area. The overall design uses a high-contrast color scheme, maintaining a professional look while ensuring visual clarity, perfectly conveying the formation's tactical intent and the coordinated movement of the players.

D. DYNAMIC DEDUCTION MECHANISM DRIVEN BY OFFENSE-DEFENSE CONFRONTATION

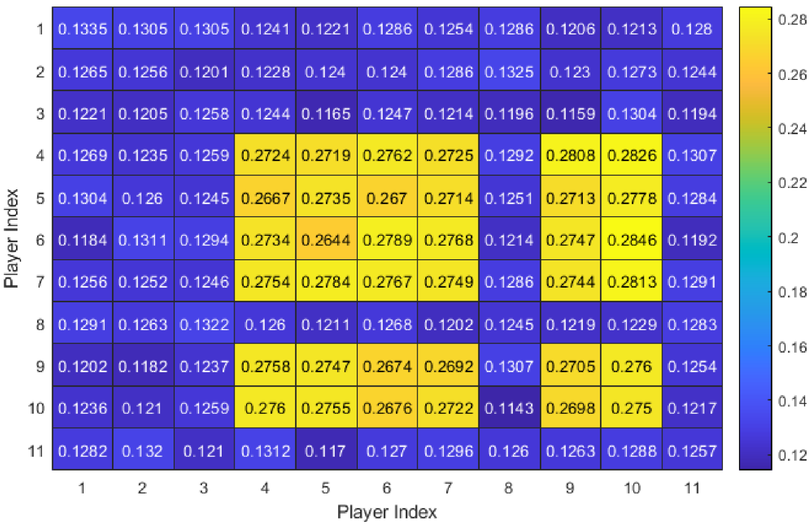

FIGURE 2. Average Multi-Head Graph Attention Weight (Tactical Label: Central Penetration)

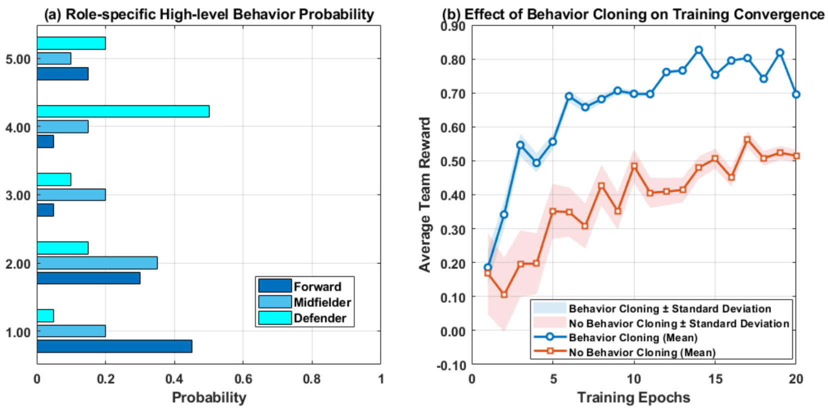

FIGURE 3. Strategy Selection Probability and Training Convergence Performance

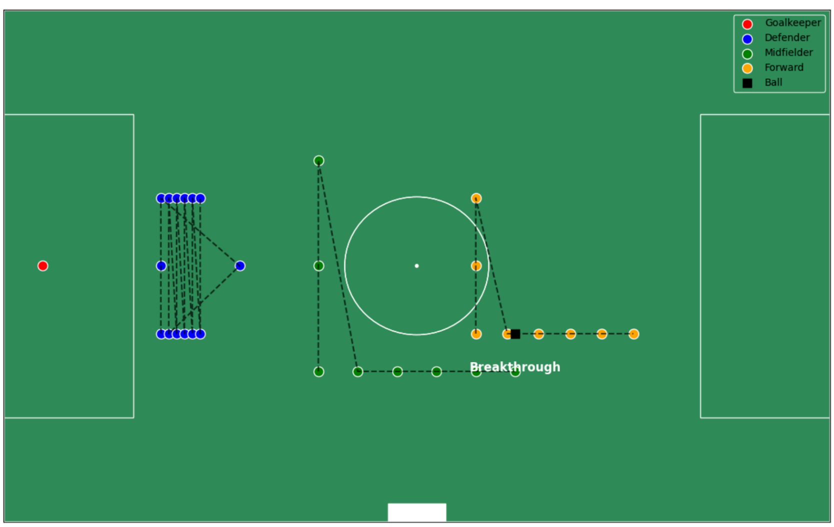

FIGURE 4. Visualization of a Wing-Break Tactic

1) Multi-Agent Confrontation Environment Design and Strategy Co-Optimization

The simulation environment consists of two teams of 22 agents each, representing offense and defense, executing action decisions in a continuous physical space. Each agent's observation range is limited to a local field of view, encompassing its own state, the ball's position, the relative coordinates of its five nearest teammates and opponents, and their velocity vectors. It also receives global event signals, such as changes in possession, offside triggers, and out-of-bounds situations. Action execution utilizes an asynchronous update mechanism. All agents synchronously receive observations and output action commands each frame. These are integrated into the environment dynamics model and updated to the next state, forming a closed-loop interaction. This setup simulates the characteristics of partial observability and concurrent decision-making in real-world competitions, forcing agents to rely on local perception and role presets for collaboration. Strategy optimization for both agents utilizes the MAPPO framework. Each agent maintains an independent actor-critic network structure, sharing the same team's critic parameters to enhance training stability. The actor network outputs action probability distributions, while the critic network evaluates the value function of the current joint strategy. Policy updates utilize a clipping mechanism to prevent parameter mutations. The loss function is defined as:

$$L_{\text{clip}} = \mathbb{E}_t \left[\min \left(r_t(\theta)\hat{A}_t, \text{clip}(r_t(\theta), 1-\epsilon, 1+\epsilon)\hat{A}_t \right) \right] \tag{8}$$

$r_t(\theta)$ is the probability ratio of the new and old strategies, $\hat{A}_t$ is the generalized advantage estimate, and ϵ is the clipping threshold. Advantage calculation uses the generalized advantage estimation (GAE) method, combining multi-step temporal difference errors with a discount factor to balance bias and variance. During training, the experience replay buffer stores complete adversarial trajectories, and multiple rounds of policy updates are performed in batches to ensure stable gradient direction.

2) Compound Reward-Driven Tactical Consistency Evolution

The reward function design integrates physical goal achievement with tactical logic compliance to guide the strategy toward high-level competitive behavior. Each step's reward is a weighted composite of three core components: ball possession gain, spatial control gain, and a penalty for tactical consistency deviation. The final reward function is formulated as:

$$R_t = \omega_p \cdot R_{\text{possession}} + \omega_s \cdot R_{\text{space}} + \omega_t \cdot R_{\text{tactic}} \tag{9}$$

The weights $(\omega_p, \omega_s, \omega_t) = (0.5, 0.3, 0.2)$ were determined through an extensive grid search over a held-out validation set of tactical scenarios.

A sensitivity analysis revealed that an over-emphasis on ball possession ($\omega_p > 0.7$) led to myopic and frantic play with high possession turnover (as reflected in the high frequency in Table 2), while over-weighting tactical consistency ($\omega_t > 0.4$) could result in overly rigid and predictable behaviors, hindering adaptive counterplay. The chosen weights strike a balance, enabling agents to engage in dynamic contests while adhering to recognizable tactical patterns.

TABLE 2. Stability Performance Against Dynamic Evolution

Method	Strategy Volatility Index ($\times 10^{-2}$)	Strategy Volatility Index (Std)	Frequency of Ball Control Switching (per round)	Frequency of Ball Control Switching (Std)
Diffusion+MARL	3.1	0.4	14.2	2.1
VAE-GAN	6.8	0.9	23.7	3.5
InterGen	5.4	0.7	20.3	2.9
TrajGen	6.1	0.8	21.8	3.2
STGAN	6.5	0.9	22.5	3.4

[Note: "Std" is the abbreviation for standard deviation.]

Ball possession gain is dynamically assigned based on ball possession and pass success rate. The team in possession receives positive rewards for completing effective passes or entering the final third, while negative rewards are applied for losing the ball or being intercepted. Spatial control gain is calculated by dividing the field into zones using a Voronoi diagram and calculating the percentage of the team's controlled area. This encourages agents to expand their effective range through movement and discourages densely packed or empty areas.

The tactical consistency term is used to constrain the degree of deviation of behavioral sequences from preset tactical patterns. The system pre-stores several typical tactical templates, such as "wing-center combination," "rapid transition," and "low-post dense defense." Each template defines key behavioral sequences and spatial distribution characteristics. During real-time deduction, the current behavioral stream is dynamically matched with the template, and a semantic similarity score is calculated as a reward adjustment factor. This score is determined by the cosine similarity between the behavioral sequence encoding and the template vector within the sliding time window:

$$sim_t = \frac{e_{\text{act}}(t) \cdot e_{\text{tmpl}}}{\|e_{\text{act}}(t)\| \|e_{\text{tmpl}}\|} \tag{10}$$

$e_{\text{act}}(t)$ is the embedded representation of the current behavior sequence, and e_{tmpl} is the target template encoding. High similarity triggers positive rewards, incentivizing the agent to maintain tactical consistency; low similarity introduces slight penalties to prevent strategy drift into atypical patterns. This mechanism does not force behavioral replication, but rather guides evolution in a direction consistent with common professional game logic.

IV. MULTI-DIMENSIONAL QUANTITATIVE ANALYSIS OF SCENARIO GENERATION AND COUNTERMEASURE SIMULATION

A. EXPERIMENTAL DATA

The experimental data comprise 200 professional matches from Europe's top five leagues (2021–2023 seasons), covering high-dimensional spatiotemporal trajectories in formations like 4-3-3 and 3-5-2. This dataset's size and diversity ensure the model learns generalized tactical distributions rather than domain-specific patterns. Collected using a multi-view camera system and sensors, it covers high-dimensional spatiotemporal trajectories of players in popular formations such as 4-3-3 and 3-5-2. The data include fields such as two-dimensional coordinates, instantaneous velocity, and ball possession status, with a sampling frequency of 10 Hz. The raw data are smoothed using a Kalman filter to remove outliers and noise. A manual annotation system then supplements the data with tactical labels (high-pressure pressing, wing penetration) and character behavior sequences (dribbling, passing, and assisting defense). To validate the model's generalization capabilities, the dataset was split into training, validation, and test sets with a 7:2:1 ratio. Perturbed test subsets were set for different tactical scenarios to simulate formation shifts and ball position variations in real-world matches. Furthermore, Voronoi field segmentation and convex hull analysis tools were introduced to quantify team spatial distribution characteristics, providing a structured evaluation benchmark for generation and deduction.

B. PHASED OPTIMIZATION IN TACTICAL TRAJECTORY GENERATION

To evaluate the denoising effectiveness of the diffusion model, a synthetic dataset was constructed to simulate real-world match trajectories, including initial noise distribution and tactical labeling. The physical error is calculated by measuring the deviation of the generated trajectory from the true kinematic constraints (velocity/acceleration thresholds). A phased sampling strategy is used to verify the correction capability of different denoising steps. The tactical score is based on a pre-trained tactic classifier and quantifies the semantic match between the generated trajectory and the target tactic label. Formation compactness is calculated by normalizing the convex hull area of the player positions, while speed variance analyzes the dispersion of a team's instantaneous speed. All metrics were repeatedly tested at multiple noise levels, using a sliding window to smooth data fluctuations. Finally, time-series correlation analysis was used to verify whether the evolutionary logic of each metric conformed to theoretical expectations.

Figure 5 demonstrates the phased optimization of the diffusion model for tactical trajectory generation. Figure 5(a) shows that the physical error rapidly decays from 0.9 initially, reflecting a coarse-to-fine process that prioritizes reconstructing basic kinematic constraints, followed by local fine-tuning. Concurrently, the tactical consistency score shows nonlinear, gradual growth, indicating the hierarchical learning of semantic guidance: starting with individual trajectory rationality, then establishing player synergy, and finally achieving alignment with pre-set tactical labels. Figure 5(b) analyzes team structure evolution: the formation compactness monotonically increases from 0.2 to 0.8 as the model corrects the spatial distribution from random to tactically defined. Simultaneously, the speed variance decays, signifying a transition from disordered to coordinated player movement.

C. CONDITIONALLY GENERATED TACTICAL TRAJECTORIES

To evaluate the semantic control capabilities of the conditional diffusion model in generating tactical trajectories, the paper used a phase space projection method to analyze the kinematic characteristics of the generated samples. Three typical tactical patterns were selected, and the horizontal width, vertical velocity, and spatial stretch index of each trajectory were collected as dynamic variables. Kernel density estimation was used to construct a two-dimensional probability manifold under each condition, and contour topology was extracted to characterize the distribution boundaries. For each sample type, the centroid position, covariance matrix, and principal component directions were calculated, and a 95% confidence ellipse was drawn to reflect the distribution's discreteness and main motion trends.

Figure 6 shows the kinematic distribution characteristics of trajectories generated under different tactical conditions. The horizontal axis represents the horizontal attack width (the standard deviation of the x-coordinates of the attacking players, unit: meters), and the vertical axis represents the vertical advance rate (the average y-speed of the players in possession, unit: meters per second). The color map represents the spatial stretch index (the rate of change in the distance between the striker's and midfielder's center of mass, unit: meters per second). Wing penetration samples (red circle) are concentrated in a lateral width of approximately 6–10 m, with a push rate of approximately 1–1.75 m/s. Their spatial stretch index is high, reflecting this tactic's emphasis on synergistic lateral stretching and flank advancement, resulting in a lateral expansion strategy. Central penetration samples (green circle) are concentrated in a narrower lateral width with a push rate of approximately 1.5–2.5 m/s. Their spatial stretch index is low, reflecting their vertical penetration capability. Low post defense samples (blue circle) are located in the lower left corner, with compact contour lines and an overall negative spatial stretch index, indicating a contracted formation. The direction of the principal axis of the 95% confidence ellipse reflects the direction of the principal component of the strategy.

Figure 7 shows the evolution of a soccer player's trajectory at different stages of processing, including the original trajectory, the noisy trajectory, and the denoised trajectory. Deviation plots are used to compare the differences between these trajectories. The original trajectory represents the actual movement path of a player on the soccer field. The original trajectory is smooth and free of external interference, accurately reflecting the player's movement patterns.

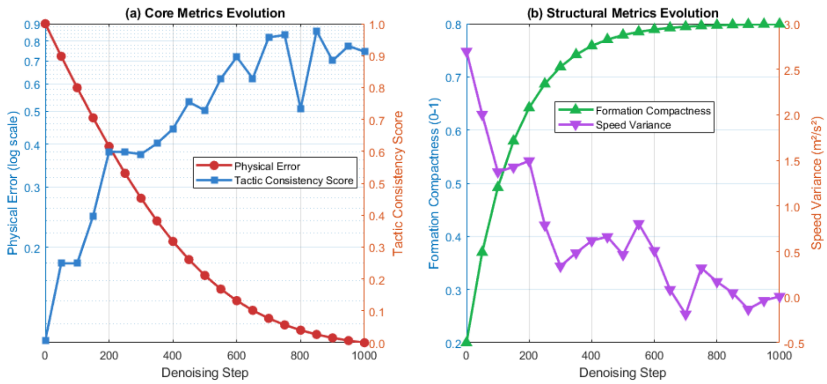

FIGURE 5. Phased Optimization Characteristics in Tactical Trajectory Generation

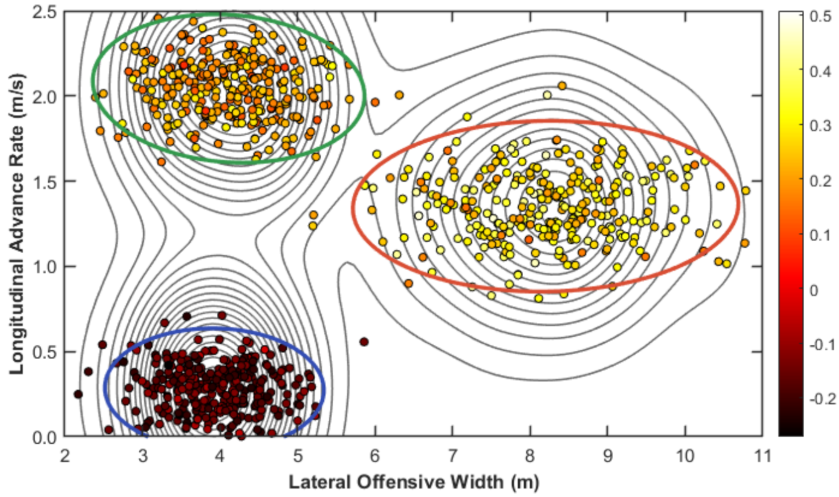

FIGURE 6. Phase Space Profiles of Tactical Trajectories

The trajectory after adding noise represents the potential deviations in trajectory data caused by sensor errors, data loss, or environmental factors in real applications. The noisy trajectory is more chaotic, exhibiting significant errors and irregularities, reflecting the impact of these external factors on the trajectory. The denoising algorithm aims to recover more realistic trajectories from noisy data. The trajectories in this section are very close to the original trajectories, demonstrating that the denoising process has effectively reduced the noise interference and restored the player's true movement trajectory.

D. TRAJECTORY FIDELITY ANALYSIS

Trajectory fidelity evaluation focuses on the approximation capability of the generated results in terms of geometric accuracy and distribution consistency. In the context of football tactical simulation, the acceptability of trajectory error must be assessed relative to the scale of the playing field and the typical distances between players. The pitch is approximately 100×64 m, and the average inter-player distance during structured play ranges from 10 to 20 m. Within this scale, an error of approximately 1 m represents a relative error of 1–5% in terms of player positioning, which is comparable to the inherent noise and uncertainty in real-world tracking data. Therefore, an absolute trajectory error (ATE) under 1.5 m is generally considered sufficient for capturing macroscopic tactical patterns, rather than requiring centimeter-level precision for micro-actions.

Figure 8 shows a comparison of trajectory fidelity for different generative models under three typical tactical phases, using metrics such as ATE and maximum mean difference (MMD). The Diffusion + multi-agent reinforcement learning (MARL) method proposed in this paper achieved the lowest ATE and MMD values across all scenarios, with mean ATE values ranging from 0.82 to 1.03 m and mean MMD values ranging from 0.0215 to 0.0267. This indicates that its generated trajectories are closest to real data in terms of frame-by-frame accuracy and overall distribution. The error in the ball-control phase is generally lower than that in positional offense and rapid counterattacks.

E. TACTICAL CONSISTENCY EVALUATION

Tactical consistency evaluation focuses on the degree of semantic alignment between the agent's behavior in the generated scenario and the pre-set tactical intent. Tactical execution accuracy is defined as the proportion of matches

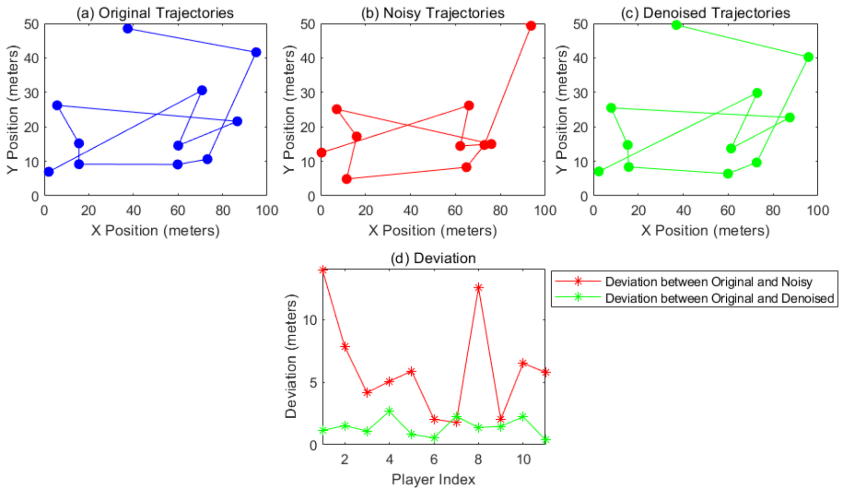

FIGURE 7. Trajectory Diffusion, Denoising Process, and Deviation

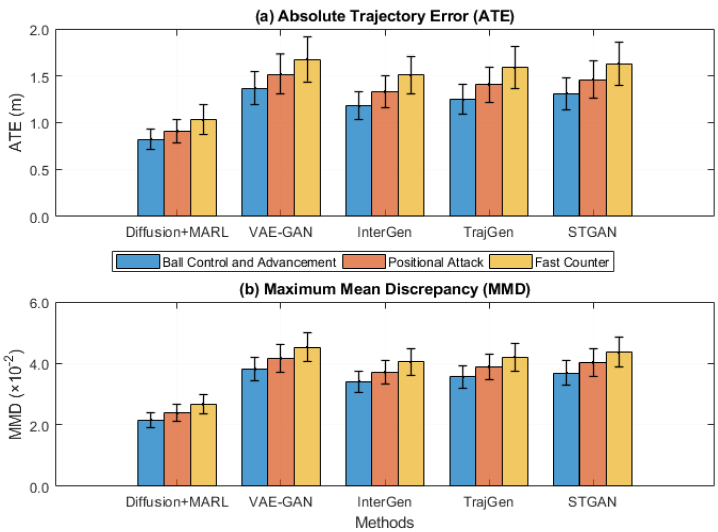

FIGURE 8. Trajectory Fidelity Comparison (ATE and MMD) with Baseline Methods

between the agent's actual behavior pattern at the beginning of the simulation and the standard behavior sequence corresponding to the input tactical label. This is quantified using dynamic time warping and behavior encoding similarity metrics. Role behavior fit measures the degree to which individual behavior conforms to the role's functional specifications. A pre-trained role classifier is used to reverse-infer the generated action sequence and calculate a consistency score between the predicted category and the actual role. The evaluation was conducted independently under three different pressing intensity conditions: low, medium, and high, corresponding to situations with one, two to three, and more than three players, respectively, to test the robustness of the strategy guidance mechanism under different pressure loads. All compared models were run ten times on the same test set, and the average was taken to eliminate random bias.

Figure 9 shows the tactical consistency performance of different models at multiple levels of pressing intensity, covering two key metrics: tactical execution accuracy and role behavior fit. The proposed method significantly outperforms the comparison model at low, medium, and high pressing intensities, demonstrating the robustness and contextual adaptability of its tactical guidance mechanism. At high pressing intensity, tactical execution accuracy and role behavior fit reach 80.1% and 83.4%, respectively. As pressure intensity increases, the performance of all models decreases, reflecting the increasing difficulty of behavioral control in highly confrontational environments. However, the proposed method maintains high performance, demonstrating that diffusion generation and multi-agent strategies effectively maintain tactical consistency. Generative models such as VAE-GAN and STGAN, due to their lack of explicit semantic anchoring, are prone to behavioral drift in complex confrontations, potentially leading to imbalances such as forwards failing to track back or defenders overcommitting.

F. STABILITY OF DYNAMIC EVOLUTION OF THE CONFRONTATION

The dynamic stability of the confrontation process is quantitatively assessed using the strategy volatility index and the frequency of ball possession switches. The strategy volatility index is defined as the normalized weighted variance of the frequency of agent behavior pattern switches between consecutive time steps. It reflects the consistency of individual decisions and the degree of strategic convergence. A low value indicates a stable behavioral sequence. The possession switch frequency measures the total number of times the ball transfers between the red and blue teams during each simulation round. It is used to measure the rationality of the game rhythm and the smoothness of the offense-defense transitions. A high frequency suggests an imbalance in the confrontation or strategy volatility, while a low frequency may reflect a stalemate or insufficient exploration. The evaluation was conducted over 100 rounds of simulation, each lasting 90 time steps, covering the full attack and defense cycle. All methods were run with the same initial distribution to ensure comparability.

Table 2 shows the stability of five methods in adversarial dynamic evolution over 100 simulation rounds. The policy volatility index reflects the continuity of the agent's behavior sequence and the degree of policy convergence. The Diffusion+MARL framework proposed in this paper achieved the lowest value (3.1×10^{-2}), indicating a high degree of consistency in its decision-making process, avoiding frequent and meaningless behavior switching, and demonstrating good policy stability. In contrast, generative models like VAE-GAN and STGAN decouple trajectories from strategies, leading to significant behavioral jitter and strategic fluctuations during simulation. The frequency of possession shifts reflects the fluidity of offensive and defensive transitions and the balance of play. The proposed method averages 14.2 shifts per round, which is within a reasonable range and demonstrates a balanced and rhythmic interaction between offense and defense. Other methods generally exceed 20, exposing issues of rapid possession turnover and fragmented tactical organization.

G. SCENARIO DIVERSITY TESTING

Scenario diversity is quantitatively assessed using the dual metrics of trajectory diversity score and strategic entropy. The trajectory diversity score is based on cluster analysis of the generated trajectory set. It uses dynamic time warping distance to measure the differences between sequences. After performing hierarchical clustering, the average distance between clusters is calculated to reflect the breadth of coverage of spatiotemporal patterns. Strategic entropy constructs a strategic vector distribution by discretizing tactical behavior sequences and calculates its uncertainty using the information entropy formula. High entropy values indicate that the system can generate rich and balanced tactical combinations. To test robustness, a ± 2-meter ball position offset was applied to the initial possession position, and a ± 1-row formation offset was introduced to the formation position to simulate the subtle state differences in actual combat. Two hundred scenarios were independently generated under each perturbation condition, and the mean and standard deviation of the indicators were statistically analyzed to eliminate randomness.

Table 3 shows the scenario diversity performance of each model under two types of perturbations: ball position offset (± 2 m) and formation offset (± 1 row). Trajectory diversity scores and strategic entropy jointly characterize the breadth and richness of the generated tactical space. The proposed Diffusion+MARL framework achieved the highest diversity scores and strategic entropy values under both perturbation conditions, ranging from 1.86 to 1.88 and 3.22 to 3.26, respectively. This demonstrates that its generated results cover a wider range of spatiotemporal patterns and tactical combinations, demonstrating strong generalization capabilities. Methods like VAE-GAN and STGAN suffer from a loose latent space structure, making generated trajectories prone to local patterns and limiting diversity. While InterGen and TrajGen introduce interaction modeling, their responses to changes in initial conditions are rigid, resulting in limited improvement in diversity. Notably, Diffusion+MARL achieves minimal fluctuations in metrics across different perturbation types, demonstrating its high robustness to minor state variations and its ability to maintain the structure and discreteness of the generated space.

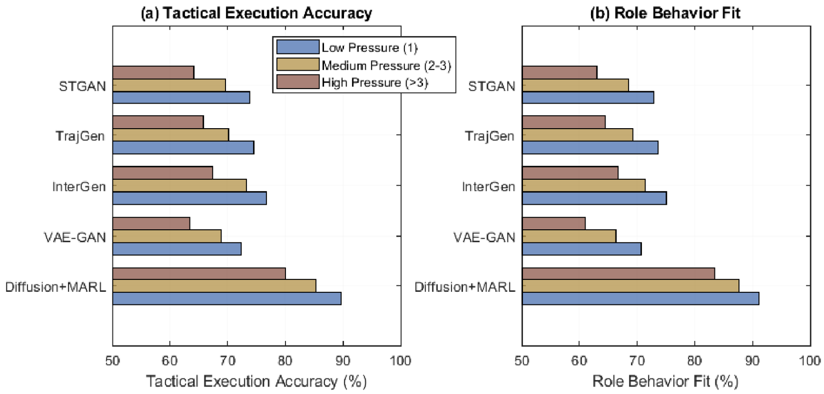

FIGURE 9. Tactical Consistency Performance at Multiple Levels of Pressing Intensity

TABLE 3. Scenario Diversity Performance

Method	Perturbation Type	Trajectory Diversity Score	Trajectory Diversity Score (Std)	Strategic Entropy (bit)	Strategic Entropy (Std)
Diffusion+MARL	±2 m Ball Position Offset	1.86	0.11	3.22	0.13
Diffusion+MARL	±1-Row Formation Shift	1.88	0.09	3.26	0.10
VAE-GAN	±2 m Ball Position Offset	1.31	0.16	2.49	0.22
VAE-GAN	±1-Row Formation Shift	1.33	0.14	2.53	0.20
InterGen	±2 m Ball Position Offset	1.48	0.14	2.65	0.2
InterGen	±1-Row Formation Shift	1.50	0.12	2.69	0.18
TrajGen	±2 m Ball Position Offset	1.40	0.15	2.56	0.21
TrajGen	±1-Row Formation Shift	1.42	0.13	2.60	0.19
STGAN	±2 m Ball Position Offset	1.37	0.17	2.43	0.23
STGAN	±1-Row Formation Shift	1.39	0.15	2.47	0.21

V. CONCLUSION

This research constructs a football tactical simulation framework that integrates a diffusion generative model with multi-agent reinforcement learning—a novel combination directly transferable to the industry for developing sophisticated, multi-agent systems that coordinate complex manufacturing and supply chain processes. This system implements a closed-loop process that learns tactical distributions from high-dimensional trajectory data and drives dynamic adversarial deduction. This approach overcomes the limitations of traditional generative models in terms of semantic controllability and structural rationality. By introducing tactical label guidance and a graph attention mechanism, it precisely controls the intent expression and team coordination structure of generated trajectories during the denoising process. Generated scenarios serve as the initial state for simulations. Role-based strategy space partitioning and behavioral cloning pre-training significantly improve the rationality of multi-agent decision-making starting points. In a red-blue confrontation environment, strategy optimization based on the MAPPO algorithm, combined with rewards for ball possession, space, and tactical consistency, encourages both offense and defense to autonomously evolve high-level game behaviors under diverse initial conditions. Experimental validation demonstrates that the method achieves mean ATE values of 0.82–1.03 m and mean MMD values of 0.0215–0.0267. Tactical execution accuracy and role-behavior consistency under high-pressure conditions reach 80.1% and 83.4%, respectively. This framework boasts high trajectory fidelity and tactical consistency, outperforming existing mainstream methods in terms of confrontational stability and scenario diversity. This research provides a new paradigm for football tactical simulation that is generative, deducible, and interpretable, advancing the transition of intelligent systems from passive backtracking to active rehearsal. It has the potential for widespread application in tactical design, risk assessment, and confrontational strategy generation; more broadly, this advancement offers the industry a powerful tool for strategic decision-making, enabling the active rehearsal of new production strategies, risk assessment for supply chain disruptions, and the design of dynamic resource allocation plans in a smart factory environment.

AUTHOR CONTRIBUTIONS

Juhai Wang designed, collected and analyzed the data, and drafted the manuscript. Juhai Wang conducted the study, critically revised the manuscript for important intellectual content, and gave final approval of the version to be published. Juhai Wang participated fully in the work, take public responsibility for appropriate portions of the content, and agreed to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

CONFLICTS OF INTEREST

The author declares no conflict of interest.

FUNDING

This research received no external funding.

ACKNOWLEDGEMENTS

Not applicable.

REFERENCES

[1] U. S. Rozmatovich and D. Elyor, "Teaching attack tactics to football players," International Journal of Research in Commerce, IT, Engineering and Social Sciences, vol. 16, no. 10, pp. 125-132, 2022, [Online]. Available: https://gejournal.net/index.php/IJRCIESS/article/view/1119.

[2] J. M. M. M. Junior, R. G. de Souza Vale, D. B. de Mello, R. A. M. Nunes, L. A. dos Santos, and G. Rosa, "Effects of scoring method on the physical, technical, and tactical performances during football small-sided games (SSGs): A systematic review," Retos: Nuevas Tendencias en Educación Física, Deporte y Recreación, vol. (49), pp. 961-969, 2023, doi: 10.47197/retos.v49.98459.

[3] B. Wang, "Use of network technologies in teaching football tactics: Cooperation, engagement, creativity," Interactive Learning Environments, vol. 32, no. 9, pp. 5078-5088, 2024, doi: 10.1080/10494820.2023.2209608.

[4] M. Ötting and D. Karlis, "Football tracking data: A copula-based hidden Markov model for classification of tactics in football," Annals of Operations Research, vol. 325, no. 1, pp. 167-183, 2023, doi: 10.1007/s10479-022-04660-0.

[5] B. Low, R. Rein, D. Raabe, S. Schwab, and D. Memmert, "The porous high-press? An experimental approach investigating tactical behaviours from two pressing strategies in football," Journal of Sports Sciences, vol. 39, no. 19, pp. 2199-2210, 2021, doi: 10.1080/02640414.2021.1925424.

[6] Q. Shao, "Virtual reality and ANN-based three-dimensional tactical training model for football players," Soft Computing, vol. 28, no. 4, pp. 3633-3648, 2024, doi: 10.1007/s00500-024-09634-x.

[7] D. Steffi, S. Mehta, and K. A. Venkatesh, "Bayesian probabilistic modeling in robosoccer environment for robot path planning," Bulletin of Electrical Engineering and Informatics, vol. 13, no. 1, pp. 465-472, 2024, doi: 10.11591/eei.v13i1.6080.

[8] G. A. Whitaker, R. Silva, D. Edwards, and I. Kosmidis, "A Bayesian approach for determining player abilities in football," Journal of the Royal Statistical Society: Series C (Applied Statistics), vol. 70, no. 1, pp. 174-201, 2021, doi: 10.1111/rssc.12454.

[9] A. Dutta, H. Saikia, J. Gogoi, and D. Bhattacharjee, "Forecasting the opening goal in second-half of a football match: Bayesian and frequentist perspectives," Computational statistics, vol. 40, no. 4, pp. 1959-1984, 2025, doi: 10.1007/s00180-024-01558-2.

[10] J. Won, D. Gopinath, and J. Hodgins, "Physics-based character controllers using conditional vaes," ACM Transactions on Graphics (TOG), vol. 41, no. 4, pp. 1-12, 2022, doi: 10.1145/3528223.3530067.

[11] X. Cheng and P. Zhang, "Enhanced Soccer Training Simulation Using Progressive Wasserstein GAN and Termite Life Cycle Optimization in Virtual Reality," The International Arab Journal of Information Technology, vol. 21, no. 4, pp. 549-559, 2024, doi: 10.34028/iajit/21/4/1.

[12] X. Zhan, J. Sun, Y. Liu, N. J. Cecchi, E. Le Flao, O. Gevaert, et al., "Adaptive machine learning head model across different head impact types using unsupervised domain adaptation and generative adversarial networks," IEEE Sensors Journal, vol. 24, no. 5, pp. 7097-7106, 2024, doi: 10.1109/JSEN.2023.3349213.

[13] W. Zhu, X. Ma, D. Ro, H. Ci, J. Zhang, J. Shi, et al., "Human motion generation: A survey," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 4, pp. 2430-2449, 2023, doi: 10.1109/TPAMI.2023.3330935.

[14] M. Zhang, Z. Cai, L. Pan, F. Hong, X. Guo, L. Yang, et al., "Motiondiffuse: Text-driven human motion generation with diffusion model," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 46, no. 6, pp. 4115-4128, 2024, doi: 10.1109/TPAMI.2024.3355414.

[15] A. Croitoru F, V. Hondru, T. Ionescu R, and M. Shah, "Diffusion models in vision: A survey," IEEE transactions on pattern analysis and machine intelligence, vol. 45, no. 9, pp. 10850-10869, 2023, doi: 10.1109/TPAMI.2023.3261988.

[16] L. Kong, D. Pei, R. He, D. Huang, and Y. Wang, "Spatio-temporal player relation modeling for tactic recognition in sports videos," IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 9, pp. 6086-6099, 2022, doi: 10.1109/TCSVT.2022.3156634.

[17] B. Jiang, J. Du, C. Jiang, Z. Han, and M. Debbah, "Underwater searching and multiround data collection via AUV swarms: An energy-efficient AoI-aware MAPPO approach," IEEE Internet of Things Journal, vol. 11, no. 7, pp. 12768-12782, 2023, doi: 10.1109/JIOT.2023.3336055.

[18] Y. Guan, S. Zou, H. Peng, W. Ni, Y. Sun, and H. Gao, "Cooperative UAV trajectory design for disaster area emergency communications: A multiagent PPO method," IEEE Internet of Things Journal, vol. 11, no. 5, pp. 8848-8859, 2023, doi: 10.1109/JIOT.2023.3320796.

[19] F. V. Coimbra and M. R. O. de Albuquerque Máximo, "Team modeling with deep behavioral cloning for the RoboCup 2D soccer simulation league," IEEE Latin America Transactions, vol. 21, no. 2, pp. 288-294, 2023, doi: 10.1109/TLA.2023.10015221.