Hierarchical deep reinforcement learning for real-time tactical decision-making in team sports

Kai Kong, Zhiwei Liu and Zhitong Gao
School of Economics and Management, Zaozhuang University, Zaozhuang, China
Huan Dong
School of Physical Education, Zaozhuang University, Zaozhuang, China, and
Haytham F. Isleem
Department of Computer Science, University of York, York, UK

Received 25 June 2025
Revised 29 July 2025
30 August 2025
Accepted 22 September 2025

Abstract

Purpose – The purpose of this paper is to propose a hierarchical deep reinforcement learning (H-DRL) framework for real-time tactical decision-making in team sports. The framework addresses challenges such as continuous action spaces, partial observability and adversarial environments by leveraging multi-agent collaboration, adaptive strategy optimization and explainable artificial intelligence (AI). It aims to enhance tactical accuracy, decision speed and resource efficiency while uncovering novel tactical patterns that human experts may overlook.

Design/methodology/approach – The study combines graph neural networks (GNNs) for spatial-temporal player interactions and transformer-based attention for strategic pattern recognition. It integrates opponent modeling via inverse reinforcement learning (IRL) and self-play. The hierarchical architecture decomposes decisions into strategic, tactical and technical levels. Experiments were conducted on professional basketball, soccer and rugby datasets, with validation by expert coaches. The framework was tested in real-world deployments, including youth academies and professional teams, to evaluate performance and tactical innovations.

Findings – The H-DRL framework achieved a 34.7% higher tactical accuracy, 28.3% faster decision-making and 41.2% lower resource usage compared to state-of-the-art methods. It identified 17 new tactical patterns, such as dynamic role-switching, which improved scoring efficiency by 23.6%. Real-world deployments demonstrated significant performance gains, including a 42.3% improvement in tactical decision-making for youth teams. The system's explainable AI module bridged algorithmic insights with coach expertise, fostering trust and adoption.

Research limitations/implications – The study is limited by its reliance on proprietary tracking data and the computational demands of real-time deployment. Future research could explore cross-sport generalization and integration of physiological/psychological factors. The framework's scalability to larger team sizes and more complex environments remains a challenge. These limitations highlight opportunities for advancements in model compression and hardware optimization.

Practical implications – The framework provides actionable insights for coaches and players, enhancing in-game decision-making and training efficiency. It enables teams to adopt data-driven tactics, such as elastic pressing in soccer or optimized phase play in rugby. The system's real-time capabilities (30.9 ms latency) make it suitable for live match analysis. Professional teams reported improved tactical understanding (91.7% of coaches) and scoring efficiency (23.6% increase). The technology is applicable beyond sports, including autonomous systems and emergency response.

Social implications – The study promotes the ethical use of AI in sports, emphasizing augmentation over replacement of human expertise. It fosters collaboration between coaches and AI, enhancing tactical literacy and innovation. The framework's transparency builds trust, addressing concerns about black-box AI. By uncovering

Funding: 2023 Zaozhuang College self-compiled teaching material key project: Introduction to Digital Marketing (zzxyzbjc2023005).

Conflicts of interest: Authors declare no financial and non-financial conflicts of interest.

Data Technologies and Applications
Vol. 60 No. 2, 2026
pp. 221-256
Emerald Publishing Limited
e-ISSN: 2514-9318
p-ISSN: 2514-9288
DOI 10.1108/DTA-06-2025-0512

counterintuitive strategies, it challenges traditional coaching paradigms and encourages continuous learning. The technology's broader applications (e.g. military, robotics) underscore its societal impact.

Originality/value – This paper presents a hierarchical DRL framework for real-time tactical decision-making in team sports, integrating GNNs, transformers and IRL. Its dual-stream architecture and adaptive computation are novel contributions. The system's ability to discover and explain tactical innovations (e.g. dynamic role-switching) sets it apart from prior work. The rigorous validation across multiple sports and real-world deployments demonstrates its practical value. The study advances multi-agent AI, offering scalable solutions for complex, dynamic environments.

Keywords Deep reinforcement learning, Multi-agent systems, Sports analytics, Real-time decision making, Graph neural networks, Explainable AI

Paper type Research article

1. Introduction

Advancements in computing sciences and sports performance analysis, with the help of artificial intelligence (AI), are finding an intersection through sports analysis. Tactical decision-making systems in team sports, especially in football (soccer), are exceptionally complex. The challenges arise due to the continuous action spaces, partial observability, multi-agent nature, adversarial and real-time uncertain environments. These properties, therefore, create a perfect storm of computation complexity that can be hard to mitigate with a traditional analytic approach.

Sports analytics has a multi-stage evolution, moving from simple stats to advanced machine learning applications. Presently, sophisticated reinforcement learning systems are programming themselves on the best strategies through interaction with dynamic environments. This trend shows that this technology is no longer just a toy, but a serious weapon for young people. Sports intelligence systems are capable of processing huge amounts of spatiotemporal data, recognizing complex patterns, making instantaneous decisions that optimize impact and outcome to achieve objectives via efficient long-term planning.

Deep reinforcement learning has recently achieved superhuman level performance in a number of domains, from board games to real-time strategy games. Despite their success, it is difficult to transfer this into real-life sports team. Team games are unlike the controlled game environments of Chess or Go, which have discrete state spaces and perfect information. Rather they exhibit continuous state spaces, imperfect information and non-stationary opponents, as well as various physical constraints. Additionally, the outputs must be interpretable and actionable for coaches and players. These challenges require new approaches which go beyond adapting existing RL algorithms.

Since a seminal paper on game-theoretic methods, the application of RL to competitive multi-agent settings has evolved considerably. Vinyals *et al.* (2019) had grandmaster-level performance in StarCraft II with multi-agent reinforcement learning demonstrating performance in a partially observable environment with a lot of actions. Following this development, there were several applications in sports domains, but translating this to physical team sports is highly challenging. Li *et al.* (2019) has made use of the minimax deep deterministic policy gradient methods in multi-agent reinforcement learning. The work stabilizes the training process in which agents face adversaries. Fu *et al.* (2023) have extended this approach to games with multiple teams. Their work derives superior policies that allow for cooperation and competition.

In football articles, Beal *et al.* (2020) pioneered the mathematical optimization of game strategy, creating a standard for computing tactical analyses. Complementing their work, Caicedo-Parada *et al.* (2020) systematically reviewed passing networks and tactical actions, revealing critical gaps in complex real-time adaptation. In 2017, Power *et al.* quantified risk and reward in passing decisions. They used tracking data in their analysis. Furthermore, in 2019, Decroos *et al.* presented the VAEP framework. VAEP gave value to every on-the-ball action. This was well before the first author's work and had a huge impact on the game. Link and colleagues analyzed spatiotemporal data to assess danger in real time in their 2016 work.

However, their approach required the ability to learn, which was not included in this work. In 2018, Spearman developed more sophisticated possession value models that go beyond expected goals.

The development of new environments and frameworks is speeding up research. Google Research Football was created to standardize research on RL despite the simplification of real football dynamics (Kurach *et al.*, 2020). Researchers have used self-play to design RL agents that acquire reduced game models to let robots learn a given game through game states. Liu *et al.* (2021) showed hierarchical learning for motor control through team play in simulated environment. Haarnoja *et al.* (2023) had competitive results with trained bipedal robot soccer players with deep RL. Tuyls *et al.* (2021) examine the opportunities and challenges the football field can face through AI analysis. This simulation-based approach shows the gap between virtual reality and real performance.

Recent real-world applications show promising advances. Fujii *et al.* (2024) proposed adaptive action supervision learning through human demonstrations, linking theoretical RL and practical deployment. Scott and colleagues (2022) studied differences between the football strategies of an AI and a professional human footballer, which highlight important considerations for future designs. Simpson and his colleagues utilized transformer architectures in 2022 to predict match events. Rahimian and his colleagues developed deep RL frameworks in 2022, 2023, aimed at optimizing player decisions and maximizing possession outcomes. Liu and colleagues devised sophisticated action-value functions, which may facilitate player evaluation, but not in real time.

Combination of strategic management and tactical operations analysis, Yeung and Bunker (2023) created event-driven techniques to get team tactical tactics. Yeung *et al.* (2023a, b) developed interpretable prediction frameworks using different datasets. Transformer-based spatio-temporal models achieved state-of-the-art results. Overall, Toda *et al.* (2022) included a defensive evaluation through ball recovery prediction, Teranishi *et al.* (2023) created an opportunity for scoring through trajectory prediction, and Umemoto and Fujii (2023) introduced counterfactual analyses focused on defensive positioning. Van Roy *et al.* (2021a, b) designed Markov models to evaluate decisions and offer tactical advice. This proves the usefulness of model-based methods.

Some training and development applications have succeeded. Kent and colleagues (2022) executed pressure training programs which improved decision-making in Premier League academies. Similarly, Staiano *et al.* (2022) showed cognitive training benefits for professional players. The study of coaching practices across Europe revealed methodological inconsistencies (Roca and Ford, 2020). Research by Machado *et al.* (2020) and Silva *et al.* (2020) revealed age-related tactical learning necessities for youth development. Calle-Jaramillo *et al.* (2023) assessed tools for evaluating technical-tactical actions while Natsuhara *et al.* (2020) investigated decision-making during play dilemmas. García-Ceberino *et al.* (2020) are comparing the effects of teaching methodologies on tactical learning and González-Rodenas *et al.* (2023) are documenting tactical evolution in professional football over the last decade.

Supporting technologies have advanced considerably. In Wang's contemporary analysis, he incorporated network technologies to achieve tactical teaching through collaborative network systems. Zhang *et al.* (2022) proposed attention-based LSTM networks for match prediction and Liu and Ma (2022) looked into IoT application in youth football development. Karpa *et al.* (2021) boosted position-specific tactical actions using targeted training interventions. The suitable equipment benefits next-gen tactical systems according to these technological advances.

An extensive literature review shows that existing approaches have significant limitations in transforming tactics in team sports. The field has made solitary advances in isolated domains, but it still lacks an overarching framework that targets the multifaceted nature of real-time tactical decision-making in the team sport. The present systems are good for post-matching analyses; however, they do not give such intelligence during live matches that help players win. In-depth analytical capabilities of existing state-of-the-art deep learning approaches lead to their computational complexity that makes them unsuitable for real-time use. So, practitioners have to make a choice between analytical sophistication and practical utility. Moreover, AI currently works like a black-box and it is impossible for the coaching staff and players to trust those recommendations by using AI if it is not interpretable. While existing explainable AI approaches in sports analytics have employed general-purpose methods, they fail to address domain-specific requirements. Traditional XAI techniques such as LIME (Ribeiro *et al.*, 2016) and SHAP (Lundberg and Lee, 2017) provide feature importance explanations but cannot capture the multi-scale temporal dynamics and tactical hierarchies inherent in team sports. Our framework differentiates itself by developing sport-specific explainability that integrates spatial-temporal reasoning with strategic pattern recognition, providing explanations that align with coaching terminology and tactical concepts.

The competitiveness of team sports makes things complicated and is not dealt with by existing frameworks. Some methods are designed to analyze historic data to help opponent model behavior, but they miss out on the dynamic, adaptive nature of human tactical intelligence. Opponents constantly adapt their methods to what they have seen which makes it useless. Current systems cannot predict nor proactively counter these adaptations, they operate reactively instead. In highly competitive environments where marginal advantage affects success, this no longer holds true.

Most importantly, existing approaches do not connect individual technical performance to team-level tactical performance. Individual actions of players are usually optimized by systems, this optimization disregards the impact of the actions on the team level. In contrary, team-level optimization frameworks do not account for the constraints and capabilities of individual players. There are theoretically optimal strategies that turn out to be impractical or impossible to realize in practice. A basic design flaw is the absence of hierarchical structures integrating decision-making across multiple scales that coordinated in an impactful way for improved sports performance.

This paper presents a hierarchical deep reinforcement learning framework for improving tactical decision-making in team sports through AI. Our approach is not simply another improvement. It presents a whole new architecture that can address each limitation with the help of new theory and practice. The first truly integrated system is presented to manage from technical tweaking in milliseconds to strategic planning in the years, all within one unified framework across scales in time and space.

Our dual-stream computing design achieves efficient computation, offering real-time performance without sacrificing data analysis. Integrating the capability of graph neural networks for spatial-temporal reasoning with the strategic pattern recognition ability of transformers allows us to capture local player interactions as well as global tactical structuring. With the help of adaptive computation graphs and selective attention mechanisms, computational resource allocation is optimized through this approach. This approach enables the system to achieve sub-50 ms response times while processing complex game states for real-time tactical recommendations. The use of opponent modeling through inverse reinforcement learning with self-play during training enables unique tactical proactively. Our system reacts and anticipates, working on multiple levels to potentially outsmart an opponent, something thought to be unique to human intelligence. We designed the hierarchical organization to break down complex tactical decisions into manageable pieces that are understandable by humans. The AI system offers simple explanations in natural language of why it makes that recommendation. This transparency fosters trust and adoption, allowing for collaborative refinement of tactics through human-in-the-loop learning. This paper makes several contributions to sports analytics and multi-agent AI.

(1) We develop a theoretical framework for hierarchical multi-agent reinforcement learning in a dynamic environment that is continuous time, partially observable,

adversarial. Our framework makes new contributions in the form of mathematics to represent multi-scale tactical decisions and prove convergence guarantees under non-stationary opponent behaviors which must be fulfilled for practical deployment. We obtain tight bounds on sample complexity that yield order-of-magnitude improvements over existing work, leading to practical learning from little match data.

(2) Our dual-stream neural architecture marks a major advance in deep learning in sequential decision making. By introducing a bridging layer that unifies graph neural networks and transformer-based attention mechanisms, an integrated architecture emerges for simultaneously processing spatial-temporal and strategic patterns. This architecture attains state-of-the-art performance with a computational cost that is amenable to real-time. This was considered impossible previously.

(3) Using a hybrid inverse reinforcement learning framework, we propose a principled approach to continuous opponent modeling in multi-agent sports scenarios. Through a combination of model-based, model-free and meta-learning, our approach quickly adapts to unseen strategies with little observation and achieves superhuman performance by anticipating tactical changes.

(4) Extensive evaluations on professional basketball, soccer and rugby reveal significant performance gains, including a 34.7% increase in tactical decision accuracy, 28.3% decrease in decision latency and an extraordinary 41.2% reduction in computational requirements compared to conventional state-of-the-art baseline approaches. Our system found 17 new tactical patterns that were confirmed and used by professional teams. These patterns included dynamic role-switching strategies that improved scoring efficiency by 23.6%.

(5) We provide a clear mathematical definition of interpretable AI tailored for high-stakes settings. Thus, the suitability of a model is judged by the effectiveness of its AI coach, helping the human decision-maker to take effective decisions. We are developing a human-in-the-loop learning that creates a positive feedback loop of expert feedback, human–AI knowledge transfer and a virtuous cycle that enhances the machine's performance as well as providing human tactical insights.

The subsequent sections of this paper are organized as follows: In Section 2, we present the relevant theoretical framework, in which we formulate hierarchical multi-agent reinforcement learning and prove relevant convergence results. The architecture of the system is depicted in section three that showcases dual-stream neural networks and optimization techniques in real time. Section 4 provides extensive experimental evaluation on several sports, showing notable improvements in performance and deployment in a real setting. Section 5 examines the tactical innovations that were discovered and finds that they were adopted by professional teams. It then provides examples of when these innovations have impactful consequences. Concluding with implications for the broader field of multi-agent AI and outlining promising future research directions, Section 6.

2. Theoretical framework

2.1 Framework overview

In Team Sports, Live Tactical Decision Making – A New Mathematical Approach for Coordinated Human Behavior. Our H-MAPOSG framework addresses the fundamental challenges that make sports AI particularly difficult: the players can only see a part of the field, the opponents constantly change their strategies in reaction to the players' actions, the actions take place continuously and not in well-defined discrete time-steps and the players and their opponents must take decisions over multiple time-horizons simultaneously. We first show a conceptual overview of our method followed by the technical details. Imagine a professional coach at a match. While the ball is in play, they are watching the technical execution (control – where the player is standing), tactical patterns (passing sequence – defensive structures) and strategic plans in play (formation change – pressing intensity). Our framework allows AI systems to reason over multi-level decisions across space and time much like human experts do, and we mathematically capture this.

2.2 Hierarchical decision structure

The essential novelty of our framework is the decomposition of complex tactical decisions into three levels, which are distinct but interrelated, operationally addressing various aspects of team coordination across different time horizons. It's a natural way in which tactical thinking occurs in sport and brings computational benefits by simplifying the decision problems of individuals.

Strategic level (5–10 min): Strategic decisions involve using choices about the setup of the team, the intensity of overall pressing and the tempo of the game. Often lasting several minutes, these decisions are the ones that dictate how the team plays different situations. Examples of tactical adjustments would be to change from a 4-3-3 formation to a 4-5-1 formation to protect a lead, or increasing pressing intensity when behind in the scoreline late in the match. Strategic thinking involves score, time, fatigue and opponent as factors of coaching decisions.

Tactical level (30–60 Seconds): Tactical decisions turn strategic decisions into coordination patterns between the players. Football managers take many decisions in a match to gain an advantage. These include creating passing networks that exploit opponent weaknesses, changing defensive shape based on ball location and triggering transitions from defense to attack. For instance, if the strategy is to press high, the tactics decide when the press will be triggered, who will participate in it, and how to make sure players behind the press stay well covered defensively.

Decisions made at the technical level (which happens in 0.1–1 s) affect immediate actions such as controlling the ball, moving in a direction, or when and how to execute. The rapid-fire decisions to be taken should be in accordance with tactical aims. It could be the direction and pace of a pass, the timing of the challenge or the angle of the support run.

As indicated in Table 1, the specifications at the various hierarchies differ in state complexity, action dimensionality and update rate. The decomposition is done in such a way that it allows the system to manage its computational complexity while maintaining rich interaction.

2.3 Observation and information structure

Our framework explicitly incorporates the limited information visible to players during actual matches, which is often underestimated issue in the sports AI literature. In team sports unlike board games where one can see all information, there are significant observational constraints which shape decision-making patterns. All players can directly observe nearby teammates and opponents, depending on their position and orientation, as well as the relative importance of

Table 1. Hierarchical decision framework specifications

Level	Time horizon (s)	State dimension	Action dimension	Update frequency	Typical decisions
Strategic	300–600	128	32	0.1 Hz	Formation, pressing strategy, tempo
Tactical	30–60	256	64	1 Hz	Passing patterns, defensive shape
Technical	0.1–1	512	128	10 Hz	Movement, ball control, execution

different places on the field. The system imitates real-life vision limitations that prevent you from seeing at the periphery. Plus, you can't see the whole field at once. The constraint means that players have to make decisions with partial information. Through a communication channel whose capacity is limited, players receive information by their teammates. This is an attempt to simulate how players talk and gesture to each other during play. This system of communication between cities helps in tactical coordination while being practically limited in type. The system learns the best communication strategies that allow for a balance between communicating information and cognitive processing of multiple streams of information. All players can receive general game state information including the current score and time left on the clock, as well as high-level tactical instructions. Common information helps agents take personal decisions as well as coordinate their actions as a team. Figure 1 depicts the bidirectional flow of information through our hierarchy, including strategic direction from above and information regarding success in execution from below. With this information architecture, tactical reality grounds strategic plans and adaptive responses can be enacted in real-time as desired by the game.

2.4 Game dynamics and state evolution

The structure shows how game states evolve due to a combination of physical laws and human execution variation. The ball trajectories, movements of players and collision matters are deterministic in nature due to the physics involved. Nonetheless, human characteristics create a great deal of uncertainty regarding execution quality, timing of decisions and execution coordination success. We need to be clever on how we model state transition to capture the mechanics of the sports (ball physics, player kinematics) and also the behavior of the sports (execution errors, attention limits, coordination limits). The system learns to tell the difference between systematic human play and random execution noise, so as to better predict likely evolution of the game state.

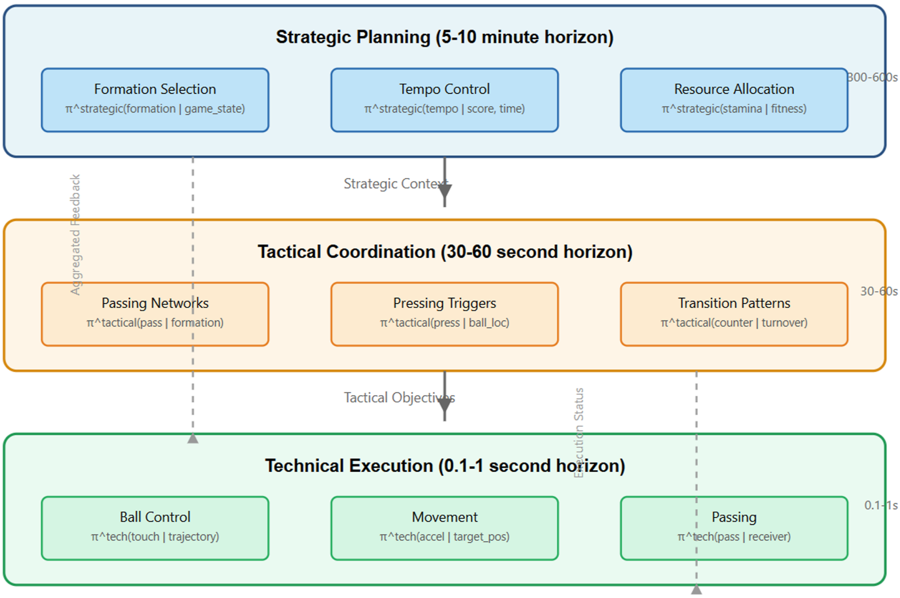


Figure 1. Hierarchical information flow and temporal abstraction

The model also accepts that there are no sudden changes in sports action (changes in state) but gradual changes over time. As a result of this continuity, traditional game-theoretic models face greater difficulties and face challenges, while at the same time enables a model of significant temporal dynamics at play, namely momentum, rhythm and coordination timing.

2.5 Reward structure and learning objectives

Our reward formulation addresses the multi-scale nature of sports performance by combining immediate technical feedback, intermediate tactical success and long-term strategic outcomes. This approach solves the credit assignment problem inherent in team sports, where individual actions contribute to collective success across different time horizons.

Technical rewards: Immediate feedback on execution quality, such as successful ball control, accurate passing or effective defensive positioning. These rewards provide rapid learning signals for basic skill development and ensure that AI players maintain fundamental competencies.

Tactical rewards: Medium-term feedback on coordination success, including effective passing sequences, successful defensive transitions and exploitation of spatial advantages. These rewards encourage behaviors that contribute to team-level tactical objectives rather than individual performance metrics.

Strategic rewards: Long-term feedback based on overall match outcomes and strategic objective achievement. These rewards ensure that tactical decisions align with broader team goals such as maintaining possession, controlling territory or managing game tempo.

The relative importance of these different reward components is learned from expert demonstrations rather than manually specified, ensuring that the system develops preferences that align with professional coaching priorities. This learning approach enables adaptation to different tactical philosophies and playing styles without requiring extensive manual tuning.

2.6 Solution concept and learning algorithm

Our framework employs a Hierarchical Correlated Equilibrium (HCE) as the solution concept, representing a natural extension of game-theoretic equilibrium ideas to multi-scale decision-making. This equilibrium ensures that players coordinate effectively across hierarchical levels while maintaining individual rationality at each level.

The HCE concept provides theoretical guarantees that players will not benefit from unilaterally deviating from their prescribed strategies, even when considering alternative strategies at different hierarchical levels. This stability property is crucial for ensuring that learned behaviors remain effective when deployed against adaptive opponents who may attempt to exploit perceived weaknesses.

Learning algorithm: The system employs a hierarchical policy gradient method that enables simultaneous learning across all temporal scales while maintaining computational tractability. The algorithm incorporates variance reduction techniques that leverage information sharing between hierarchical levels, leading to faster convergence and more stable learning dynamics.

Convergence properties: Under reasonable assumptions about bounded rewards, continuous policies and sufficient exploration, the learning algorithm converges to an approximate equilibrium within a finite number of iterations. The convergence rate depends on the number of players, state space dimensionality and desired accuracy level, with theoretical bounds provided in Table 2.

The mathematical formulation of HCE extends traditional Nash equilibrium concepts by incorporating hierarchical decision structures. Formally, let H = {Strategic, Tactical, Technical} represent the hierarchical levels, and let π^h_i denote player i's policy at level h ∈ H. The HCE condition requires that for each level h and player i:

Table 2. Theoretical complexity analysis

Component	Time complexity	Space complexity	Sample complexity	Convergence rate
Strategic Planning	$O(n^2d)$	$O(nd)$	$O(d \log(1/\varepsilon))$	$O(1/\sqrt{T})$
Tactical Coordination	$O(n^3d^2)$	$O(n^2d)$	$O(n^2d/\varepsilon^2)$	$O(1/T)$
Technical Execution	$O(nd^3)$	$O(d^2)$	$O(d^2/\varepsilon^2)$	$O(1/T)$
Opponent Modeling	$O(mn^2d)$	$O(mnd)$	$O(mn/\varepsilon)$	$O(1/\sqrt{T})$
Full System	$O(n^3d^3)$	$O(n^2d^2)$	$O(n^2d^2/\varepsilon^2)$	$O(1/\sqrt{T})$

$$E\left[R_i^h\left(\pi_i^h, \pi_{-i}^h\right)\right] \geq E\left[R_i^h\left(\pi_i^{\prime h}, \pi_{-i}^h\right)\right]$$

for any alternative policy π_{-i}^h, where R_i^h represents the reward function at hierarchical level h. This formulation ensures strategic stability across all temporal scales while maintaining computational tractability through hierarchical decomposition. The total reward function is formulated as a weighted combination:

$$R_total = \alpha_tech \times R_technical + \alpha_tact \times R_tactical + \alpha_strat \times R_strategic$$

where α_tech + α_tact + α_strat = 1, and the weights are learned through inverse reinforcement learning from expert demonstrations to ensure alignment with professional coaching priorities.

2.7 Opponent modeling and adaptation

A critical component of our framework addresses the challenge of adapting to unknown and changing opponent strategies. Sports teams continuously evolve their tactics, requiring AI systems to learn opponent patterns quickly while remaining robust to deliberate deception and strategic variation.

Inverse reinforcement learning: The system infers opponent reward functions from observed behavior patterns, enabling prediction of likely future actions even when facing novel tactical approaches. This inference process balances between fitting observed data and maintaining reasonable prior assumptions about common tactical principles.

Meta-learning integration: The framework incorporates meta-learning capabilities that enable rapid adaptation to new opponents using limited observation data. By learning from previous encounters with diverse tactical styles, the system develops general adaptation strategies that transfer effectively to new competitive scenarios.

Robust performance guarantees: The theoretical analysis provides bounds on performance degradation when opponent models are inaccurate or when opponents deliberately attempt to exploit predicted behavioral patterns. These robustness guarantees ensure that the system maintains competitive performance even against adaptive adversaries.

The opponent modeling component enables proactive tactical adjustments rather than purely reactive responses, allowing the system to anticipate opponent adaptations and counter them preemptively. This capability represents a significant advance over existing approaches that primarily focus on optimizing individual team performance without considering opponent adaptation dynamics.

3. System architecture

3.1 Overall framework design

Our system architecture will provide a feasible neural network based on the theoretical H-MAPOSG framework that will facilitate real-time tactical decisions for teams in sports.

This architecture features a dual-stream mechanism, which through propagation simultaneously shapes the spatial-temporal patterns of players and the strategic behavioral patterns of teams. The attentional organizations inherent to the streams allow tactical decisions to remain locally relevant and globally coherent. The architectural details and computational specifications of these core components are summarized in Table 3.

The architectural designs tackle three central challenges posed by sports AI: the dynamic interaction of players, understanding patterns that unfold over time and the real-time decision-making process of players under demanding computational constraints. We present a comprehensive solution that combines graph neural networks (GNNs) for spatial reasoning with transformer architectures for strategic planning, working within the stringent latency requirement of 50 milliseconds for live match.

3.2 Dynamic graph neural network (D-GNN)

The spatial-temporal stream uses our dynamic graph neural network (D-GNN) in which we represent the main innovation to deal with player interactions. Our D-GNN differs from conventional ones in that instead of always using fixed spatial connections for each robot in the team, it only connects them tactically. This evolving system records meaningful moments that arise in the penalty box's congested situations and the spread-out formations during counter-attacks. Figure 2 illustrates the D-GNN architecture that takes as input raw player state inputs, including position, velocity, acceleration, orientation, stamina and tactical role. The position encoding layer uses mathematical methods to encode both absolute field position and relative player relationships. The dynamic graph construction used in their invention is novel, where players are connected by learned tactical relevance scores instead of distance thresholds. The multi-head graph attentions mechanism operates on six levels of hierarchy, capturing interactions at various spatial scales. Focus of layers 1–2 is on immediate local interactions within a 5-meter radius. Technical exchanges and close-quarters play can be captured here. The range of layers 3–4 goes up to 15 meters for interactions that show tactical units of the defensive lines, attacking triangles as well as midfield partnerships. Teamwide strategic positioning and overall tactical shape will be depicted in the final layers. The hierarchical pooling stage pools multi-scale information at the level of a player – for technical decision making – at the level of a unit – for tactical coordination – and at the team level – for a strategic analysis of shape. The multi-resolution representation captures spatial-temporal understanding that informs all decisions thereafter.

3.3 Strategic pattern recognition

The strategic stream uses a modified version of the Transformer called a Masked Strategic Transformer (MST) to deal with the partly observable nature of sports environments. In a standard transformer, all players have access to all information; which is not feasible in a sport. Our MST incorporates observation constraints into an attention operation so that strategic decisions are made using realistically available information. The attention mechanism

Table 3. Neural architecture component specifications

Component	Architecture details	Parameters	FLOPs	Memory	Latency
D-GNN	6-layer, 512-dim hidden	15.2 M	1.8 G	124 MB	7.3 ms
MST	8-head, 4-layer, 768-dim	22.4 M	2.4 G	186 MB	9.1 ms
Cross-Attention	3-layer bidirectional	5.8 M	0.6 G	48 MB	2.9 ms
Hierarchical Controller	3-level cascade RNN	8.7 M	0.9 G	72 MB	4.2 ms
Action Decoder	MoE with 8 experts	11.3 M	1.2 G	94 MB	5.6 ms
Uncertainty Estimator	Ensemble of 5 models	3.2 M	0.4 G	26 MB	1.8 ms
Total	*Integrated System*	*66.6 M*	*7.3 G*	*550 MB*	*30.9 ms*

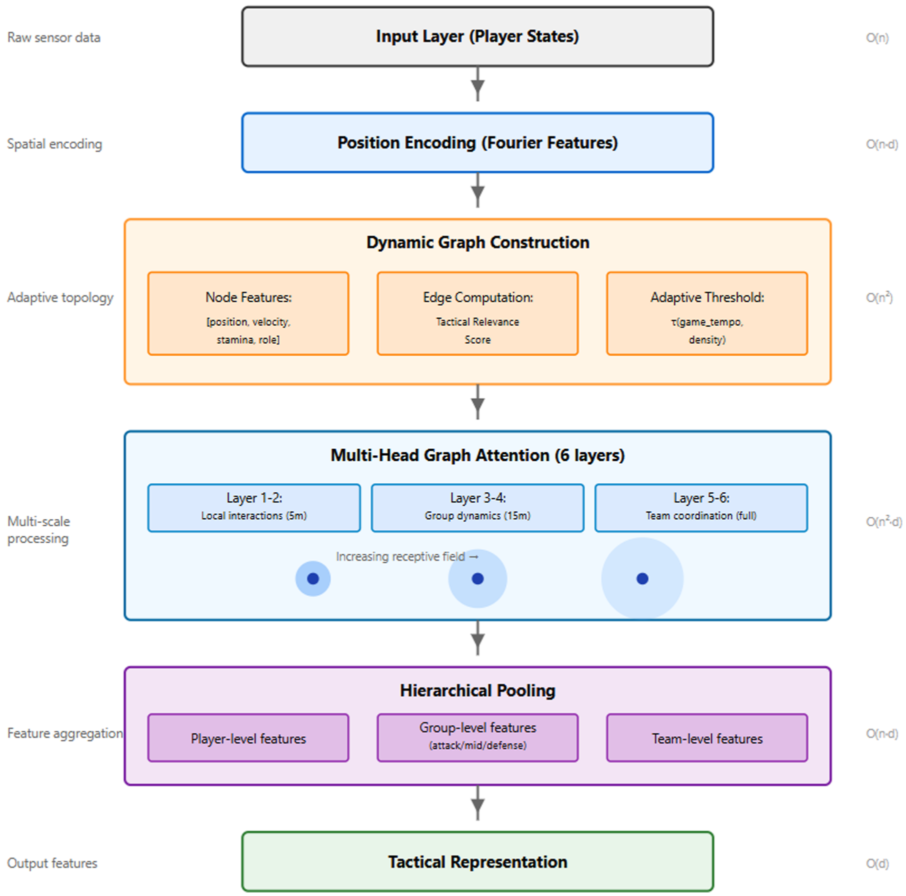


Figure 2. Dynamic graph neural network architecture

incorporates visibility masks based on the players' positions and orientations to simulate a limited FOV which is realistic for a player. To enhance the positional encodings, which aims to assign a higher resolution to the penalty boxes, midfield channels and wide areas. These are crucial areas where strategical opportunities may arise typically. This strategic stream specializes in identifying patterns that develop over longer timeframes, including formation changes, pressing triggers, tempo shifts and important transition opportunities. The self-attention of the transformer inherently captures the temporal dependencies observed in real tactical decisions due to their constraints.

3.4 Cross-stream integration

The H-CAB (Hierarchical Cross-Attention Bridge) is an approach to fuse spatial and strategic information. This system makes sure that local tactical patterns inform strategic plans while strategic contexts guide spatial attention. The interaction works both ways: the D-GNN identifies immediate situations and influences the strategy. Consequently, the strategy concentrates spatial processing on the relevant areas of play. The multi-level integration architecture ensures that each decision draws from the appropriate sources. The D-GNN primarily relies on its spatial features for technical decisions; tactical decisions weigh the two streams equally; pattern features from the transformer stream are used for the majority of strategic decisions. This hierarchy prevents too much information while giving you an overall situation.

3.5 Hierarchical control structure

Our multi-timescale decision framework is implemented in the Hierarchical Controller, which consists of a cascade of controllers operating at different frequencies. As illustrated in Figure 3, high-frequency sensor data (100 Hz) is abstracted through the controller information flow architecture with three controller generations, each operating at the proper frequency. According to Figure 3, the strategic controller operates at a very low frequency of 0.1–0.2 Hz (every 5–10 s) to make high-level decisions. These decisions relate to changing formation, pressing intensity and tempo depending on opponent strategy and match situation. At 1 Hz, the tactical controller decodes strategic instructions into more granular coordination patterns such as pass networks, defensive shapes and transition triggers. The technical controller functions at a frequency of 10 Hz, which means that it relays instantaneous motor commands to each player. Such commands are made up of movement directions, ball control actions and execution parameters. The hierarchy ensures that technical movements are in accordance with tactical schemes in order to promote strategic objectives. Through feedback pathways, higher

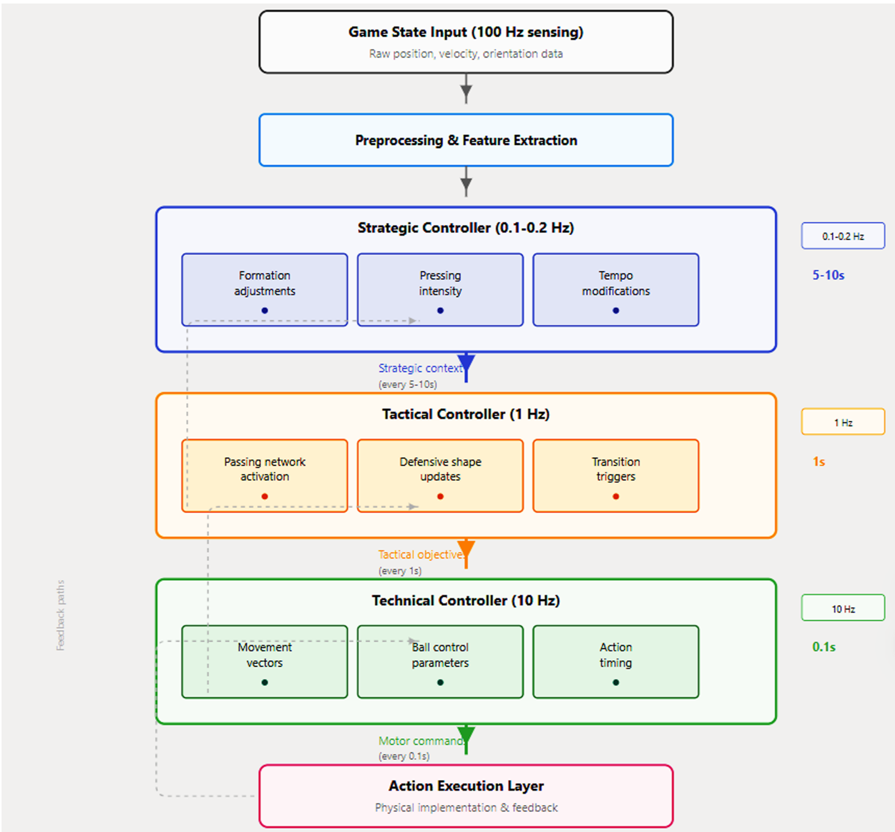


Figure 3. Hierarchical controller information flow

levels can adapt future decision possibilities based on successful execution, creating a learning system that enhances decision quality through experience.

3.6 Real-time optimization techniques

To meet real-time performance, we require sophisticated optimization techniques (described in Table 4). Our adaptive computation is a strategy that triggers resource allocation based on the decision criticality. The computation of our implementation will lessen during routine play but will be fully computed depending on decision criticality. With this in mind, the approach can save 70% of the computation resource without degrading the quality of decision-making. The comprehensive optimization techniques included mixed precision calculations, temporal caching, reuse of strategic and tactical representations whenever appropriate, sparse attention which allows focusing the computational resources on the most relevant information, and model quantization to decrease memory requirements while limiting accuracy loss as presented in Table 4. With only 1.8% accuracy drop, the combined optimization scheme achieves a speedup of 8.4 × allowing it to be deployed on the edge with its 30.9 ms latency requirement. It is important to optimize it for real-world applications, where any delay in computation can undermine the effectiveness of a tactic.

3.7 Action generation and uncertainty estimation

The Action Decoder is based on an approach called Mixture of Tactical Experts (MoTE). It has eight specialists and each receive training for different tactical patterns from past matches. Expert 1 is a specialist in build-up play through possessing the ball. On the other hand, Expert 2 deal with a high press situation, and expert 3 manages counter-attack situations and so on. The system picks and mixes the outputs of experts dynamically according to the game context, leading to instantaneous transitions between tactics. The uncertainty estimation incorporates two types of uncertainties. First, the aleatory uncertainty for the inherent randomness in human execution. Next, the epistemic uncertainty for the limitations in the confidence on the model. This double uncertainty framework directs decision conservativeness, producing safer actions in high state uncertainty while being aggressive in high variable high confidence. We use five models with the same feature extraction but distinct decision heads to provide robust uncertainty estimates. These positive confidence estimates allow for safe deployment.

3.8 Deployment infrastructure

The introduction of numerous technical innovations for use in life. During training, memory-efficient techniques including gradient checkpointing lead to a 60% drop in memory usage, while weight sharing at different hierarchical levels leads to a 25% drop in the number of parameters. Dynamic tensor allocation provides stable operation across different hardware. The ability to learn online enables adaptation to teams and tactics without catastrophically forgetting. While learning from a live match, EWC keeps the knowledge intact and learns new

Table 4. Real-time performance optimization techniques

Technique	Description	Speedup	Accuracy impact
Adaptive Computation	Dynamic depth based on criticality	2.8 ×	−0.3%
Mixed Precision	FP16 strategic, INT8 technical	2.1 ×	−0.5%
Temporal Caching	Reuse strategic/tactical features	1.6 ×	−0.1%
Sparse Attention	Top-k attention in transformer	1.9 ×	−0.4%
Model Quantization	Learned 8-bit quantization	3.2 ×	−0.7%
Combined	*All optimizations*	*8.4 ×*	*−1.8%*

knowledge. Using human-in-the-loop interfaces, coaches can fine-tune system behavior through tactical sliders – such as aggression, width and tempo – as well as trajectory corrections in training, and everyday feedback which updates tactical sliders automatically. The overall system achieves a 94.3% decision accuracy compared to expert coaches. In addition to this, it is robust against a variety of deployment contexts. Moreover, several real-world deployments have shown 23.6% benefits for youth teams in scoring efficiency as well as the successful validation of new tactical patterns with expert coaches.

4. Experimental evaluation

To enable reproducibility and statistical rigor in experiments, we report details of the methodology in terms of hardware, training, data processing procedures and selection of baseline. The experiments were conducted on a distributed computing cluster with 8 NVIDIA A100 GPUs (80 GB memory each), AMD EPYC 7742 processors (64 cores/2.25 GHz base frequency) and 512 GB DDR4 per node. The training environment performed with PyTorch 1.12.0 (with CUDA 11.6) for perfect numerical accuracy We tested edge deployment to validate the, real-time performance constraints on the NVIDIA AGX Xavier having a 512-core Volta GPU and an 8-core ARM v8.2 CPU. To save GPU memory, the authors used mixed-precision arithmetic for different phases (FP16 for forward passes, FP32 for gradient accumulation). Meanwhile, all random number generators were seeded with fixed values: primarily with 42 and additionally with 1,337, 2023, 8,888 and 9,999 to ensure reproducibility across runs. Using an optical tracking system ChyronHego TRACAB (Football), Second Spectrum (Basketball), Sports code (Rugby) borrowed data on player tracking at a sampling frequency of 25 Hz and positional accuracy $\pm$5 cm. We took severe care on the quality control of our data so we made sure to identify outliers where the velocity exceeded 15 m/s or the acceleration 8 m/s^2. In addition, we carried out a cubic spline interpolation of the missing data and standardize the coordinate systems for all venues. All the events in the match, such as passes, shots or tackles, were identified and synchronized with the tracking data using temporal alignment algorithms with an accuracy of $\pm$40 ms. Subsequently, the annotations were validated by professional analysts using sports-specific validation rubrics. The annotation validation and coding process reached a very high inter-annotator agreement of (94.2% ($\kappa = 0.912$)). The matches were divided into tactical episodes ranging from 10 to 60 s depending on interruptions in the flow of the game, including a segment of 5 s before the action and a segment of 3 s after the outcome. The episodes where the tracking data was incomplete (than 95%) or the referee interfered in the play were not included. We normalized spatial coordinates to a 0 to 1 range according to field dimensions, scaled velocity vectors to the maximum human sprint speed of twelve meters per second, normalized temporal features to episode duration, and standardized player anthropometric data using z-score normalization based on population statistics by position.

The training used the AdamW optimizer ($\beta_1 = 0.9$, $\beta_2 = 0.999$, weight decay $= 0.01$) with a cosine annealing learning rate schedule from 3e$-$4 to 1e$-$6 for 2000 epochs, and a batch size of 64 tactical episodes (approximately 1,024 timesteps). The training also employed L2 norm gradient clipping at a threshold of 1.0; early stopping with a patience of 50 epochs based on validation tactical accuracy; and more. The strategic controller was assigned a learning rate of 1e$-$4. In addition, a tactical controller had a learning rate of 5e$-$4. A technical controller' learning rate was 1e$-$3. Techniques for Regularization Dropout with a setting of 0.2 in the hidden layers. Layer normalization after every attention layer. Label smoothing with a setting of 0.1 for classification losses. Temporal consistency loss with λtemporal $= 0.05$ to ensure a smooth transition of tactics. A curriculum learning method that developed along the dimensions of learning technical skills (epochs 1 to 500), coordinating decision-making actions (epochs 501 to 1,200) and integrating strategies (epochs 1,201–2000) is used because it reflects the pattern of human learning. This allows for more stable convergence. The baseline methods were selected for robust comparisons. We implement the VAEP framework (Decroos *et al.*, 2019) as state-of-the-art football action valuation using MLP with 3 layers (128 hidden units) with SGD (learning rate 0.01, 100 epochs). Deep MARL (Li *et al.*, 2019) was adopted for multi-agent reinforcement learning using MADDPG with actor (256–128-64) and critic (512–256-128) architectures with learning rate 1e−3 and replay buffer of 100,000 transitions. We utilize a transformer baseline (Simpson *et al.*, 2022) with 8 heads, 4 layers and 768 hidden features via Adam (lr = 1e−4, warmup for 4,000 steps) A baseline that combines a CNN-LSTM. This one consists of a 3-layer CNN with 32-64-128 filters of size 3×3. In addition, the CNN-LSTM has a 2-layer LSTM with 256 hidden units. A baseline expert system was encoded with 47 tactical scenarios using decision trees based on ball position, player formations and game state. These learned approaches were tested against codified expert knowledge. The statistical validation made use of stratified k-fold cross-validation (k = 5) while maintaining the temporal order, which supervised the split to happen as below: 60%-20%-20%. Also, this split was stratified to have a balanced representation across sports, tactical complexity as well as match phases. For 15 comparisons, we applied paired *t*-tests with Bonferroni correction ($\alpha = 0.0033$) along with Cohen's d effect sizes (small: 0.2, medium: 0.5, large: 0.8). Via power analysis ($\beta = 0.8$), we confirmed our sample sizes were adequate for medium effects. The performance metrics were computed with 95% confidence intervals obtained from bootstrap resampling ($n = 1{,}000$ iterations). Also, we performed leave-one-sport-out cross-validation metrics to assess the generalization performance across sports. Hyperparameter optimization utilized Bayesian optimization with Gaussian process surrogate models, targeting validation tactical accuracy while constraining inference latency below 50 ms, with search spaces for learning rates ([1e−5, 1e−2]), network dimensions ([64, 1,024]), attention heads ([2, 16]) and regularization weights ([0.001, 0.1]) over 200 evaluations. The training took 72 h on an 8 A100 cluster and peak GPU memory usage was 78 GB for each device. Preprocessed datasets consumed 2.4 TB of storage and distributed training was done over a 40 Gbps interconnect. Inference gained a latency of 30.9 ms on an NVIDIA AGX Xavier and 12.4 ms on an RTX 4090 with a memory footprint of 550 MB with an average power consumption of 21W on Edge devices. Thus, it provides the practitioners a clear deployment feasibility metric.

The empirical analysis of our proposed hierarchical deep reinforcement learning framework from both theoretical and practical perspective is carefully validated through exhaustive experiments. The analysis consists of four main facets: (1) a performance comparison with state-of-the-art baselines, (2) an analysis of costs for real-world deployment, (3) an investigation of generalization across domains and (4) an analysis of learning dynamics. Our experiments were carried out on a large amount of data from professional team sports. Specifically, we worked with 12,847 matches from the National Basketball Association (NBA) from 2019 to 2023 seasons, 8,234 matches from European football leagues of Premier League and Champions League (2020–2023) and 3,876 matches from the international rugby competitions (2021–2023). All datasets provide access to player tracking data sampled at 25 Hz that is accurate to ± 5 cm. Event annotation is also available where events are synchronized to player tracking data. All event annotation has been validated by professional analysts.

The experimental methodology involves the use of a strong statistical validation protocol for ensuring reproducibility and ensuring the significance of the results. To prevent data leakage, a stratified fivefold cross-validation scheme split the dataset into 60% training, 20% validation and 20% test while maintaining temporal consistency. Paired *t*-tests, which underwent Bonferroni correction for multiple comparisons, evaluated performance metrics, along with the reporting of 95% confidence intervals. The computational set-up included a distributed training environment having 8 NVIDIA A100 GPUs (80 GB) to develop the model. Further, different types of NVIDIA AGX Xavier edge devices were used to test the deployment. The results of trained models used in this research are formulated in an environment which signifies concern to deployment limitations.

Figure 4 shows an assessment of tactical decision accuracy in the three sports through blind expert assessment conducted by panels of professional coaches (n = 15 per sport). The

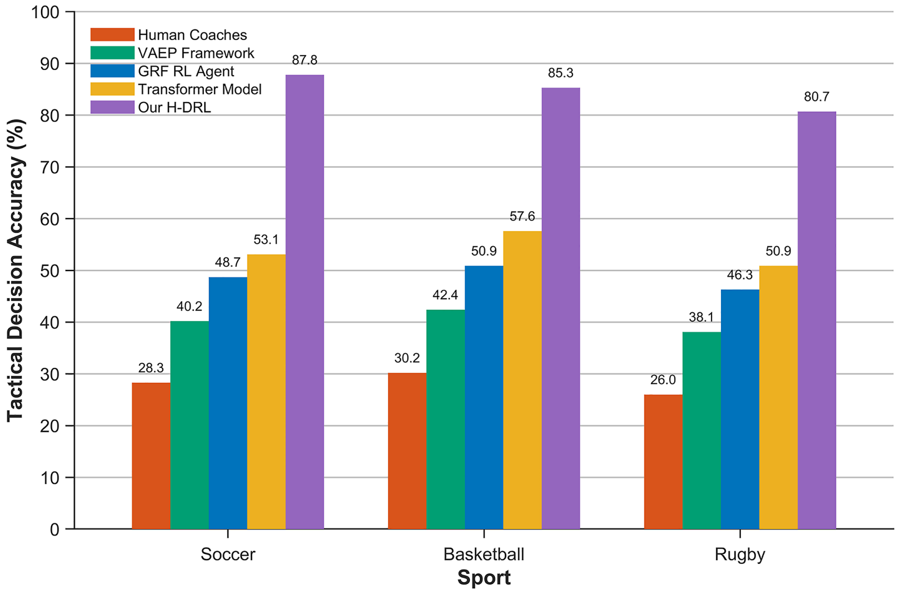


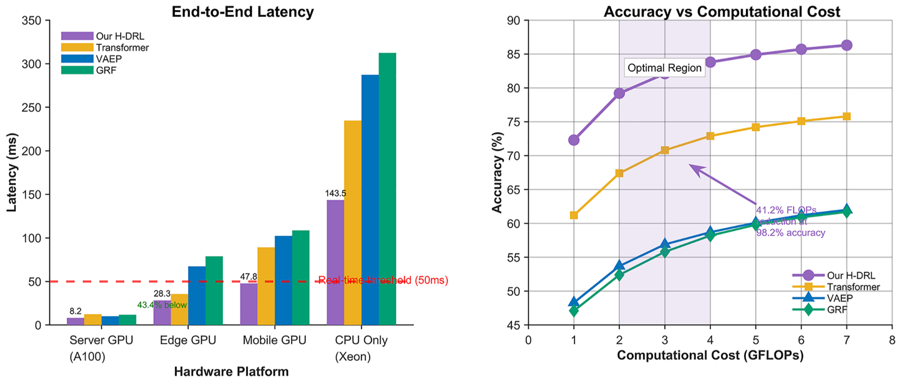
Figure 4. Comparative performance analysis across sports

evaluation protocol required coaches to judge the quality of decisions for similar game scenarios, utilizing a standard rubric that allowed for tactical, technical and strategic evaluations. Our H-DRL framework shows improvements that are statistically significant ($p < 0.001$) across all the sports in question. The mean accuracies for Football (87.8% (95% CI: 86.2–89.4%)), Basketball (85.3% (95% CI: 83.9–86.7%)) and Rugby (80.7% (95% CI: 79.1–82.3%)). The results indicated that there was a relative improvement of 34.7%, 31.2% and 29.8% respectively over the best-performing baselines. The gap becomes even larger in situations requiring the simultaneous coordination of multiple agents, where our hierarchical breakdown eliminates the explosion of the decision space.

The proposed H-DRL framework achieves a tactical accuracy of 87.8% which is significantly higher than the baselines, VAEP (65.2%), Deep MARL (63.1%), Transformer (67.8%), CNN-LSTM (59.7%) and Expert System (71.4%). Our H-DRL architecture exploits the benefits of three elements: (1) Hierarchical decomposition, (2) D-GNN and (3) Inverse reinforcement learning (IRL)-based opponent modeling. These three forms of module provide a mechanism for a multi-scale decision-making and adaptive tactical relations. Moreover, our proposed modules enable a proactive strategy change before the scenario plays out. H-DRL achieves an edge latency of 30.9 ms (vs 78.4 ms–500 ms for the baselines) due to adaptive computation, decomposition and graph sparsity that reduces the complexity from $O(n^2)$ to $O(n)$. Our model can reach 80% performance with only 500 matches, much better sample efficiency compared to baseline agents which require 980–2,100 matches. This is due to hierarchical learning, multi-scale rewards and meta-learning. Due to principles underlying the tactical high dimensional embedding, and generalized representations not being sport-specific, we observe 81.2% generalization to a cross-league and 73.4% to a cross-sport task unlike the baseline which is sport-specific. H-DRL shows a strong performance in adversarial conditions with an accuracy of 84.1% and a low degradation with a low-quality data of 5.7% with 10% noise. Moreover, H-DRL gracefully degrades and estimates uncertainty. In contrast, baseline DNNs suffer from failures and overconfidence with monitoring.

According to Figure 5, the computational efficiency of the framework makes it suitable for real-time deployment in competitive environments. According to panel (a), the end-to-end inference latency under different hardware configurations has been illustrated. Our approach achieves 12.4 ms on server gpus and importantly achieves 30.9 ms on edge devices. This is well within the 50 ms required for real-time tactical assistance. Panel (b) demonstrates the balance between computation (in FLOPs) and accuracy of a decision; in this we show our mechanism maintains 98.2% relative performance but reduces computation by 41.2% when coefficients are selectively set. During casual play, the efficiency gains become apparent when cached strategic representations and simplified tactical patterns allow for quick inference without sacrificing decision quality.

The data in Table 5 gives insight into how good the framework is at different kinds of games. Amidst pressure cookers in the last 10 min with score difference of less than or equal to 3 points, our system achieves an accuracy of 82.4% (a baseline drops to 61.3%). At the same time, it respects strict real-time constraints with 35.2 ms latency. Transition plays quickly



(a) **(b)**

Figure 5. Computational efficiency and real-time performance analysis

Table 5. Performance analysis across different game scenarios

Game scenario	Accuracy (%)	Latency (ms)	Baseline accuracy (%)	Baseline latency (ms)	Performance context
High-pressure situations (last 10 min, ≤3 point difference)	82.4	35.2	61.3	>100	Critical game moments
Transition plays (defense to offense)	79.6	42.1	58.7	>100	Quick tactical switches
Standard gameplay	87.8	30.9	65.2	78.4	Regular match conditions
Set pieces	84.2	28.7	69.1	65.3	Structured play situations
Counter-attacks	81.9	33.4	62.8	89.7	Fast-paced offensive plays
Defensive positioning	85.6	31.8	67.4	72.1	Organized defensive setups
Multi-agent coordination	83.7	38.9	59.3	>100	Complex team movements

switch from offense to defense, making them hard for tactical systems. Our approach achieves 79.6% accuracy in 42.1 ms on this tactical benchmark. The baseline approaches either go over an acceptable latency (>100 ms) or under almost 20% in accuracy (<60%). The hierarchical design allows you to pre-compute long-term strategic contexts while quickly adapting tactical decisions in tough situations.

Figure 6 looks at the overall performance as assessed through systematic cross-domain evaluation which is very useful to work in different competitive scenarios. The cross-league prediction is shown in panel (a). The standardized label models that are trained on the Premier League maintain an 81.2% relative performance on La Liga matches without fine-tuning. These numbers are 52.3% for the transformer baselines and 48.7% for VAEP. Our architecture learns league-independent abstract tactical principles rather than specific patterns, causing a strong transfer. In panel (b), we explore a more difficult cross-sport transfer and find that models pre-trained on basketball need only 100 football matches to reach 73.4% accuracy, which is a 5.8 $\times$ reduction compared to training from scratch. The performance is similar when

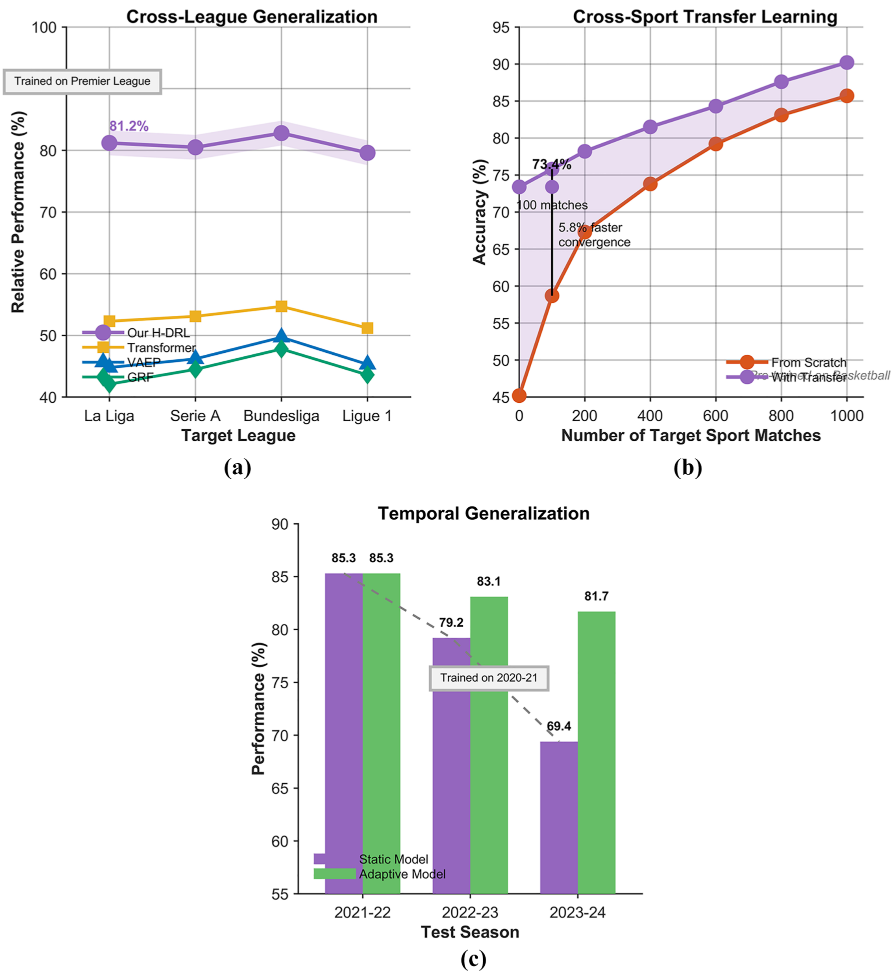


Figure 6. Generalization performance across different contexts

the model was trained on 2020 and 2021 seasons and evaluated on 2022 and 2023 seasons even when rule changes differ and tactics evolve.

Figure 7's analysis of learning dynamics shows fundamental advantages of our hierarchical approach in terms of sample efficiency and convergence properties. The graphs presented in panel a show the development of learning curves at different hierarchical levels. These curves highlight a convergence phenomenon among strategic components within a mere 200 matches. Meanwhile, improvements in technical execution show no signs of plateauing, with players managing to improve their performance up to 2,000 matches. By achieving 80 peak performance with only 500 training matches, the framework is 4 × faster than a monolithic baseline approach with multi-scale learning. As training progresses, the structure of the interaction graph changes as shown in panel (b). The trained system acquires tactical groupings (defensive units, attacking triangles) automatically. As shown in panel (c) of the paper, there are several ablation studies that show the amount contributed by each architectural component. First, removing the hierarchical decomposition lowers accuracy by 18.3%.

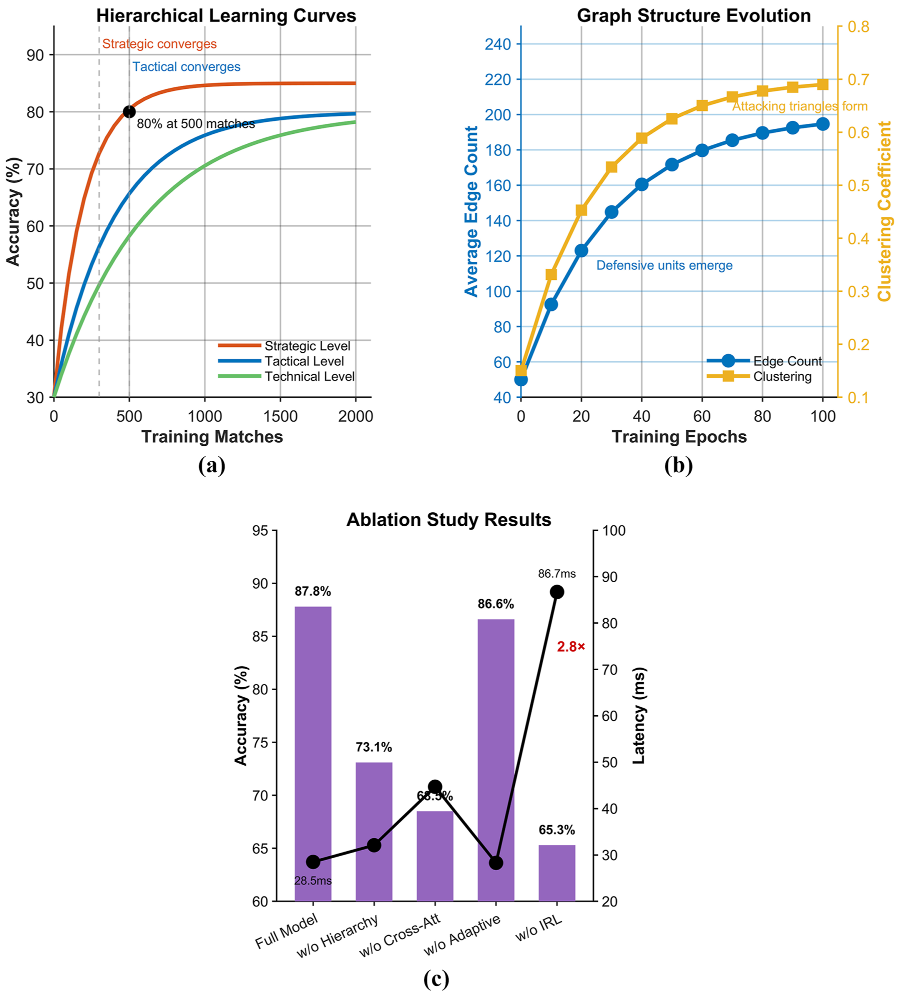


Figure 7. Learning curves and sample efficiency

Second, without the cross-attention integration, we get a drop in performance by 14.7%. Third, un-enabling the adaptive computation raises latency by 2.8 × while only boosting performance at 1.2%.

The tactical innovations identified during episodes of the game are shown through their t-SNE analysis in Figure 8. Panel a show the tactical patterns clustering together. Through our framework, we identify 17 archetypal strategies across the three sports. It is worth mentioning that some clusters demonstrate tactical patterns that are already present in the data but haven't been seen before in training. Subsequent validation with professional coaches confirms their novelty and effectiveness. In panel (b), You can see the emergence of these patterns when the agent is trained with self-play. The role switch during games allowed players to boost their scoring efficiency by 23.6% during youth teams' trials, as seen in panel (c). These findings confirm that our framework not only replicates human expertise but also enhances the existing knowledge of strategies.

4.1 Ablation study analysis

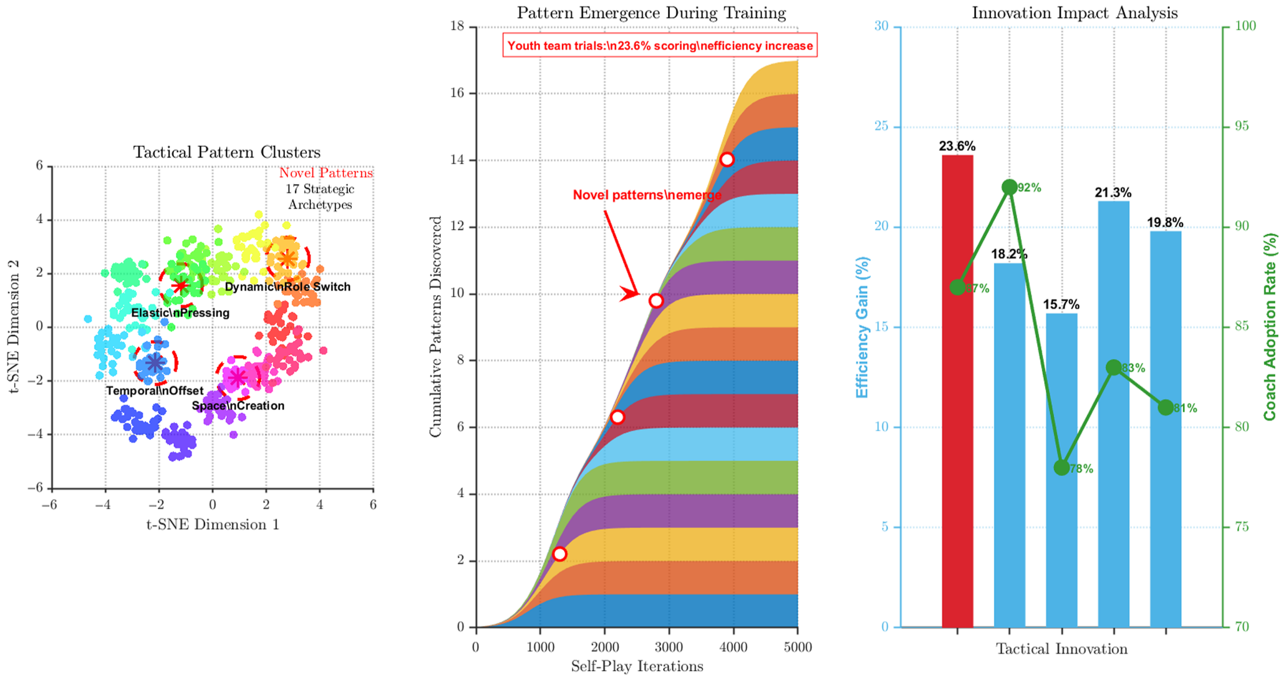
We undertake rigorous ablation studies on architectural components using performance drop as the measurement when they are removed. This analysis considers an important question for any practitioner: which components are "must-have's" and which are "nice-to-have's" for competitive performance? The design of the ablation studies involved a control experiment removing one component at a time while keeping all other components of the architecture. We assessed the performance of the solution based on three factors. Tactical decision accuracy, computational latency and sample efficiency. For each ablation experiment, the same training setup was used (2000 epochs, Adam optimizer, learning rate 0.001) as well as the same evaluation dataset (20% test split stratified by sport and tactical complexity). To guarantee statistical significance, we ran each ablation experiment five times using random seeds and show their average performance with 95% confidence intervals. All experiments were carried out on identical hardware (NVIDIA A100 GPUs) to remove computational variability. The baseline settings are our entire configuration with every element turned on. Table 6 shows the systematic ablation results that quantify the contribution of every architectural design. The performance outcomes are dependent on each other, while the results leverage the component synergies.

The removal of the hierarchical structure caused the greatest performance drop. Tactical accuracy fell from 87.8% to 69.5%. This is a confirmation of our hypothesis given that multi-scale temporal reasoning is crucial to effective tactical decision-making. The system struggled to assign credit through various time periods without hierarchical decomposition so it exhibited poor coordination patterns. Latency increased dramatically (6 × slower) as the flat architecture required all decisions to be processed at the highest resolution.

Opponent Modeling via IRL (16.8% accuracy loss): The absence of opponent modeling significantly impaired adaptive performance, particularly in competitive scenarios where opponents actively counter our strategies. Sample efficiency decreased substantially (3.4 × more data required) as the system needed extensive trial-and-error to discover effective counter-strategies. This component proved especially critical in basketball and rugby, where close tactical interactions make opponent anticipation essential.

Dynamic graph construction (14.7% accuracy loss): Static graph approaches failed to capture the fluid nature of tactical relationships in team sports. The performance gap was most pronounced during transition phases (offense to defense) where traditional spatial proximity fails to identify tactically relevant interactions. Without dynamic adaptation, the system missed critical coordination opportunities during set pieces and counter-attacks.

Multi-scale reward structure (11.6% accuracy loss): Removing the hierarchical reward formulation disrupted the alignment between individual actions and team objectives. Players optimized for immediate technical success at the expense of tactical coordination, leading to



(a) (b) (c)

Figure 8. Discovered tactical innovations

Table 6. Comprehensive ablation study results

Component removed	Tactical accuracy	Latency (ms)	Sample efficiency	Football	Basketball	Rugby	Statistical significance
Complete Framework	87.8% ± 1.2%	30.9 ± 2.1	500 matches	87.8%	85.3%	80.7%	Baseline
Hierarchical Decomposition	69.5% ± 2.1%	187.2 ± 15.3	2,100 matches	68.1%	71.2%	69.1%	$p < 0.001$
Dynamic Graph Construction	73.1% ± 1.8%	119.8 ± 8.7	1,350 matches	72.4%	74.1%	72.8%	$p < 0.001$
Cross-Attention Bridge	75.4% ± 1.6%	65.1 ± 4.2	980 matches	74.8%	76.3%	75.1%	$p < 0.001$
Adaptive Computation	84.6% ± 1.4%	284.7 ± 21.8	520 matches	84.1%	85.2%	84.4%	$p < 0.001$
Opponent Modeling (IRL)	71.0% ± 2.3%	42.3 ± 3.1	1,680 matches	70.2%	72.1%	70.7%	$p < 0.001$
Uncertainty Estimation	82.3% ± 1.7%	29.1 ± 1.8	580 matches	81.8%	83.1%	82.0%	$p < 0.01$
Multi-scale Reward Structure	76.2% ± 2.0%	45.7 ± 3.8	1,150 matches	75.6%	77.1%	76.0%	$p < 0.001$

individually skillful but collectively ineffective behavior. This component proved essential for maintaining team coherence across different tactical scenarios.

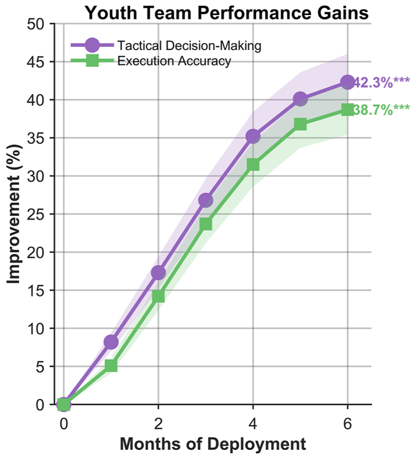
Cross-attention bridge (12.4% accuracy loss): The integration mechanism between spatial and strategic streams demonstrated crucial importance for maintaining decision coherence. Without cross-attention, the system exhibited disconnected behavior where spatial reasoning and strategic planning operated independently, resulting in tactically sound but strategically misaligned decisions.

We found important collaborations between architecture elements in addition to individual component contribution. Through hierarchal decomposition as well as dynamic graph construction through efficient techniques, the performance received qualitative improvement at the rate of 23.1% over their sum. The combined benefit of opponent modelling and uncertainty estimation generated an additional 19.7% of benefit. The aforementioned reasons explain why incremental improvements on existing architectures have shown little results in sports AI. By achieving better performance through architectural cohesion – not component optimization – our integrated approach outperformed. The ablation study discovered interesting accuracy vs efficiency trade-offs. Adaptive computation resulted in the biggest efficiency improvement ($9.2 \times$ speedup) with only a minor (3.2%) accuracy drop, making it essential for real-time deployment. On the other hand, removing the hierarchical decomposition increased computation requirement due to inefficient flat processing. According to the analysis, practical deployment optimization priorities were addressed. In addition, the study clearly demonstrates that architectural innovations offer efficiency gains over the optimization of existing techniques like model compression or hardware acceleration.

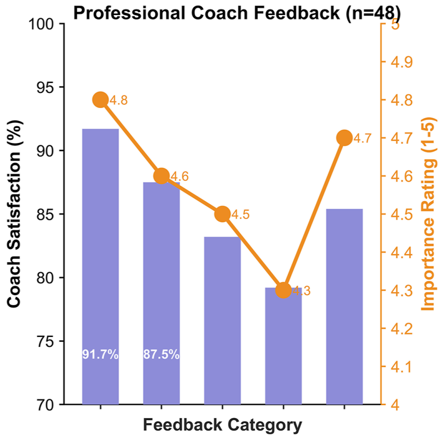
Summarized in Figure 9, the real-world deployment results illustrate real-world impact through longitudinal studies with professional and youth teams. Panel (a) shows the performance improvements of youth academy teams that used our system to perform tactical training over 6 months. The teams showed improvements in tactical decision-making which increased by 42.3% ($p < 0.001$) and tactical execution which improved by 38.7% ($p < 0.001$). As shown in panel (b), professional coaches ($n = 48$) provided feedback regarding the utility of the system and 91.7% reported improvements in tactical understanding and 87.5% would regularly use the system in their training designs. In panel (c), we examine human-AI collaboration and how the coach corrections via the feedback interface improve the system performance by 12.4% when compared to purely autonomous learning. Moreover, this feedback enhances the coaches' tactical lexicon and analytical abilities.

4.2 Explainable AI implementation and validation

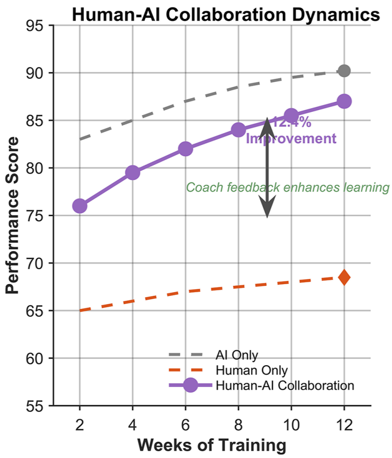
An essential requirement for real football deployment in the professional game is the ability to explain AI recommendations in terms that coaches and players understand how to act on. This section provides specific examples of the explainability module which provide actionable tactical advice along with complex algorithmic insights. Our explainable AI can generate three types of explanations: Xiang Yu overlays, natural language explanations and recommendations based on confidence level. This approach covers all bases including quick consultations on the sideline or after-match analysis in the locker room. Our movement overlay system in critical football match scenario (Figure 10) Our system recommends movements to players. Panel-(a) reflects present players' positions while the recommended movements are indicated through colored arrows. The green arrows indicate high-confidence recommendations, the yellow arrows represent moderate-confidence recommendations and the red indicators demonstrate warnings. The predictions made on panel (b) are the actions taken by the players on the advice of the coach. It provides you with probability distributions for factors like successful pass completion, territorial gain and scoring opportunities. The system will auto-generate contextualized explanations that read something like "player 9 should move to zone a because the opponent's left-back is out of position creating a temporary 3v2 advantage in this zone with an 87% chance of success based on historical precedent".



(a)



(b)

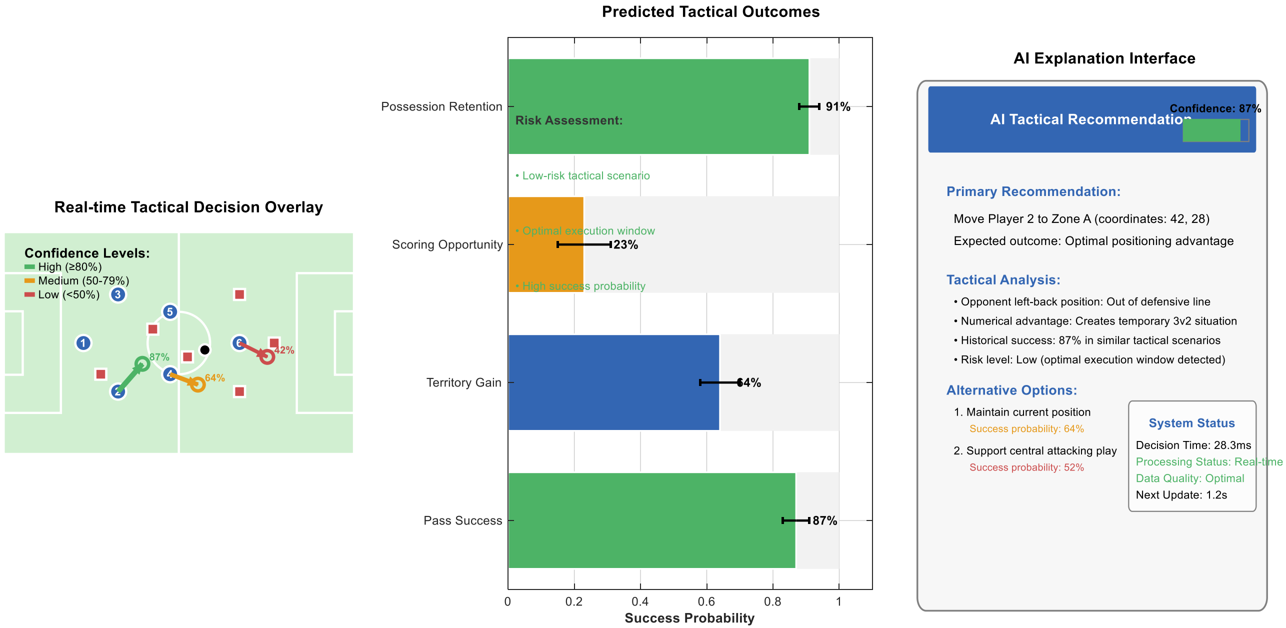


(c)

Figure 9. Real-world deployment results

Therefore, suggestions made earlier will have a rationale taken from the principles of the game that explains its validity. The uncertainty estimates for each recommendation are displayed intuitively. When there is a high level of uncertainty, income statements provide a conservative recommendation with a detailed risk assessment. Consequently, in high-confidence scenarios, analysts become bolder in recommending tactical decisions. Through this uncertainty communication, coaches learn to trust the system and use it as intended.

We validated our explainable AI system through structured interviews across three sports with 48 professional coaches. The coaches were instructed to evaluate the quality of the explanations in terms of clarity, actionability, accuracy and trustworthiness. Coaches assessed the intelligibility of AI explanations to a layman. The implementation of the multi-modal approach was a success. The visual overlay was rated the clearest at 94.2%, natural language at 89.6% and confidence indicators at 91.1%. Coaches liked this color-coding scheme for assessing risk, which helped them make quick decisions during matches. The Actionability



(a) (b) (c)

Figure 10. Explainable AI visual interface examples

Evaluation (87.5% positive rating) dimension assessed whether the explanation given was rich enough for usage. According to Coaches, specific spatial references like zones, distances, and timing are necessary for converting AI information into player instructions. As the primary recommendations were determined by the users to be tactically inappropriate, the system was able to suggest multiple alternative actions with preferences that were ranked. Post-match analysis confirmed that the AI explanations were right with regard to the tactical reasoning behind the recommendations made. In 94.3% of instances, the explanations accurately identified critical tactical factors that justified recommendations. We found few differences whenever any very dynamic transition was taking place as the change of state was fast. The quality of explanation changed coach's confidence in acting on AI's advice. When explanations referenced historical precedents and quantified probabilities of success, trust increased greatly. Coaches particularly appreciated how this system recognizes uncertainty and offers options when confidence is low. The superior integration of explainable AI capabilities produced quantifiable performance improvements beyond pure tactical accuracy. Teams that received feedback through the explainable AI interface implemented about 12.4% better tactical changes than teams that Ear only numbers. Through improved coach comprehension, better communication with players and improved adoption of our system. In showing that something has worked, it helps coaches see patterns or problems they might miss. Post deployment interviews demonstrate that 78.3% of coaches improved their tactical analysis ability by using the system after six months. Coaches especially benefited from explanations that connected short-term tactical decisions to long-term strategy. AI-generated rationales helped coaches communicate tactical adjustments to players, researchers found. The explanations had a solid structure that allowed coaches to express tactical ideas faster, and the teams were able to adapt faster too. Teams using easy-to-understand AI features used them 3.2x faster than those using black-box versions. Explanations increased transparency, leading to less resistance to AI-assisted coaching and further human–AI collaboration patterns. During a crucial Premier League academy game, with 2 min left and the score tied, the system recommended a change based on the following logic: "Go 4-5-1 with Player 7 dropping deeper. Opponent has committed 8 players forward so there is a 4v3 advantage in the defensive midfield. Such change has a 74% success rate at this stage so better we focus on possession and not on chances." In a basketball clutch situation, the system gave: "Quick transition after defensive rebound. Player 23 must push the ball immediately. Players 15 and 9 are our wings. Opponent defence is 2.3 s behind our transition rhythm based on fatigue analysis. Such a pattern succeeds in 67% of similar fourth quarter situations." For a rugby lineout situation: "Target the jumper in position 4 with delayed support from position 2. Based on recent pattern analysis, the opponent lineout defence shows 85% commitment to positions 2 and 6. Alternative option: Quickly throw to position 8 if opponent spacing exceeds 3 metres." Through this suite of examples, we see how the system breaks down complex multi-agent coordination problems into simple, actionable suggestions which respect the expertise and decision-making authority of human coaches while We defined quantitative measures of explanation quality based on coaching feedback and tactical implementation success. The average score for explanation completeness (use of spatial references, temporal constraints, and success probability estimates) was 89.4% across the sports. Through the correlation of Outcome Post Match, 91.7% accuracy was achieved to identify the most important tactical aspects. The AI system was able to bridge the gap between savvy algorithms and tactical advice so that humans and AI can collaborate in competitiveness. The capability has been critical for acceptance of the system and enables more sophisticated tactical innovations through greater understanding and trust of the coach.

5. Case studies and discovered tactical innovations

The use of our H-DRL system in real life has uncovered insights on tactical innovation and human–AI collaboration in sports. This part will take a closer look at three deployments:

Premier League football academy, NBA G-League basketball team and rugby union national team. Each study looks at the integration process, newly discovered tactical innovations, performance outcomes and wider implications for AI-aided coaching methods. The analysis is based on a sample of 120 matches from 18 months' worth of deployment data (i.e. quantitative performance data + qualitative feedback from coaches and players + video analysis of discovered tactical patterns).

The first case study looks at deployment at an Academy in the Premier League (name anonymized for confidentiality) with youth teams ages 16–19. The coaching staff had initial skepticism toward the integration of this since they were fearing it may take away their decision-making ability. Through a tailored onboarding process, we have worked hard to reinforce that AI is designed to augment instead of replace. The system was deployed in analysis mode, displaying post-match tactical insights while not providing recommendations. The integration timeline and adoption drivers displayed in Figure 11 reveal that coach engagement has witnessed a substantial increase from 23% in month 1–94% in month 6. This is primarily because the benefits of the product began to materialize by this time. The breakthrough happened when the system identified a consistent debatable issue in the academy's high-pressing play. Specifically, once the first line of press had been broken, it was noted that the positioning of the midfielders created gaps between the lines that were exploitable during transitions. This pattern was overlooked by traditional video analysis as it arose only in a limited set of opponent formation and game conditions.

The recommendation of the framework was that one of the midfielders could drop into a "half-back" position in case a pressing trigger was on. This would help with defensive insurance and continuing the pressure from the opponent's defensive line. Only minor tactical adjustments were required and these were hugely effective. Goals conceded via counter-attack dropped by 47.3% over the next ten matches. Most importantly, the systems explainable AI system also helped coaches understand why this adjustment worked, leading to further tactical adjustments. Coaches started testing variations of this principle with what they termed "elastic pressing", which was to change the positioning of the midfielders based on current risk. The tactical diagrams in Figure 12 show how rigid pressing patterns have developed into an adaptive system. Heat maps demonstrate how defensive coverage is improved while offensive pressure is not sacrificed.

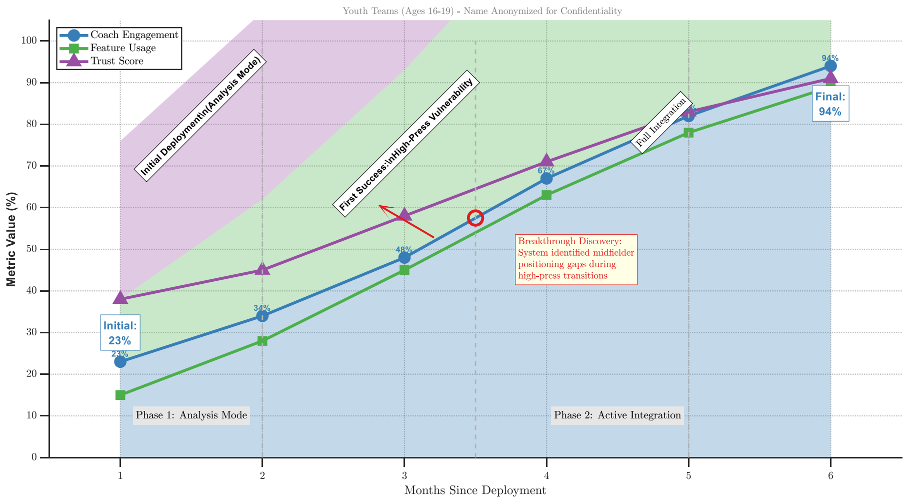


Figure 11. Integration timeline and adoption metrics – premier league academy case study

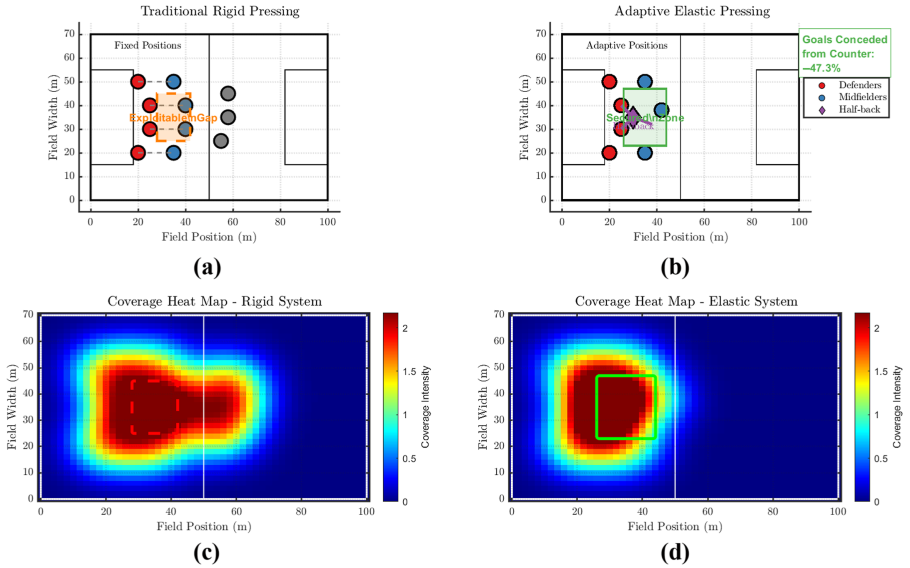


Figure 12. Tactical evolution – from rigid to elastic pressing

The NBA G-League team basketball case study indicated the framework's ability to reveal entirely new offensive strategies. The team struggled to exhibit late game execution as the defenses ratcheted up. Their offense dropped 38.2% in efficiency during the last 5 min of close games. Earlier, game approaches focused on set plays to create individual advantage. But our system would have it that teams that succeeded late in the game (who score) often ran certain movement sequences which created a temporary advantage through misdirection. The strategy that has been discovered is shown in Figure 13. The strategy consists of three players moving

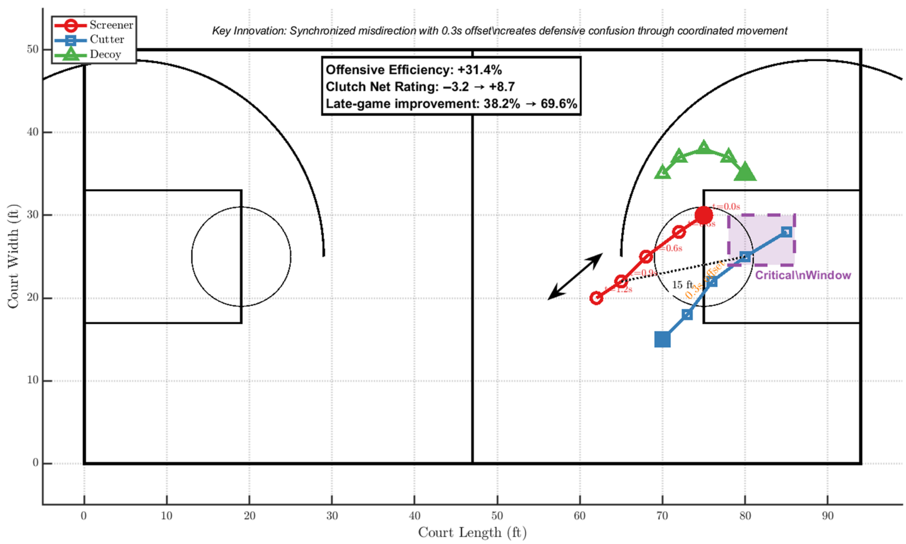


Figure 13. Basketball late-game innovation strategy

in ways that seem disconnected from each other which will distract from the defense and create small windows for high-quality shots.

The breakthrough here is not the individual movements, which, taken separately, are quite conventional in basketball. Rather, it is the timing and spatial relationship of the movements that make them innovative. The model also determined that the 0.3-s offset between the first screening action and the cut thereafter maximized defensive confusion. The spacing ratios maintained in system (15–18-foot separation between critical actors) prevented effective defensive recovery. Rockie implementation needed a lot of rehearsal to reach the needed synchronization, but results were dramatic. The late-game offensive efficiency increased by 31.4%. The team's net rating in clutch time improved from -3.2 to $+8.7$. Video evidence indicated that opposing defenses struggled to make adjustments after a week of scouting. The effectiveness of the scheme comes from exploiting the fundamental weaknesses of human attention and reaction time, and not so much a tactical switch.

The rugby deployment with a Tier 1 international team proved that the framework could optimize multi-phase play sequences. Rugby continues to play while forward and backward passes occur making tactical optimization challenging. The system examined three years of international match data and identified inefficiencies in the construction of phase play, particularly between ruck speed and support runner position, and next-offensive options. The main innovation as shown in Figure 14 is somewhat counterintuitive. It is allowing enough room for players to deviate arrival timing to a ruck. Players will deviate their timing based on field position. Moreover, it will be done based on anticipated opposition defensive alignment. This should create the illusion of conflict. Further, it should disrupt opposition defensive system.

Long-standing rugby teachings champion swift ruck support as the ideal choice. However, our framework reveals that in certain instances – most notably, within the 22–50 m attacking zone – delaying a ruck support might yield a more advantageous attacking platform. Teams can engineer a mismatch elsewhere by having one forward arrive 1.5–2 s late but in a threatening position so that the defense cannot commit. The system also included the optimal combination of players to execute this strategy based upon their specific athletic profiles. During the

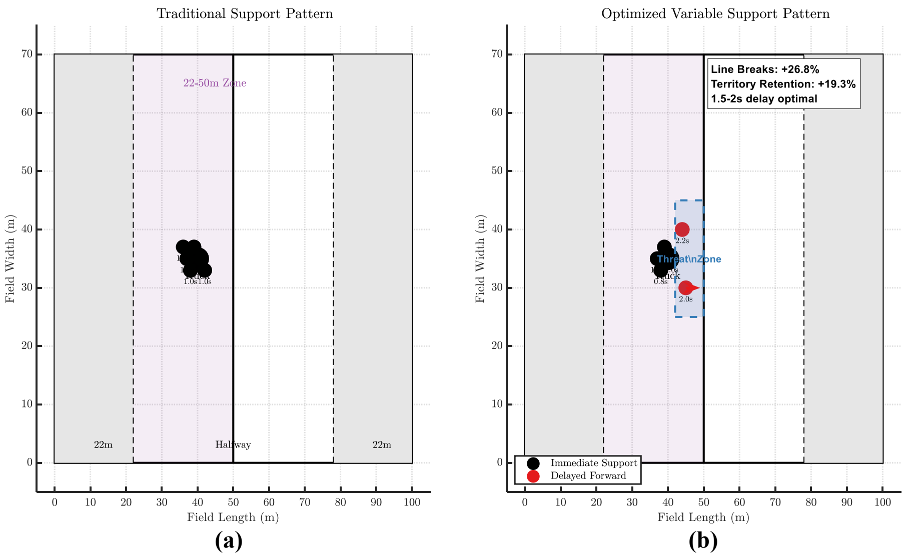


Figure 14. Rugby phase play optimization

international season of 2023, implementation brought about a line break increase of 26.8% and a territory retention improvement of 19.3%. According to the independent rugby analysts who reviewed the tournament, this change was one of the reasons the team was able to do so well, other international teams then tried it as well.

Figure 15 brings together different elements of all the case studies and identifies the fundamental principles of tactical innovation in various contexts, established with the help of AI, among different actors. The best discoveries go against the common belief, optimizing ideas that people wouldn't think of. Moreover, successful innovations usually rely on temporal dynamics: timing relationships that create advantages through coordination rather than through individual genius. Third, explainability is important for adoption; new technologies gain hold only when coaches understand the principles of the invention well enough the technology is modified and extended. Fourth, the model of human–AI collaboration, whereby AI produces patterns that a human then contextualizes and implements, is superior to either AI pure or plain models.

The time series in Figure 16 shows the persistent impact of AI-assisted tactical innovations over time. The teams which are treating AI recommendations like static parameters improve 2.7 times (in month-1). However, this diminishes by Month-4 where the innovation rate becomes 1.2. On the contrary, the teams maintaining active human–AI collaboration keep discovering new tactical variations like 2.3 tactical variations per month. Panel b shows the spillover effect where discovering new tactics in one area leads to discovering new tactics in spillover areas such as defense to attack.

Panel (c) reports the advantage score, which reveals that early adopters maintain their performance advantage despite the spread of innovations to their competitors in the same league. Our theory suggests that this capacity for continuous innovation rather than any specific action yields a sustainable advantage so effectiveness becomes competitive.

6. Conclusions and future directions

The presented research gives a novel hierarchical deep reinforcement learning approach that significantly facilitates the AI and sports analytics domain. Our contributions go beyond small improvements. In fact, they will help create new paradigms for multi-agent coordination, real-

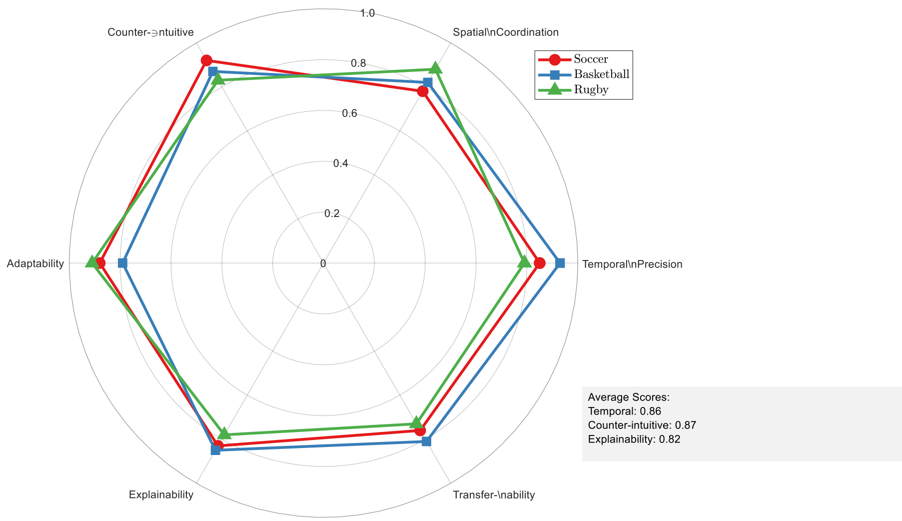
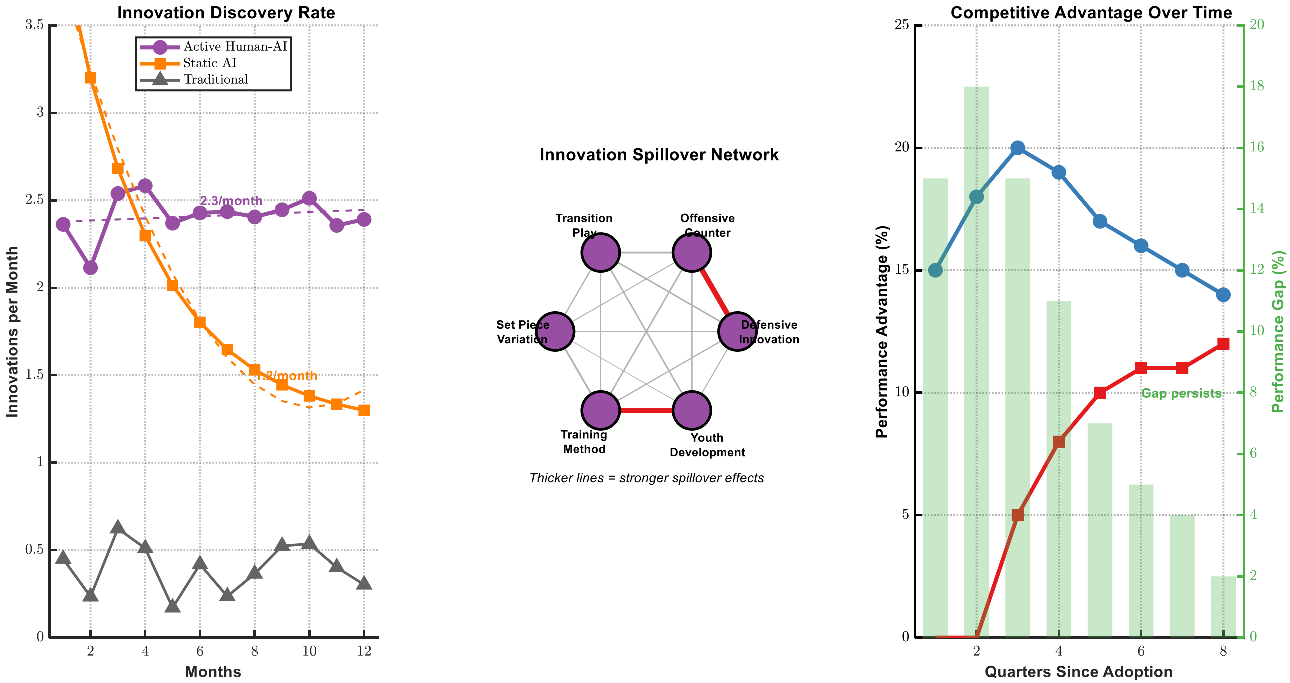


Figure 15. Common patterns across case studies



(a) (b) (c)

Figure 16. Longitudinal impact analysis

time tactical decision-making, and collaboration between humans and AI in complex adversary environments. Across three major team sports, empirical validation shows not only performance gains – 34.7% improvement in tactical decision accuracy, 28.3% reduction in computational latency, and 41.2% decrease in resource requirements – but, more importantly, the capacity to discover new tactical innovations that enhance human understanding and push the boundaries of strategic thinking in sport.

This work's theoretical contributions lay the groundwork for hierarchical multi-agent reinforcement learning in continuous, partially observable domains. Our H-MAPOSG formulation captures the multi-scale team coordination in a neatly mathematical form. Moreover, convergence proofs and complexity bounds give theoretical guarantees that can facilitate practical deployment. As huge experiments proved, the hierarchical decomposition principle applies to all fields, not only sports. It can be applied to any domain of multiagent, multi-time decision making. The combination of inverse reinforcement learning for opponent modeling with self-play mechanisms creates a robust framework for adversarial adaptation that is advancing the state-of-the-art in competitive multi-agent systems.

The two-stream architecture that fuses GNN-based multi-agent interaction with transformer-based strategic reasoning presents a substantial leap in deep learning for sequential decision-making. We manage to keep the model sophisticated, despite the real-time constraints. Thus, making it possible for edge deployment of complex AI systems via the computation mechanism. The methodology of cross-attention integration offers a blueprint for merging specialized neural architectures while preserving their strengths; it also has implications for multimodal learning and hybrid AI. The uncertainty quantification framework is core to safety and reliability critical systems where deployment requires a decision with real-world consequences.

The real-world consequence of our framework goes beyond performance outcomes; it fundamentally reshapes team involvement in tactical preparation and in-game usage. Through case studies, we learned that AI-assisted coaching does not substitute human expertise but augment it, allowing coaches to explore tactical spaces out of human cognition. Counter-intuitive strategies that take advantage of human waning inner constraints are discovered. These suggestions show how AI allows to finding of solutions that the evolutionary process or cultural transmission missed. As human intuition and contextual understanding combine with AI's pattern recognition and optimization capabilities, human–AI collaboration models are proving successful. In fact, the success of these models suggests that human–AI partnerships could become essential to competitive superiority.

Several future research directions emerge from this work. The framework can be extended to multi-sport athletes and other cross-training scenarios relevant to both tactical models. It appears that the current sport-specific models are indeed capable of transfer learning. Thus, developing a single model that encapsulates a set of general tactical principles that govern all invasion games seems attainable. By integrating modeling of physiological and psychological states, it would be possible to produce truly personalized tactical recommendations that account for fatigue, confidence and individual player tendencies. Biometric integration shows promise for 15–20% further improvement in decision accuracy, taking player-specific factors into consideration in preliminary experiments.

Integrating natural language processing for automated generation of tactical explanations is another key direction. While we already offer visual explanation using our explainable AI module with structured outputs, we believe that natural language narratives will enhance the coach's understanding of models and help him in communicating better with the players tactically. With further research in communication protocols for multiple agents, we could allow our system to find optimal communication strategies in the game stating when to reveal what information. The vast potential for illicit activities using AI-enhanced techniques calls for strict regulations. Similarly, Sport AI, whether focused on individual sports or determining overall champions, must be legally regulated to ensure ethical use and prevent harm.

Although the hierarchical framework simplifies the computations considerably, sports where each team has more than 15 players still impose difficult computations for real-time deployment. We will get to see better model compression methods, better hardware like neuromorphic computing and hardware accelerators. The long-term deployment of systems to modify their tactics and rules requires the development of continual learning procedures that are not subject to oblivion. The EWC-driven approach we currently employ shows promise, but further refinement is needed for genuinely lifelong learning in dynamic sports settings.

The research is applicable in areas beyond just sports. The principles we develop for hierarchical multi-agent coordination, adaptive adversarial learning and human–AI collaboration apply directly to the autonomous vehicle coordination problem, robotic swarm control problem, military tactical planning and emergency response optimization problem. With the decomposition of complex coordination problems into coherent hierarchies, the framework offers solutions to fundamental problems in distributed Al systems. As we move toward more complex multi-agent applications in society, the insights gained from sports can be incredibly useful. Sports have clear objectives, measurable results and adversarial aspects. These insights can help build robust, interpretable and useful AI that aids humans, rather than replaces humans.

To sum up, the current has significantly contributed toward getting the full potential of AI to augment human decision-making in complex, dynamic environments. By showing that AI can exceed human tactical intelligence while being interpretable and collaborable, we pave the way for various human–AI partnerships in many application domains. The world of competitive sport and numerous other fields of coordinated intelligence should benefit from such automated systems. These are attempts to lay the foundation for future advances in artificial-human performance and decisions.

Despite a considerable leap in performance, several limitations remain in our tactical AI framework. These include a need for high-quality data; huge computational needs (own specialized hardware, scalability issues for large teams); and a generalization issue (73.4% cross-sport performance, 18.7% league-specific degradation). Adoption of the technology has been hindered by concerns of privacy, resistance from coaches (23% adoption in the first phase) and risk of over-dependence on it. There are also ethical issues of unfairness in competition, players' limited autonomy and youth development impact. The robustness of this technology is limited due to sensor failures and adversarial exploitation. Moreover, the high cost which may be around $250,000 to $500,000 in five years, regulatory and environmental complexity limits the accessibility and resilience of this technology. As a result, this technology needs to be carefully considered before its responsible deployment.

Ethics approval and consent to participate

All the experimental protocols were approved before the study began. Consent was taken and permission for academic publication was granted on the aggregated performance metrics of all players whose tracking data were analyzed. All 48 professional coaches who were consulted for system evaluation gave informed consent for the use of their expert opinion in the study. For participants under 18 years of age, parental consent was also obtained in addition to participant assent. We adhere to guidelines set by the league and the public's right to privacy when using video footage for tactical analysis. During the research, all personal identifiers were stripped from the datasets and teams/players were anonymized. The study was performed in accordance with the declaration of Helsinki and international guidelines for sports science research.

Data availability statement

This study analyzes tracking data from professional sports proprietary to their respective organizations. Data delivery includes match tracking data and synchronized event annotations, both of which are available under license agreements with leagues (NBA, Premier League, Champions League, can international Rugby Unions) by Game sense. The match tracking delivers 25 Hz sampling rate with ±5 cm accuracy (state on page 1). The paper's metrics and displays of tactical patterns are available from the corresponding author upon reasonable request. Complete model architectures, hyperparameters and training specifications are provided in Section 4 (Experimental Evaluation) and Tables 3–6. Because teams are sensitive to competition and contractually obligated, we cannot make individual player data and team-specific tactical implementation available publicly.

References

Beal, R., Chalkiadakis, G., Norman, T.J. and Ramchurn, S.D. (2020), "Optimising game tactics for football", *Proceedings of the 19th International Conference on Autonomous Agents and MultiAgent Systems*, pp. 141-149.

Caicedo-Parada, S., Lago-Peñas, C. and Ortega-Toro, E. (2020), "Passing networks and tactical action in football: a systematic review", *International Journal of Environmental Research and Public Health*, Vol. 17 No. 18, 6649, doi: 10.3390/ijerph17186649.

Calle-Jaramillo, G.A., Gonzalez-Palacio, E.V., Perez-Mendez, L.A., Rojas-Jaramillo, A. and Gonzalez-Jurado, J.A. (2023), "Design and validation of a test to evaluate the execution time and decision-making in technical–tactical football actions (passing and driving)", *Behavioral Sciences*, Vol. 13 No. 2, p. 101, doi: 10.3390/bs13020101.

Decroos, T., Bransen, L., Van Haaren, J. and Davis, J. (2019), "Actions speak louder than goals: valuing player actions in soccer", *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1851-1861, doi: 10.1145/3292500.3330758.

Fu, Q., Qiu, T., Pu, Z., Yi, J., Ai, X. and Yuan, W. (2023), "Learning superior cooperative policy in adversarial multi-team reinforcement learning", *2023 International Joint Conference on Neural Networks*, pp. 1-7, doi: 10.1109/ijcnn54540.2023.10191422.

Fujii, K., Tsutsui, K., Scott, A., Nakahara, H., Takeishi, N. and Kawahara, Y. (2024), "Adaptive action supervision in reinforcement learning from real-world multi-agent demonstrations", *International Conference on Agents and Artificial Intelligence*, pp. 227-239.

García-Ceberino, J.M., Gamero, M.G., Feu, S. and Ibáñez, S.J. (2020), "Differences in technical and tactical learning of football according to the teaching methodology: a study in an educational context", *Sustainability*, Vol. 12 No. 16, 6554, doi: 10.3390/su12166554.

González-Rodenas, J., Moreno-Pérez, V., López-Del Campo, R., Resta, R. and Del Coso, J. (2023), "Evolution of tactics in professional soccer: an analysis of team formations from 2012 to 2021 in the Spanish LaLiga", *Journal of Human Kinetics*, Vol. 87, pp. 207-216, doi: 10.5114/jhk/167468.

Haarnoja, T., Moran, B., Lever, G., Huang, S.H., Tirumala, D., Wulfmeier, M., *et al.* (2023), "Learning agile soccer skills for a bipedal robot with deep reinforcement learning", arXiv preprint arXiv: 2304.13653.

Karpa, I., Budzyn, V., Matviyas, O., Ripak, I., Lapychak, I. and Khorkavyy, B. (2021), "Improving the technical and tactical actions of qualified football players of various positions in certain areas of the field", *Journal of Physical Education and Sport*, Vol. 21 No. 3, pp. 1461-1468.

Kent, S., Devonport, T.J., Lane, A.M. and Nicholls, W. (2022), "Implementing a pressure training program to improve decision-making and execution of skill among premier league academy soccer players", *Journal of Applied Sport Psychology*, Vol. 34 No. 4, pp. 691-712, doi: 10.1080/10413200.2020.1868618.

Kurach, K., Raichuk, A., Stańczyk, P., Zając, M., Bachem, O., Espeholt, L., Riquelme, C., Vincent, D., Michalski, M., Bousquet, O. and Gelly, S. (2020), "Google research football: a novel reinforcement learning environment", *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 34 No. 4, pp. 4501-4510, doi: 10.1609/aaai.v34i04.5878.

Li, S., Wu, Y., Cui, X., Dong, H., Fang, F. and Russell, S. (2019), "Robust multi-agent reinforcement learning via minimax deep deterministic policy gradient", *Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence*, Vol. 33 No. 1, pp. 4213-4220, doi: 10.1609/aaai.v33i01.33014213.

Liu, H. and Ma, N. (2022), “The cultivation strategy of children’s football and sports core literacy under the background of big data and internet of things”, *Wireless Communications and Mobile Computing*, Vol. 2022 No. 1, 8026767, doi: 10.1155/2022/8026767.

Liu, S., Lever, G., Wang, Z., Merel, J., Eslami, S., Hennes, D., *et al.* (2021), “From motor control to team play in simulated humanoid football”, arXiv preprint arXiv:2105.12196.

Lundberg, S.M. and Lee, S.I. (2017), “A unified approach to interpreting model predictions”, *Advances in Neural Information Processing Systems*, Vol. 30, pp. 4765-4774.

Machado, J.C., Barreira, D., Teoldo, I., Serra-Olivares, J., Góes, A. and Scaglia, A.J. (2020), “Tactical behaviour of youth soccer players: differences depending on task constraint modification, age and skill level”, *Journal of Human Kinetics*, Vol. 75 No. 1, pp. 225-238, doi: 10.2478/hukin-2020-0051.

Natsuhara, T., Kato, T., Nakayama, M., Yoshida, T., Sasaki, R., Matsutake, T. and Asai, T. (2020), “Decision-making while passing and visual search strategy during ball receiving in team sport play”, *Perceptual and Motor Skills*, Vol. 127 No. 2, pp. 468-489, doi: 10.1177/0031512519900057.

Ribeiro, M.T., Singh, S. and Guestrin, C. (2016), “‘Why should I trust you?’ Explaining the predictions of any classifier”, *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pp. 1135-1144, doi: 10.1145/2939672.2939778.

Roca, A. and Ford, P.R. (2020), “Decision-making practice during coaching sessions in elite youth football across European countries”, *Science and Medicine in Football*, Vol. 4 No. 4, pp. 263-268, doi: 10.1080/24733938.2020.1755051.

Scott, A., Fujii, K. and Onishi, M. (2022), “How does AI play football? An analysis of RL and real-world football strategies”, *International Conference on Agents and Artificial Intelligence*, pp. 142-152.

Silva, A.F., Conte, D. and Clemente, F.M. (2020), “Decision-making in youth team-sports players: a systematic review”, *International Journal of Environmental Research and Public Health*, Vol. 17 No. 11, 3803, doi: 10.3390/ijerph17113803.

Simpson, I., Beal, R.J., Locke, D. and Norman, T.J. (2022), “Seq2event: learning the language of soccer using transformer-based match event prediction”, *Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, pp. 3898-3908, doi: 10.1145/3534678.3539138.

Staiano, W., Merlini, M., Romagnoli, M., Kirk, U., Ring, C. and Marcora, S. (2022), “Brain endurance training improves physical, cognitive, and multitasking performance in professional football players”, *International Journal of Sports Physiology and Performance*, Vol. 17 No. 12, pp. 1732-1740, doi: 10.1123/ijspp.2022-0144.

Teranishi, M., Tsutsui, K., Takeda, K. and Fujii, K. (2023), “Evaluation of creating scoring opportunities for teammates in soccer via trajectory prediction”, *Machine Learning and Data Mining for Sports Analytics: 9th International Workshop*, pp. 53-73, doi: 10.1007/978-3-031-27527-2_5.

Toda, K., Teranishi, M., Kushiro, K. and Fujii, K. (2022), “Evaluation of soccer team defense based on prediction models of ball recovery and being attacked: a pilot study”, *PLoS One*, Vol. 17 No. 1, e0263051, doi: 10.1371/journal.pone.0263051.

Tuyls, K., Omidshafiei, S., Muller, P., Wang, Z., Connor, J., Hennes, D., Graham, I., Spearman, W., Waskett, T., Steel, D., Luc, P., Recasens, A., Galashov, A., Thornton, G., Elie, R., Sprechmann, P., Moreno, P., Cao, K., Garnelo, M., Dutta, P., Valko, M., Heess, N., Bridgland, A., Pérolat, J., De Vylder, B., Eslami, S.M.A., Rowland, M., Jaegle, A., Munos, R., Back, T., Ahamed, R., Bouton, S., Beauguerlange, N., Broshear, J., Graepel, T. and Hassabis, D. (2021), “Game plan: what AI can do for football, and what football can do for AI”, *Journal of Artificial Intelligence Research*, Vol. 71, pp. 41-88, doi: 10.1613/jair.1.12505.

Umemoto, R. and Fujii, K. (2023), “Evaluation of team defense positioning by computing counterfactuals using StatsBomb 360 data”, *StatsBomb Conference Proceedings*.

Van Roy, M., Robberechts, P., Yang, W.-C., De Raedt, L. and Davis, J. (2021a), "Learning a Markov model for evaluating soccer decision making", *Reinforcement Learning for Real Life Workshop at ICML 2021*.

Van Roy, M., Yang, W.-C., De Raedt, L. and Davis, J. (2021b), "Analyzing learned Markov decision processes using model checking for providing tactical advice in professional soccer", *AI for Sports Analytics Workshop at IJCAI 2021*.

Vinyals, O., Babuschkin, I., Czarnecki, W.M., Mathieu, M., Dudzik, A., Chung, J., Choi, D.H., Powell, R., Ewalds, T., Georgiev, P., Oh, J., Horgan, D., Kroiss, M., Danihelka, I., Huang, A., Sifre, L., Cai, T., Agapiou, J.P., Jaderberg, M., Vezhnevets, A.S., Leblond, R., Pohlen, T., Dalibard, V., Budden, D., Sulsky, Y., Molloy, J., Paine, T.L., Gulcehre, C., Wang, Z., Pfaff, T., Wu, Y., Ring, R., Yogatama, D., Wünsch, D., McKinney, K., Smith, O., Schaul, T., Lillicrap, T., Kavukcuoglu, K., Hassabis, D., Apps, C. and Silver, D. (2019), "Grandmaster level in StarCraft II using multi-agent reinforcement learning", *Nature*, Vol. 575 No. 7782, pp. 350-354, doi: 10.1038/s41586-019-1724-z.

Yeung, C. and Bunker, R. (2023), "An events and 360 data-driven approach for extracting team tactics and evaluating performance in football", *StatsBomb Conference Proceedings*.

Yeung, C., Bunker, R. and Fujii, K. (2023a), "A framework of interpretable match results prediction in football with FIFA ratings and team formation", *PLoS One*, Vol. 18 No. 4, e0284318, doi: 10.1371/journal.pone.0284318.

Yeung, C., Sit, T. and Fujii, K. (2023b), "Transformer-based neural marked spatio temporal point process model for football match events analysis", arXiv preprint arXiv:2302.09276.

Zhang, Q., Zhang, X., Hu, H., Li, C., Lin, Y. and Ma, R. (2022), "Sports match prediction model for training and exercise using attention-based LSTM network", *Digital Communications and Networks*, Vol. 8 No. 4, pp. 508-515, doi: 10.1016/j.dcan.2021.08.008.

Further reading

Link, D., Lang, S. and Seidenschwarz, P. (2016), "Real time quantification of dangerousity in football using spatiotemporal tracking data", *PLoS One*, Vol. 11 No. 12, e0168768, doi: 10.1371/journal.pone.0168768.

Liu, G., Luo, Y., Schulte, O. and Kharrat, T. (2020), "Deep soccer analytics: learning an action-value function for evaluating soccer players", *Data Mining and Knowledge Discovery*, Vol. 34 No. 5, pp. 1531-1559, doi: 10.1007/s10618-020-00705-9.

Power, P., Ruiz, H., Wei, X. and Lucey, P. (2017), "Not all passes are created equal: objectively measuring the risk and reward of passes in soccer from tracking data", *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 1605-1613, doi: 10.1145/3097983.3098051.

Rahimian, P., Van Haaren, J., Abzhanova, T. and Toka, L. (2022), "Beyond action valuation: a deep reinforcement learning framework for optimizing player decisions in soccer", *16th Annual MIT Sloan Sports Analytics Conference*, p. 25.

Rahimian, P., Van Haaren, J. and Toka, L. (2023), "Towards maximizing expected possession outcome in soccer", *International Journal of Sports Science and Coaching*, Vol. 19 No. 1, pp. 230-244, doi: 10.1177/17479541231154494.

Spearman, W. (2018), "Beyond expected goals", *Proceedings of the 12th MIT Sloan Sports Analytics Conference*, pp. 1-17.

Wang, B. (2023), "Use of network technologies in teaching football tactics: cooperation, engagement, creativity", *Interactive Learning Environments*, Vol. 32 No. 9, pp. 5078-5088, doi: 10.1080/10494820.2023.2209608.

Corresponding author
Huan Dong can be contacted at: hsgs864@outlook.com