HELIOS2026: Team Description Paper

Hidehisa Akiyama[1][0000−0002−9793−9123] and
Tomoharu Nakashima[2][0000−0002−1443−0816]

[1] Nagasaki University, Nagasaki, Japan
hidehisa.akiyama@gmail.com
[2] Osaka Metropolitan University, Osaka, Japan
tomoharu.nakashima@omu.ac.jp

Abstract. This paper presents an overview of the previous work and recent research efforts of Team HELIOS2026, a leading team in RoboCup competitions. Since its inception, Team HELIOS has participated in RoboCup competitions, achieving eight championships, including four consecutive victories since 2022. The team has consistently ranked among the top four teams since 2005. This paper not only highlights the team's historical achievements but also delves into recent advancements, focusing on formation analysis and Situation Score evaluation. Building on prior research, the team has developed innovative methods for unsupervised formation clustering using game logs from the RoboCup2D dataset, enabling the quantification of tactical characteristics. Additionally, the paper introduces a two-stage machine learning model to evaluate the Situation Score (SS), a metric to assess proximity to a goal. The model employs CatBoost for classification and a neural network for regression, achieving high accuracy and reliability in predicting critical game situations. These advancements underscore the team's commitment to integrating data-driven approaches with domain expertise to enhance performance in robotic soccer competitions.

Keywords: Formation analysis · Situation Score · Machine learning.

1 Introduction

Team HELIOS has been a leading participant in the RoboCup Soccer Simulation 2D League since 2000, securing eight world championships, including four consecutive victories since 2022 [2,3,4]. Our recent research emphasizes data-driven tactical analysis to handle the increasing complexity of opponent strategies.

This paper presents two key technical advancements integrated into HELIOS2026. First, we propose an unsupervised formation clustering method. By applying MiniBatch K-Means to relative player positions, we enable the quantitative identification of dynamic tactical structures in both offensive and defensive phases. Second, we introduce a refined Situation Score (SS) evaluation framework. This approach utilizes a two-stage machine learning model, incorporating CatBoost for classification and a neural network for regression, to accurately predict goal-scoring potential from complex field states. These methods demonstrate our commitment to integrating domain expertise with advanced machine learning to enhance multiagent performance.

2 Previous Works

We have publicly made some of our team's source code and debugging tools available to assist new teams in participating in the competitions [1]. Currently, the released source code is accessible on our project site[3]. We have also worked on the generation of datasets using simulated soccer. The development of tools for data generation is an ongoing process, and several datasets have already been published [5][4]. Additionally, we are organizing a competition utilizing these datasets[5].

3 Formation Analysis

Formations are crucial elements that represent a team's tactical characteristics. Identifying the opposing team's formation allows us to predict their tactics and devise effective counter-tactics. Building upon our previous research on formation identification [4], in this study we investigate a novel formation clustering method originally proposed for human soccer analysis. We clustered formations using game log data from the RoboCup2D dataset [5] and estimated the optimal number of clusters. Furthermore, leveraging this optimal clustering, we quantified formation transitions and the specific tactical characteristics of each opponent.

3.1 Formulation of the Formation Classification Problem

Formation classification involves categorizing a team's tactical structure into representative patterns based on player positioning during a match. In soccer, player positioning changes continuously; therefore, formations should not be considered static, but rather as dynamic spatial arrangements at any given time step.

In this paper, we define a formation as the positioning of the 10 field players (excluding the goalkeeper) at each time step during a match, utilizing their pitch coordinates. Consequently, the input data consists of the coordinate sets of the 10 players at each specific time step.

We formulate formation classification as the task of grouping these time-step positioning data into representative formations based on spatial similarity. Rather than relying on predefined ground-truth labels, we treat formation classification as an unsupervised clustering problem. This approach allows us to automatically extract a team's characteristic formation structures from the positional data within the match logs.

[3] `https://github.com/helios-base` (Please cite [1] when using the code from this site in your publications.)
[4] `https://github.com/hidehisaakiyama/RoboCup2D-data/`
[5] `https://sites.google.com/view/stp-challenge/`

3.2 Feature Representation and Clustering Method

Following Shaw et al. [7], we adopt a feature representation focusing on the relative positional structure of the formation. Focusing on the 10 field players (excluding the goalkeeper), the position of player $i\,(i = 1, 2, \ldots, 10)$ at time t is extracted as:

$$p_i(t) = (x_i(t), y_i(t))$$

To obtain a representation independent of absolute pitch positions, we first calculate the team's centroid (center of mass) at each time step:

$$c(t) = \frac{1}{10}\sum_{i=1}^{10} p_i(t)$$

Next, we calculate the relative position of each player with respect to $c(t)$:

$$r_i(t) = p_i(t) - c(t)$$

This enables the comparison of pure formation structures regardless of the team's overall location. The data is concatenated into a 20-dimensional feature vector $f(t)$ representing the relative coordinates of the 10 players at each time step t. Furthermore, to mitigate the impact of feature scale disparities on clustering, we standardize the data so that each dimension has a mean of 0 and a standard deviation of 1:

$$\tilde{f}_j(t) = \frac{f_j(t) - \mu_j}{\sigma_j}$$

We apply the MiniBatch K-Means algorithm to the standardized feature vectors $\tilde{f}(t)$, as it is well-suited for processing large-scale data efficiently.

We segment the match into offensive and defensive phases, performing clustering independently for each. These phases are classified by analyzing the temporal shifts in the centroid of the 10 field players. The longitudinal displacement (x-axis direction) of the players over a 10-frame window is calculated and normalized against the pitch size. We define an offensive phase when this normalized forward displacement exceeds 0.5, and a defensive phase when the backward displacement falls below -0.5.

3.3 Experiments

We conducted experiments using data from 50 matches between AETeam and HELIOS2024, sourced from the RoboCup2D dataset. Out of 322,775 total frames, 119,971 (37.2%) were classified as offensive and 123,986 (38.4%) as defensive phases. Frames meeting neither condition were deemed transitional periods lacking a clear tactical state and were thus excluded from clustering.

We performed formation clustering independently for both phases using MiniBatch K-Means with a batch size of 10,000 frames and a maximum of 100 iterations. We evaluated the number of clusters (k) ranging from 2 to 25, estimating

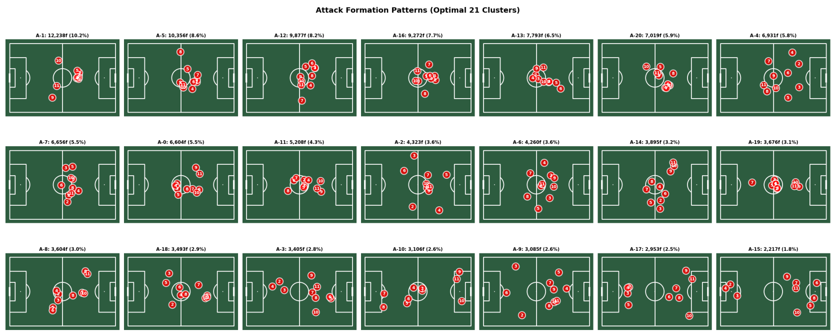

Fig. 1: Visualization results of representative formations in the offense phase based on the optimal number of clusters ($k = 21$).

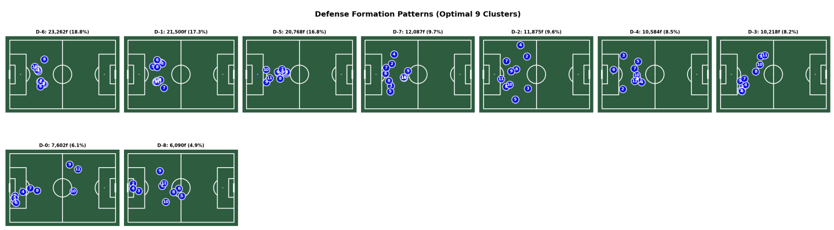

Fig. 2: Visualization results of representative formations in the defense phase based on the optimal number of clusters ($k = 9$).

the optimal k via the Elbow method and the Silhouette coefficient. While clustering was applied to all applicable frames, the Silhouette coefficient was calculated using a random sample of up to 50,000 frames for each k to reduce computational overhead.

The Elbow method indicated a distinct inflection point near $k = 21$ for the offensive phase. The corresponding Silhouette coefficient was 0.116. Conversely, for the defensive phase, the Elbow method identified $k = 9$ as the optimal number of clusters, yielding a Silhouette coefficient of 0.087. Figures 1 and 2 illustrate representative formations for the offensive and defensive phases based on these optimal cluster counts.

4 Qualitative Field Evaluation

4.1 Situation Score

Situation Score (SS) is an evaluation metric for soccer fields proposed by Nakashima [6]. It quantitatively evaluates the number of additional steps required from the current state to reach the goal. Higher SS's indicate that a goal is assumed to be soon to be achieved. Thus, from the attacker's perspective, the optimal decision is to increase the SS value, whereas the defense teams should decrease it. We define the SS value as follows:

$$SS(x) = SS_{\max} - T(x), \tag{1}$$

where $SS_{\max}$ is the pre-defined maximum SS value, $T(x)$ is the remaining steps from the current situation until the goal. In this paper, the remaining steps are measured in terms of simulation cycles in the soccer simulator. If the remaining simulation cycles until the goal exceed $SS_{\max}$, the SS value becomes negative.

4.2 SS Evaluation by Machine Learning

If the SS value for any field situation is correctly evaluated by a machine learning model, it can be used to develop team strategies or individual tactics. This paper considers a two-stage model for evaluating an SS value for a soccer field. The first stage of the model predicts whether the SS value of the current soccer field is positive, and the second stage predicts the exact SS value. Only those situations that are evaluated as positive are used in the second stage for predicting the exact SS values. Thus, the first stage performs classification, while the second stage performs regression (Fig. 3).

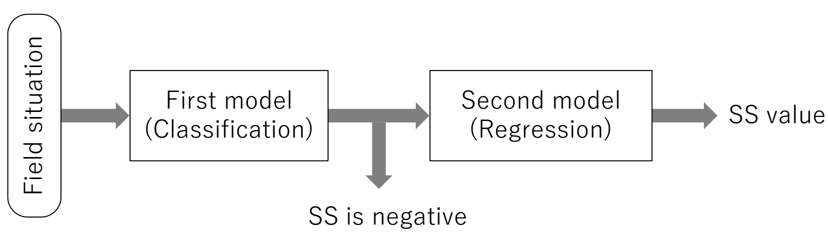

Fig. 3: Two-stage model for evaluating SS values.

The advantage of employing a two-stage model is twofold. First, by screening out the majority of cases where $SS < 0$ in the first stage, the regression model in the second stage can be specialized to evaluate critical cases in which goals or shots are likely to occur soon. Second, by separating classification and regression models, distinct feature designs and machine learning algorithms can be applied to each. This separation facilitates independent consideration of improvement strategies for each model. In this paper, we will focus on the first-stage classification task using real soccer data.

4.3 Numerical Experiments Settings

Numerical experiments were conducted to evaluate the prediction accuracy of the Situation Score using the two-stage model. The feature vector for each situation was constructed by concatenating the positions of the players and the ball at each simulation cycle, extracted from the log data obtained after the game. Situation Score data were generated for goal-scoring situations from 1,000 Helios2022 and HeliosBase [1] matches. This resulted in a dataset comprising 596,826 situation data instances with a positive Situation Score and 1,183,892 instances with a negative Situation Score. These data were used to construct the first-stage classification model and the second-stage regression model.

For the first-stage classification model, CatBoost [8] was employed. For the second-stage regression model, a three-layer neural network was adopted. The hyperparameters of these models were adjusted using a portion of the training data as a validation set. Furthermore, $SS_{\max}$ in (1) was set to 50, meaning that only situations within 50 cycles of the next goal were assigned a positive Situation Score.

4.4 Numerical Experiment Results and Discussion

The accuracy of the first-stage classification model, evaluated on the test data, reached 100%. When the second-stage regression model predicted the Situation Score for situations classified as positive by the first stage, the Mean Squared Error (MSE) was 27.23 and the coefficient of determination (R^2) was 0.87 for the training data. Furthermore, evaluating the model's performance on the test data yielded an MSE of 82.18 and an R^2 of 0.6. These results indicate that the proposed model can accurately predict the Situation Score for various game situations.

5 Conclusion

This paper summarized Team HELIOS2026's achievements and key research contributions in RoboCup. The team has won eight championships since its inception, consistently ranking among the top four teams since 2005, including four consecutive victories since 2022. Our main contributions include:

- Novel unsupervised formation clustering using game log data, enabling quantitative tactical analysis
- Development of a two-stage machine learning model for Situation Score evaluation

These advancements demonstrate our team's commitment to improving robotic soccer performance through data-driven methods.

References

1. Akiyama, H., Nakashima, T.: HELIOS Base: An Open Source Package for the RoboCup Soccer 2D Simulation, In: RoboCup 2013: Robot World Cup XVII. RoboCup 2013. Lecture Notes in Computer Science, vol 8371. Springer, Berlin, Heidelberg., pp.528–535 (2014)
2. Akiyama, H., Nakashima, T., Hatakeyama, K., Fujikawa, T.: HELIOS2022: RoboCup 2022 Soccer Simulation 2D Competition Champion, In: RoboCup 2022: Robot World Cup XXV. RoboCup 2022. Lecture Notes in Computer Science, vol 13561. Springer, Cham., pp. 253–263 (2023)
3. Akiyama, H., Nakashima, T., Hatakeyama, K., Fujikawa, T., Hishiki, A.: HELIOS2023: RoboCup 2023 Soccer Simulation 2D Competition Champion, In: RoboCup 2023: Robot World Cup XXVI. RoboCup 2023. Lecture Notes in Computer Science, vol 14140. Springer, Cham., pp. 386–394 (2024)
4. Akiyama, H., Nakashima, T., Fujikawa, T., Hishiki, A., Yano, S., Okayama, S.: HELIOS2024: RoboCup 2024 Soccer Simulation 2D Competition Champion, In: RoboCup 2024: Robot World Cup XXVII. RoboCup 2024. Lecture Notes in Computer Science, vol 15570. Springer, Cham., pp. 495–501 (2025)
5. Akiyama, H., Nakashima, T.: Soccer Gameplay Data Generation: Toward Integrating Computer Simulations and Human Sports Analysis, In: RoboCup 2024: Robot World Cup XXVII. RoboCup 2024. Lecture Notes in Computer Science, vol 15570. Springer, Cham., pp. 351–360 (2025)
6. Nakashima, T., and Pomas, T.: Evaluation of Situation in RoboCup: 2D Simulations using Soccer Field Images, In: Proc. of the RoboCup Symposium, 12 pages (2018)
7. Shaw, L., Glickman, M. E.: Dynamic analysis of team strategy in professional football, In: https://api.semanticscholar.org/CorpusID:211526838 (2019)
8. Dorogush, G., Erkin, V., Vinkov, I.: CatBoost: A Gradient Boosting Library Designed for Training High-Precision Machine Learning Models at Large Scale. In: Advances in Neural Information Processing Systems (2017).