GreenSea: Visual Soccer Analysis Using Broad Learning System

Bin Sheng, Ping Li, *Member, IEEE*, Yuhan Zhang, Lijuan Mao, and C. L. Philip Chen, *Fellow, IEEE*

***Abstract*—Modern soccer increasingly places trust in visual analysis and statistics rather than only relying on the human experience. However, soccer is an extraordinarily complex game that no widely accepted quantitative analysis methods exist. The statistics collection and visualization are time consuming which result in numerous adjustments. To tackle this issue, we developed GreenSea, a visual-based assessment system designed for soccer game analysis, tactics, and training. The system uses a broad learning system (BLS) to train the model in order to avoid the time-consuming issue that traditional deep learning may suffer. Users are able to apply multiple views of a soccer game, and visual summarization of essential statistics using advanced visualization and animation that are available. A marking system trained by BLS is designed to perform quantitative analysis. A novel recurrent discriminative BLS (RDBLS) is proposed to carry out long-term tracking. In our RDBLS, the structure is adjusted to have better performance on the binary classification problem of the discriminative model. Several experiments are carried out to verify that our proposed RDBLS model can outperform the standard BLS and other methods. Two studies were conducted to verify the effectiveness of our GreenSea. The first study was on how GreenSea assists a youth training coach to assess each trainee's performance for selecting most potential players. The second study was on how GreenSea was used to help the U20 Shanghai soccer team coaching staff analyze games and make tactics during the 13th National Games. Our studies have shown the usability of GreenSea and the values of our system to both amateur and expert users.**

***Index Terms*—Object tracking, recurrent discriminative broad learning system (RDBLS), soccer tactics, visual analytics.**

Manuscript received December 14, 2019; accepted April 15, 2020. Date of publication May 21, 2020; date of current version February 17, 2021. This work was supported in part by the National Key Research and Development Program of China under Grant 2018YFF0300903, in part by the National Natural Science Foundation of China under Grant 61872241 and Grant 61572316, and in part by the Science and Technology Commission of Shanghai Municipality under Grant 15490503200, Grant 18410750700, Grant 17411952600, and Grant 16DZ0501100. This article was recommended by Associate Editor M. Shin. *(Corresponding authors: Bin Sheng; Lijuan Mao.)*

Bin Sheng is with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China (e-mail: shengbin@sjtu.edu.cn).

Ping Li is with the Department of Computing, Hong Kong Polytechnic University, Hong Kong (e-mail: p.li@polyu.edu.hk).

Yuhan Zhang is with the SJTU-ParisTech Elite Institute of Technology, Shanghai Jiao Tong University, Shanghai 200240, China.

Lijuan Mao is with the Shanghai University of Sports, Shanghai 200438, China.

C. L. Philip Chen is with the School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China, also with the Navigation College, Dalian Maritime University, Dalian 116026, China, and also with the Faculty of Science and Technology, University of Macau, Macau, China (e-mail: philip.chen@ieee.org).

Color versions of one or more of the figures in this article are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TCYB.2020.2988792

I. Introduction

Soccer is a worldwide popular sport. The sports business industry and Internet technology bring enthusiasts easy access to high-quality videos of almost all professional soccer championships. Such video data, together with the games and players statistics these days, can be reliably collected and archived. The resulting large data, however, present both challenges and opportunities to sports professionals and businesses. To maximize the utilization of such data, advanced analytics systems are needed for soccer scouts and coaches to evaluate, choose, and train players. Researchers have thus been developing quantitative analysis to help assess the performance of players and teams. Many websites, companies, and media, such as Opta [1] and WhoScored [2], update the statistics periodically. Some of them utilize a marking system and then create a ranking list of players and teams. Despite the popularity and achievements of quantitative studies, its scientificity and effectiveness are still under suspicion.

Existing research usually has three difficulties for both professionals and amateurs. First, a highly convinced quantitative analysis model needs a bulk of data to train. Traditional deep structure and learning, such as [3]–[5], suffer from a time-consuming training process because of a large number of connecting parameters in filters and layers. Moreover, it encounters a complete retraining process if the structure is not sufficient to model the system. These drawbacks seriously discourage further applications of deep learning in this field. Second, a soccer game is a complicated entirety that is affected by multiple factors. Statistics, on the other hand, are sparse and lack relevance. It is unwise to describe a game using statistics, such as shoot or tackle only. However, most existing systems only use statistics themselves other than considering the entirety, which may result in a misunderstanding of the game. That is why many coaches still believe in their own observation rather than statistics. Third, numbers cannot give the human a direct and distinct expression. Nevertheless, regardless of whether researchers want to help coaches or attract fans, time is precious. It is important to show them the conclusion of a game at first glance. The conclusion is obtained by visual impression. Traditional charts usually can only handle low-dimensional data, and they cannot be used to evaluate a quality that coaches care about.

Thus, we designed GreenSea, a visual analysis system, to solve the above problems from three aspects. First, we use the broad learning system (BLS) [6]–[9]. BLS is constructed based on the characteristics of the flattened functional link networks, where the original inputs are transferred and placed as "mapped features" in feature nodes, and the mapped features are enhanced as enhancement nodes with randomly generated weight. The mapped features represent that BLS does not directly use the original data as the input layer, but first makes some transformations on the data, which is equivalent to feature extraction. The incremental learning algorithms can be applied for fast remodeling in broad expansion without a retraining process if the network deems them to be expanded. Second, how to make the numbers "talk?" How to make all the statistics tell us whether a player has good performance or not? We develop our marking system by referring to the existing marking websites and institutions and adjust it with the help of experts, scouts, and commentators. We use BLS to treat statistics as inputs and several performance points as outputs. Some synthetic abilities, such as offense and defensive abilities, are also available. Each of them can be scored by the related statistics. A commentator does not need to understand the game before narration. The performance is already evaluated by a score. Third, how to make the analysis intuitively? High scores are equal to good distribution on the figure. The one who has a better assessment figure performs better. Classifying statistics into groups with a different focus and organizing them well in coordinates, every ability can be known at first sight.

We also considered the real settings of different clients and designed several data-acquisition schemes. For example, for many professional clubs who need data of opponents in the same league mostly, we provide an image tracking system using TV video. For teenager training camps and schools, the videos captured by simple equipment, like digital video (DV) camera and mobile phone, can also be used. Our system has a practical application for youth training games from a low level to province level. The first study focused on ability assessment and development prediction. The second one places emphasis on analyzing the advantages and disadvantages of different teams and making tactics. We always maintain close contact with experienced coaches and players during the period of cooperation to improve our system. GreenSea employs a new way to link the quantitative and qualitative analyses. This article makes the following three main contributions.

1) We present a long-term recurrent discriminative tracking system, which is based on BLS. It has good effects on occlusion, and its detection part is trained by BLS aiming at finding and retracking lost targets. The broad learning strategy has a strong advantage in processing speed while ensuring accuracy, which also enables our system to quickly provide reliable analysis for coaches. Compared with other mainstream methods, such as deep belief networks (DBNs) [10], stacked autoencoders (SAEs) [11], and another version of stacked autoencoder (SDA) [12], our proposed recurrent discriminative BLS (RDBLS) needs much less running time and achieves state-of-the-art accuracy; it only needs 6.78 s to train with the accuracy of 98.26%, while SAE's performances are 3420.37 s and 97.45% and SDA's performances are 3489.73 s and 97.64%.
2) We introduce a marking system, applying traditional statistics and some newly defined skills as input, using BLS to train and update. Our system is able to use quantitative statistics to evaluate the performance of players, which can provide the users with a more intuitive method to assess players. Compared with the existing football marking systems that only focus on goals, our system tends to fully assess one player's abilities. As a result, GreenSea could better help coaches train their players in a more effective and more efficient way.
3) We design the visualization and interaction methods for visual analysis, performance assessment, and tactics making. Two studies on the use of our approach demonstrate the efficacy of GreenSea for player selection and tactics analysis of real games. In this article, we used the radar figure, assessment iceberg, and period-divided contour heat map to visualize players' performance during a match. Unlike the other radar figures, our radar figure shows one player's or one team's performance in the shoot, on-target and goal at the same time, which provides the coaches with an overall assessment. The iceberg map can reflect the potential ability of the player, while the heat map illustrates players' position and key area during one match. Consequently, the visualization could provide the coaches as much information as possible.

II. Related Work

In this section, we will introduce the background of related work. There are mainly four aspects: 1) object tracking; 2) quantitative analysis; 3) BLS; and 4) sports visualization.

A. *Multiobject Multiview Tracking*

The basic data that will be used to compute advanced statistics is players' positions during a 90-min game. There are some mature methods to achieve it, such as GPS wearable devices and infrared cameras. However, despite the precision problem, such equipment is not allowed in the majority of games. On the contrary, video tracking has many benefits. It has slighter demands on equipment. It captures more information for further research, and robust machine-learning methods can help recognize actions automatically. That is why our system adopts a video tracking method, which has the highest potential. Since soccer game tracking is a big-background multiobject and multiview tracking problem, we will emphatically study the existing multiobject multiview tracking methods. We have been focusing on authoritative tracking, such as [13]–[16] to find a solution, which has the best potential to be developed into multiobject multiview mode. We also researched some new multiobject tracking methods [17], [18]. Online multiobject tracking works via the following steps. Obtain detections from several detectors of each frame, then calculate trajectories by associating detections frame by frame. Previous works have different association methods, such as evaluating objects' affinities for appearances and positions [19]. There are some popular methods like using a velocity motion model to constrain the movements [20], solving the occlusion issue by online-trained classifiers [18], and associating by confidence map [21]. However, as soccer is a sport full of unknowns, the changes in the speed and trajectory of players have almost no rules to follow. Even though the algorithms mentioned are effective in short-term tracking missions, they will drift in football players tracking problems. In complicated situations, a nonlinear motion model was proposed [22]. It is an effective way to use a min-cost flow in a network flow to solve global association [23]. To merge track into final results, maximum weighted independent sets in a graph of detection pairs are applied in [24].

B. Quantitative Analysis

Compared to human judgment, statistics are more objective. Opta Sports [1] is a sports data provider. Opta records each game and will generate more than 200 statistics, including races, results, real-time scores, points ranking, team statistics, player statistics, lineup, team comparison, player rankings, and team rankings. Opta has a detailed record of position and time for each player on the field, and users can display a map of the location of a player's field, a route map, a passing roadmap, and an offensive analysis figure. Except for the detailed and accurate statistics, many websites, newspapers, and companies have also tried to use composite scores which are obtained by some synthesized algorithms using an individual action point. WhoScored [2] is a London-based website. After a game, WhoScored is able to give the final score for the first time and make minor adjustments within 10 min. In a normal situation, there are eight frequently used statistics. For different positions, the evaluation factors are different. On the stats for the team, WhoScored provides offsides, fouls, corners, extraterrestrials, dribbles, steals, and most extensive ball possession, passing success, and air confrontation success rates. However, the existing systems care too much of the goals. Even though these systems are mature and are popular among soccer fans, they are not helpful for the coaches to train the players. For example, a guard who has been breakthrough easily for 90 min was marked to be the best in the game because of a fortunate goal. On the other side, a goalkeeper who has made many outstanding saves was marked to be the worst because of a fumble, which the backs should take major responsibility. Fig. 1 is an example of Squawka [25]. The performance total score consists of four statistics. In addition, Bloomberg [26], Champion [27], Sky Sports [28], Simi Scout [29], FourFourTwo [30], Wyscout [31], and many computer games also have a marking system. Some of them calculate the total point automatically by using statistics. The other set standards and their experts and commentators will estimate and get a point.

C. Broad Learning System

These days, in addition to the growth of data in size, the data dimensions also increase tremendously. Taking the raw data with a high dimension directly to a neural network, the system cannot sustain its accessibility anymore. The challenge of solving a high-dimensional data problem becomes imperative recently. Two common practices to alleviate this problem are dimension reduction and feature extraction. Feature extraction is to seek, possible, the optimal transformation from the input data into the feature vectors. Common approaches, which have the advantage of easy implementation and outstanding efficiency, for feature selection include variable ranking, feature subset selection, penalized least squares, random feature extractions, such as nonadaptive random projections and random forest, or convolution-based input mapping, to name a few. BLS [6]–[9] is a new method in the deep structure. Some deep structures, such as a convolutional neural network (CNN) and long short-term memory (LSTM), perform well in terms of accuracy. But they suffer from long training sessions because of the large number of connecting parameters in filters and layers. In addition, in the era of big data, data are no longer stable, and we need to deal with stable data streams quickly. However, the deep structure encounters a complete retraining process if the structure is not sufficient to model the system, which requires more unnecessary time loss. The BLS is established in the form of a flat network, where the original inputs are transferred and placed as mapped features in feature nodes and the structure is expanded in the wide sense via "enhancement nodes." The incremental learning algorithms are developed for fast remodeling in broad expansion without a retraining process if the network deems to be expanded. Two incremental learning algorithms are given for both the increment of feature nodes (or filters in deep structure) and the increment of enhancement nodes. The designed model and algorithms are very versatile for selecting a model rapidly. Specifically, the system can be remodeled in an incremental way without entire retraining from the beginning.

Rank	1	2	3	4
Player	*Lionel Messi*	*Neymar*	*Harry Kane*	*Edinson Cavani*
Games	21	16	23	22
Mins	1890	1436	1952	1796
Defence	66	-9	-15	23
Attack	1521	1324	1166	903
Poss	111	98	-90	46
Total	1699	1413	1061	971

Fig. 1. This is part of a ranking list of four top players. Player point rank is used to assess player performance. This one consists of defense point, attack point, and possession point.

D. Visualization Analysis

It has been years since more visualized analysis has been applied to sports. One main application is modeling and placing data or guidelines on the original video. The other is to visualize statistics. An important previous work SoccerStories [32] proposed a visualization interface to support analysts in exploring soccer data and communicating interesting insights. It provides game phases into a series of connected visualizations. Actual movement data are usually miscellaneous and unsmooth. We need to simplify data and extract important semantic information from a mess of points and lines. A novel dynamic approach is presented in [33] that combines trajectory simplification and clustering techniques with the goal to support interpretation and understanding of movement patterns. It makes an attempt to process big soccer

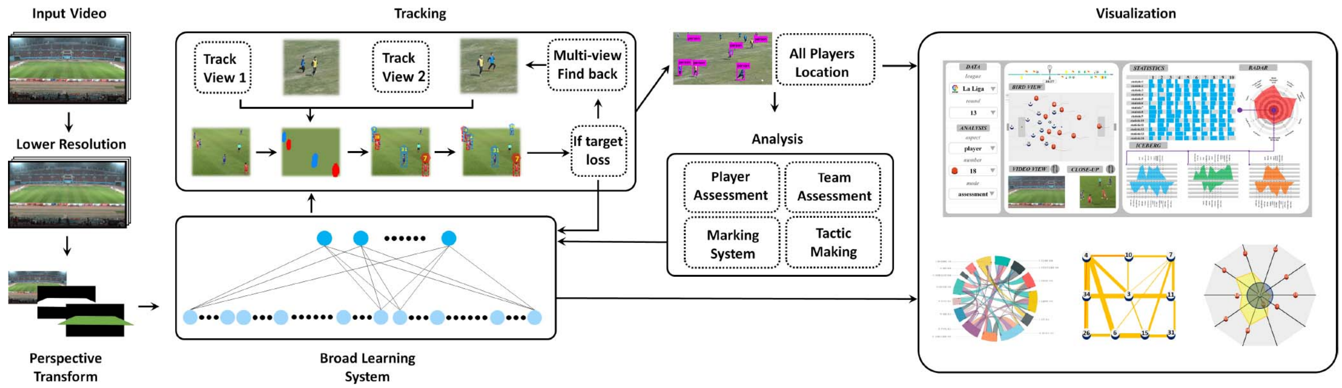


Fig. 2. Overview of the GreenSea system. First, we need a distributed system and network transmission to process the data. Input video resolution should be decreased. Then, we do the perspective transform to obtain the transition matrix. We use some algorithms to track and find back targets. The position data then will be input into Greensea. We can have a top view animation of the game. Statistics can be added by the computer and human. After calculation, we get figures and marking.

matches data. Traditional coaches watch the original video to gain information. It is first hand and all sided. However, it is also unfocused and time consuming. Statistics, figures, and tables, on the other hand, are abstract and information lost, which makes it difficult for the coaches to find relationships. One method is to combine video and movement data together [34]. It makes the data easy to understand. The work [35] starts by summarizing several related works to show how to analyze team sports. They identify important components of team sports data, exemplified by the soccer case, and explain how to analyze team sports data in general. Visualization for analyzing the soccer data is tried in [36]. It includes density information along each axis for the clustered data. The main idea is to visually represent the density distribution of each cluster along the axes. A visual analysis system is proposed in [37] for interactive identification of soccer patterns and situations that are of interest to the analyst.

III. Recurrent Discriminative Broad Learning System

We will first express the typical BLS mathematically. Then, the proposed RDBLS will be introduced. The field model will also be included to help understand how BLS helps accomplish the tracking mission. The workflow of the entire system is shown in Fig. 2.

A. Broad Learning System

In our system, we use BLS in two parts. One is to train a detection classifier to help to track, and the other is to train a marking system. We will introduce the involved algorithms in this section, while the different details will be clarified, respectively. The advantage of the BLS is that the learning can be updated dynamically and incrementally without going through a retraining process if the model deems to be expanded on additional feature nodes and enhancement nodes such that the learning is so efficient and effective. The basic network of BLS is in Fig. 3. In a general learning task, assume the input X is the input dataset, which is equipped with N samples, each with M dimensions, and Y is the output matrix which belongs to $\Re^{N\times C}$. For n feature mappings, each mapping generates k

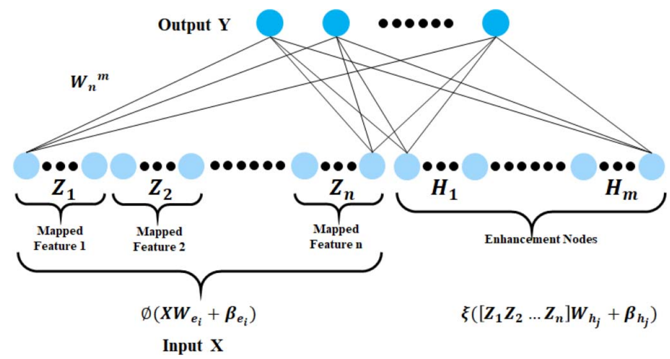


Fig. 3. Broad learning network.

nodes, can be represented as the equation of the form. The ith mapping feature could be calculated by

$$\boldsymbol{Z}_i = \theta_i(\boldsymbol{X}\boldsymbol{W}_i + \beta_i),\ \ i = 1, \ldots, n. \tag{1}$$

To speed up the training process, the enhancement nodes are obtained group by group. The enhanced feature is calculated by

$$\boldsymbol{H}_j = \xi_j\big([\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n]\boldsymbol{W}_j + \beta_j\big),\ \ j = 1, 2, \ldots, m. \tag{2}$$

The broad model can be represented as

$$\begin{aligned}\boldsymbol{Y} &= \big[\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n | \xi\big(\boldsymbol{Z}^n\boldsymbol{W}_{h_1} + \beta_{h_1}\big), \ldots, \xi\big(\boldsymbol{Z}^n\boldsymbol{W}_{h_m} + \beta_{h_m}\big)\big]\boldsymbol{W}^m \\ &= [\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n | \boldsymbol{H}_1, \ldots, \boldsymbol{H}_m]\boldsymbol{W}^m \\ &= \big[\boldsymbol{Z}^m | \boldsymbol{H}^m\big]\boldsymbol{W}^m \end{aligned}\tag{3}$$

and the updated weights are

$${}^x\boldsymbol{W}_n^m = \boldsymbol{W}_n^m + \big(\boldsymbol{Y}_a^T - \boldsymbol{A}_x^T\boldsymbol{W}_n^m\big)\boldsymbol{B} \tag{4}$$

where Y_a are the respective labels of additional X_a. This incremental learning saves time for only computing necessary pseudoinverse. This particular scheme is perfect for incremental learning for new incoming input data. In the incremental learning for BLSs, we connect each group of mapped features to a group of enhancement nodes in Fig. 4. For the input dataset, X, for n-mapped features and for n enhancement groups, the new construction is

$$\boldsymbol{Y} = \big[\boldsymbol{Z}_1, \xi\big(\boldsymbol{Z}^n\boldsymbol{W}_{h_1} + \beta_{h_1}\big)|, \ldots, \boldsymbol{Z}_n, \xi\big(\boldsymbol{Z}^n\boldsymbol{W}_{h_n} + \beta_{h_n}\big)\big]\boldsymbol{W}^n$$

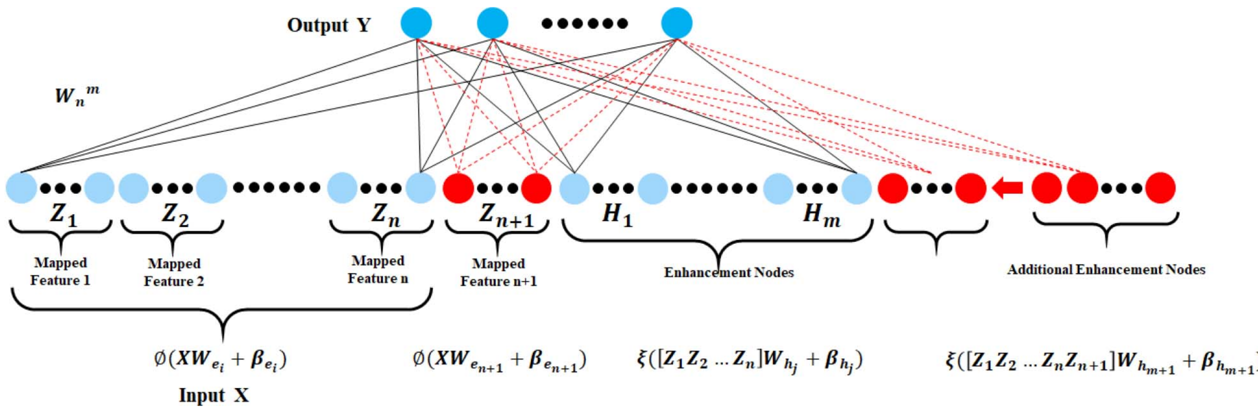


Fig. 4. Broad learning: increment of $n+1$ mapped features.

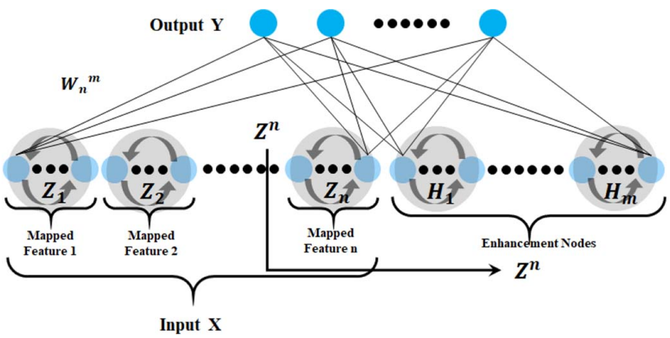


Fig. 5. BLS with alternative enhancement nodes establishment.

$$\triangleq \left[Z_1,\dots,Z_n|\xi\left(Z^nW_{h_1}+\beta_{h_1}\right),\dots,\xi\left(Z^nW_{h_n}+\beta_{h_n}\right)\right]W^m. \tag{5}$$

Algorithm 1 RDBLS Model

Input: Training samples X_n

Output: Weight ω_n

```
for i = 1; i <= n; i++ do
    Random W_ei, β_ei;
    Calculate Z_i(t) = θ_i(X W_ei + β_ei);
end for
Set feature mapping group Z_n^m(t) = [Z_1(t), ..., Z_m(t)];
for j = 1; i <= m; i++ do
    Random W_hj, β_hj;
    Calculate H_j(t) = ξ_j([Z_1(t), ..., Z_n(t)] W_j + β_j;
end for
Set the enhancement nodes group H^m = [H_1, ..., H_m];
Set A^m and calculate (A^m)^+;
W = W^{m+1};
```

B. Recurrent Discriminative BLS for Tracking

In the RDBLS, to capture the dynamic characteristics of video sequences, the enhancement nodes in each group are recurrently connected, and there are feedback connections in the enhancement units, which can be seen in Fig. 5. The basis equations now become

$$Z_i(t) = \theta_i((t)XW_{ei}+\beta_{ei}),\;\; i=1,\dots,n \tag{6}$$

and the enhancement nodes are

$$H_j(t) = \xi_j\big([Z_1(t),\dots,Z_n(t)]W_j+\beta_j\big),\;\; j=1,2,\dots,m. \tag{7}$$

Finally, the output is computed as

$$\begin{aligned} Y(t) &= [Z_1(t),\dots,Z_n(t)|H_1(t),\dots,H_m(t)]W_n^m \\ &= \left[Z^n(t)|H^m(t)\right]W_n^m. \end{aligned} \tag{8}$$

In conclusion, the RDBLS uses a recurrent structure to learn time-ordered information, which is important in object tracking. The flow is shown in Algorithm 1.

C. Field Model

This section is about the basic field model. We keep this section because it helps to understand some of the algorithms below. Our system aimed at working in complex situations and using different sources of videos. We designed some shoot schemes which are suitable for different environments. We do not want to see readers feeling confused when they find that example figures are obviously taken from different angles and distances. TV videos are usually of high quality, and researchers can obtain any match data of the top leagues. However, the disadvantage is also obvious. Initially, TV video means an ocean of useless fragments will alternate within 2 h, such as close-up and audience sidelights. Furthermore, a mobile shot means that we need to calculate the corresponding matrix in real time. Finally, the TV scene is always around the football, which means that soccer players are always outside the screen. Neither detecting image position of a panorama nor calculating matrix is easy and precision guaranteed. Now, a growing number of clubs install cameras around the stadium. If we use these videos, we can have information about all players from different angles. We can even calculate the altitude of football. Considering the complicated environments and different performances among cameras, we need several basic shoot schemes to choose from and combine. We have two basic standards. One is to make the view as wide as possible to reduce the number of cameras and enlarge the overlapping area. We will explain later that we need at least two views to calculate the altitude of football. The other is to make the player as big

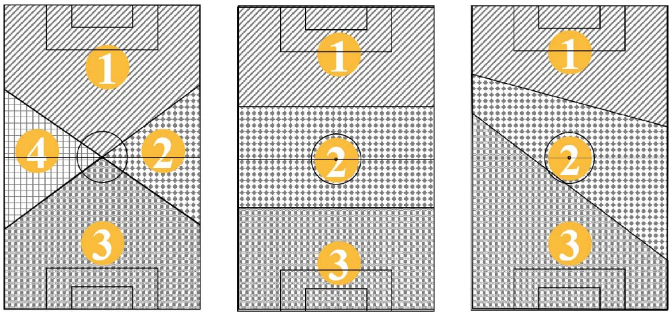


Fig. 6. Different shoot schemes of fixed cameras. We mark areas covered by different cameras with various shadings and yellow numbers. The cut lines are not visual range. They are used to distinguish the camera with the best view.

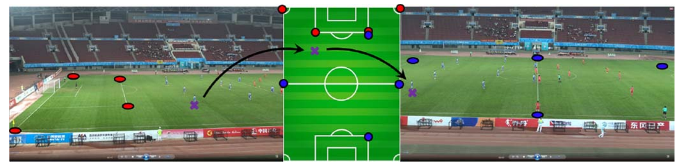


Fig. 7. Use multiview information to automatically find back the lost object. As long as the target is under detection in at least one view, we can use the perspective matrix to obtain its coordinates of bird view. Then, we use another matrix to find its location in another view. We use the original tracking parameters to search in the box and find the target back.

as possible, which is effective to improve the tracking accuracy. These two standards are in opposite directions, thus we need to weigh in Fig. 6.

Scheme A, when the places where cameras can be put are not high enough, monocular video information cannot handle the entire occlusion. When a player totally disappears behind another one, especially his teammate in the same suit, it is hard to find him back. We need different views to make a player show clearly in at least one camera, so we can use our multiview tracking method which will be introduced in the next section. According to the perspective theory, we need two angles to calculate the 3-D position of a football. In conclusion, any space on the football field needs to be recorded by two cameras. According to our test, scheme A is satisfying and economical. Scheme B, when the camera position is high, total occlusion happens hardly. An excellent monocular tracking method can handle the problem well. We can use more than two cameras collinear on the same side of the field. In each view, the target is bigger than scheme A, which promotes the tracking once again. We can find another side to put the fourth camera if we want to calculate the football altitude. Scheme C, when the spectator area is limited, or cameras of high performance are used. The advantage is that we can use several videos that have nearly the same optical center to make a panorama image, which removes the interaction problem of the player when they move past the junction of two cameras. Videos of one same optical center have high readability to researchers who may need to supervise and correct.

The size of a standard football pitch is 105 m × 68 m. We designed a 2-D pitch model which has the same scale as the pitch, and we can use its pixel location to do further research. After the tracking and detection part, we successfully obtained pixel location (x, y) of every player in each frame. When human eyes see nearer objects, those objects look bigger as compared to those which are far away. This is called perspective in a general way, whereas transformation is the transfer of an object from one state to another. So overall, the perspective transformation deals with the conversion of the 3-D world into a 2-D image. The same principle on which human vision works and the same principle on which the camera works. In this article, we need to use such a transformation matrix to find the match between the video pixel (normal view) and the 2-D model pixel (Bird's Eye View). In this article, we use a special form of perspective transformation, which is called the affine transformation. We select four pixels in the ordinary image, usually corner points of the field and the penalty area. Then, we find their coordinates in the model image. We use them to calculate the transformation matrix, thus we can obtain the new coordinates with the input of ordinary coordinates

$$\left[x', y', w'\right] = [u, v, w]\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}. \tag{9}$$

Perspective transformation needs to work twice to obtain the position of the same object in two videos. In another way, we can use the bird view as a bridge. If we lose an object in angle A and the other video tracking process who uses angle B can tell us the object is being tracked in a high-confidence level, we can use the bridge to find the area where the lost object should be and do the detection and continue tracking, as shown in Fig. 7. During a football game, football players are running on the ground, we can assume that their locations can be calculated by using pixels near their feet. However, football is different. It may be in the air. If we use the same method, the existence of height will result in an error.

Assume a line L_3 is represented as follows:

$$L_3 = t_3 + w_3 \times \text{direction}_3 \tag{10}$$

where t_3 is the position of the camera, w_3 is a constant, and direction_3 is its direction. We have two camera views, which are two unparallel lines L_1 and L_2 in the coordinate system

$$\begin{cases} L_1 = t_1 + w_1 \times \text{direction}_1 \\ L_2 = t_2 + w_2 \times \text{direction}_2 \end{cases} \tag{11}$$

we then calculate the intersection as

$$\begin{cases} \text{direction}_1 \cdot \text{direction}_3 = 0 \\ \text{direction}_2 \cdot \text{direction}_3 = 0. \end{cases} \tag{12}$$

When we have at least two different views of the ball and the parameters of cameras, we can calculate the 3-D location of the ball. It is a simple application of 3-D-reconstruction. As shown in Fig. 6, we use the left shoot scheme to set our cameras, which ensures that the system can have at least two different views at the same time. Equations (10)–(12) are actually used to calculate the 3-D position. In the ideal situation, the 3-D location (x, y, z) of the football can be calculated by the intersection of the above two lines. In a real application, the two lines are almost impossible to intersect because of

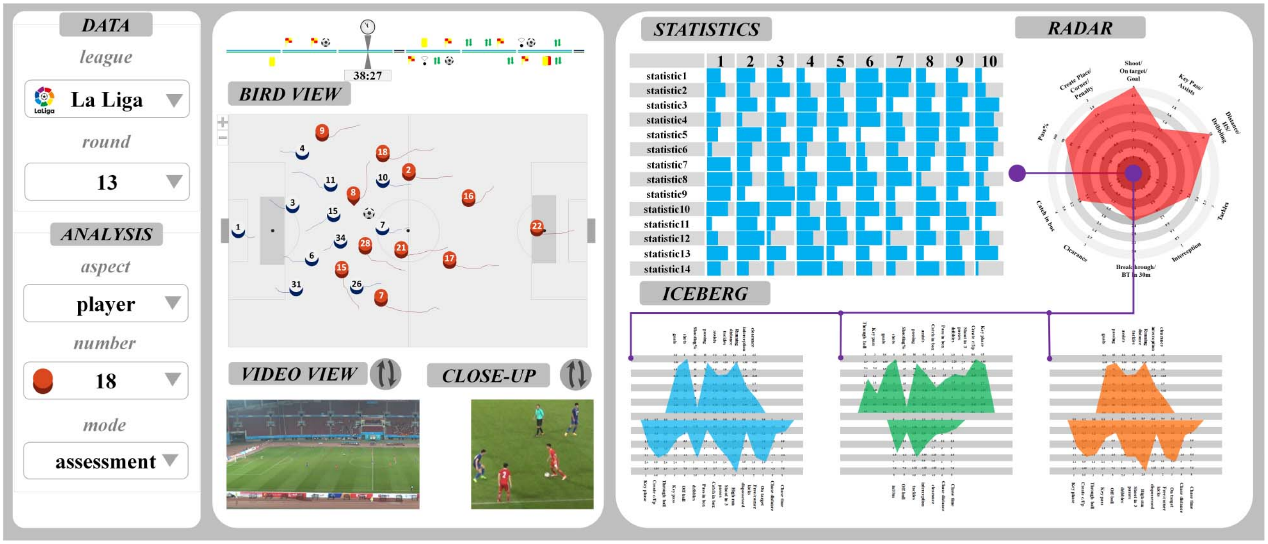


Fig. 8. Interface of GreenSea. Left column: options for choosing the players or the name of the football League. Middle column: different angles, including original video, close-up video, and bird-eye animation. The trajectories of players are also shown in this part. Right column: the statistics table and assessment figures of a player, including the radar figure and the iceberg figure, which can give the users more overall information.

Algorithm 2 Framework of Tracking

Input: Frames of soccer video I_n, confidence coefficient threshold T

Output: Sequence of a football player position $(\boldsymbol{X}, \boldsymbol{Y})$

for $i = 1$; $i <= n$; $i++$ **do**
  Lower resolution for I_i;
end for
Select origin position(x_1, y_1) on frame 1;
for $i = 2$; $i <= n$; $i++$ **do**
  Run RDBLS tracker;
  if confidence coefficient<T **then**
    Run detection to find back;
    Correct the position;
  end if
  Record position (x_i, y_i);
end for
return $(\boldsymbol{X}, \boldsymbol{Y})$;

the existence of errors. As a result, we calculate the midpoint of the perpendicular of the two lines as an approximation. We calculate the shortest distance between the two lines. The midpoint of the two intersections is a good estimation of the football location. The total tracking flow is in Algorithm 2.

IV. Quantitative and Visual Analysis

The design period of GreenSea lasts almost one year. We met coaches of different levels, recorded the training videos of regulation games, and participated in their meetings in order to understand their current workflow and how they select, improve, and evaluate a player. We also had some meetings with engineers of Champion discussing the field equipment. Our initial target clients are coaches from clubs and schools whose trainees have a wide age range, which also means the pitch situation would be complicated and several optional video capture schemes should be prepared. A pitch model that is suitable for all schemes was designed. During the longtime communication, we reached an agreement that the one-sidedness of traditional statistics bears a big responsibility in the deficiency of existing visualization. We decided to design some interaction methods, quantification approaches, and assessment methods for soccer games. The interface of the design is shown in Fig. 8.

A. New Skills Definition

Generally, soccer players' behavior can be grouped into two categories: 1) offense and 2) defense. Most researches use offensive statistics, including key pass, dribble, through ball, and shoot, and use clearance, interception, tackle and block to describe defense. According to the modern soccer concept, players have three types of actions in one offense period: create space, organize, and form a shot and three types of actions in one defense period: 1) delay; 2) restrict; and 3) compress. Some experts prefer to separate the third property: organizing ability. Sometimes when a player just stands at the right position, it is an excellent defense. All of those impel us to define more detailed statistics so that nearly every action of the player can be counted, whether it is a positive contribution or not. In Fig. 9, there are many skills defined by legend and description. Some of them are new statistics mostly designed to measure defensive and organizational abilities. Others are some traditional statistics that are also used for visualization and are presented as a supplement for those who do not know soccer well.

B. Marking System

Researchers now are trying to explain the soccer game by quantitative statistics. In most works, the key part from quantitative to qualitative is done by a human. By establishing a point-scoring system, we make the quantitative analysis itself

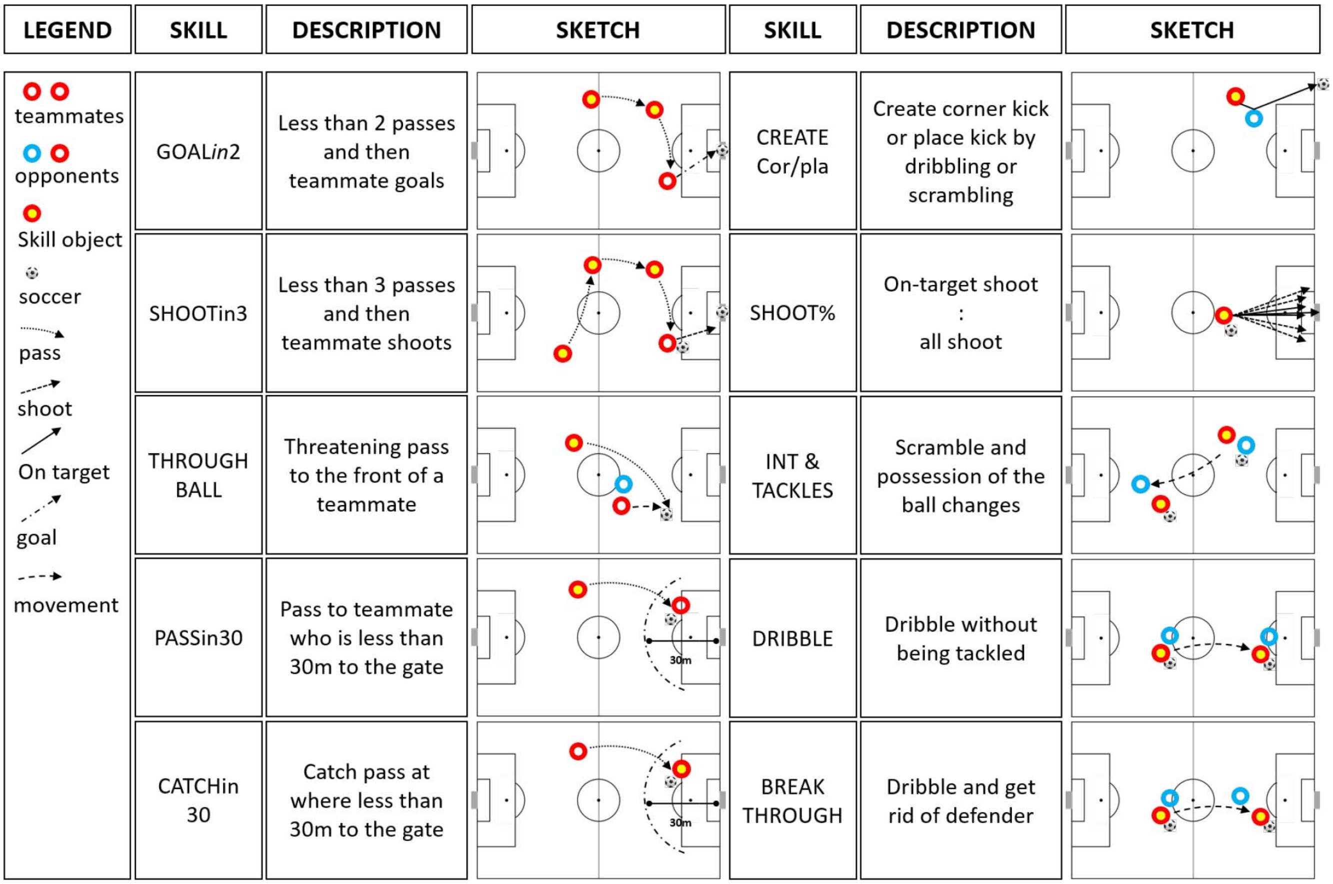


Fig. 9. Skill definition and visual description. In this figure, we define ourselves some skills by using text descriptions and illustrations. Those skills are not included in traditional skills, but the coaches find them useful and meaningful.

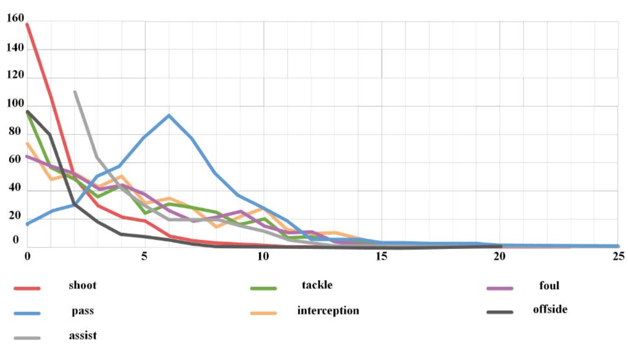


Fig. 10. Statistics have different distributions, which means the absolute value may not be a good scale.

meaningful. In other words: we can use a number to evaluate a player. For example, we will not say Messi had a great performance tonight. Instead, we can say that he got a 0.94 point which is his best performance in the last two months.

1) Normalization Standard: Statistics have been widely used. However, we are accustomed to using the absolute value to assess a player. For example, we use the shooting times to assess the offense ability of a forward. However, the absolute value may not be a good measure scale. As shown in Fig. 10, a guard hardly has the same chance as a forward. On the other hand, some statistics are dense in a little section, which makes the value of low discrimination. By studying the distribution of different statistics, we determine the scale by integrated using absolute value section and player amounts in the section. When coming to the assessment, we want the system to perform differently for different positions and among different levels. The degree of participation, which is used to describe the ball possession in one team, has three stairs, no participation, low participation, and high participation. The main purpose is to make the marking system in a good balance. On the one hand, we hope it has good discrimination. On the other hand, we hope it can buffer big change from little absolute value change.

2) Marking Rules: Real plus–minus is widely used in professional basketball. Though these days its objectivity has been questioned, it is still instructive and popular. Soccer is also a continuous game where offensive and defensive actions intertwine and change the situation on the football field. It does have periods. We can simply understand a period as from a possession change to an offense attempt or another possession change. The difference is that plenty of scores will not happen in a soccer game, and players do not have strict defensive matchups. We designed detailed statistics to make them take the place of scores and matchups. The marking system was established by giving weight points to actions in different phases. Besides, deep structure and learning suffer from a time-consuming training process because of a large number of connecting parameters in filters and layers. Moreover, it encounters a complete retraining process if the structure is not sufficient to model the system. For now, experts' manual evaluations are still more convective, but the disadvantage is the time-consuming process is obvious. We want to use the experts' evaluation as the reference to build an automarking system. By comparing the property of some deep learning methods, we finally use a BLS to help us train the network.

Considering the specificities of a soccer game, two standards are set for choice. One is to consider all offensive and defensive actions, whether it results in a score or a missed shot. The other is to only consider the scoring period. We give every action that may be involved in a period point. The game may be very dramatic, for example, a player made several mistakes while the opponent lost all of his or her shots; near the end, the player a goal and helped his or her team to win. We agreed with the experts about the above consideration, and reducing the weight involved to score is a wise choice, which means similar contributions taking place in different phases will have the same scores whether it ends up as a goal or not. Compared to the total score, sometimes the best five-phase scores and the worst five-phase scores could be a better reference. Detailed scores of the statistic are visualized.

C. Visual Design

Visual analysis has been applied in sports for a long time. A radar graph is usually used to describe the overall strength. The heat map is commonly used to show the position of a player during a soccer match. There is some visualization research [38]–[40] which equally uses the research from other sports as reference [41]–[44]. GreenSea organizes data in some promoted figures, which makes it scientific and intelligible to coaches.

1) Assessment Iceberg: The density ratio of ice to water is 0.9, which means the underwater volume of an iceberg is nine times its volume we can see over water. It is the same situation in the soccer game. Some hidden statistics are usually more significant than normal ones. When we arrange the statistics for different purposes, either upper water or underwater, we can have an intuitive judgment of a player. We have three main classification methods. As above, the first water surface is to classify the traditional statistics and advanced statistics. As shown in the figure, we can easily find good players even if their traditional data are not eye catching. Messi only ran 7000 m in a game. However, we found his high-speed running and dribbles are much more frequent than average. We can see he is still hardworking as the team core in that game. In different situations, coaches prefer different types of players. Offensive and defensive players are two frequently used features in a game. Some positions, such as forward and back, have their tendency, however, midfield players have different styles. The coach of a leading team tends to put a defender on the pitch near the end of a game. The second water surface is offense and defense.

There are different positions in a team and each of them has a different emphasis. When we want to assess a player to see how much we should pay him or her, it is better to have a parallel comparison among the same positions. We thus create different standards for different positions. On the other side, unlike basketball, soccer has a more detailed position classification which is far more than 11 in a real game, such as left wing back, left back, and left center back. It is part of the team tactics. Yet we still want to figure out which position a player will have the best performance on. As we can see in Fig. 11, the original figure is a radar figure. In fact, the radar figure is able to show the ability of one player as well as one soccer team. We use ten players trying not to make the radar figure look crowded. We can see three different shades of blue, they are three related statistics: 1) shoot; 2) on-target; and 3) goal. To be specific, the lightest blue area represents performance when shooting. The darkest blue area represents the evaluations for on-target. The middle one illustrates the assessment for the goal. The reason why we place all those different blue areas on one radar figure is that we want to better visualize all the statistics, and the users could have a more overall understanding on player's or the team's performance. We put the advanced statistics under the traditional statistics so that we can easily understand the performance. A disadvantage of the radar figure is that it has no focus. On the right is our created iceberg figure. As we discussed above, the statistics are extracted from the matrix. The surfaces are decided on different purposes. The on-water part and the below-water part have a sharp contrast.

2) Period-Divided Contour Heat Map: Physical power is vital in almost every professional sport. It is important to fully recognize the physical power condition of the entire team. However, building stamina is long-term work. In high-level matches, such as the world cup, it even has an insurmountable gap between different human species. Improving physical power is important while the distribution of physical power is also important. In short terms, it is more effective. The heat map is widely used in sports. The analysts count position points in a small area and mark it with a different color, usually warm toned means high occupied. When we talk about distributing physical power in a soccer game, we hold the view that different velocities are all important. Sprinting is important which may mean a key catch near the box. Walking is also important. A player needs to use walking cautiously because too much walking means that he cannot go back to his position in time and too little walking means he wastes too much power on useless movements. As above, middle intensity occupies nearly 30% of the entire running distance. Thus, the traditional heat map is not good enough. It is necessary to let the coaches know that what is the heat area of the high-speed movements and what is the changing tendency in a different period of a 90-min game. That is why we design the periods–velocity–heat-peak map in Fig. 12.

3) Pass Figure: There are some methods showing the passing condition in a game, which have different emphases. The first is the pass matrix. It can be filled with numbers or colors instead. If we ligature in order, we will find the key offensive route, and we can easily find the organization's core of a team. The second is the pass graph which has space information on it. The line breadth between players directly shows the pass amounts. GreenSea uses a 3-D pass figure (see Fig. 13), and we found that height is important. A place kick is an

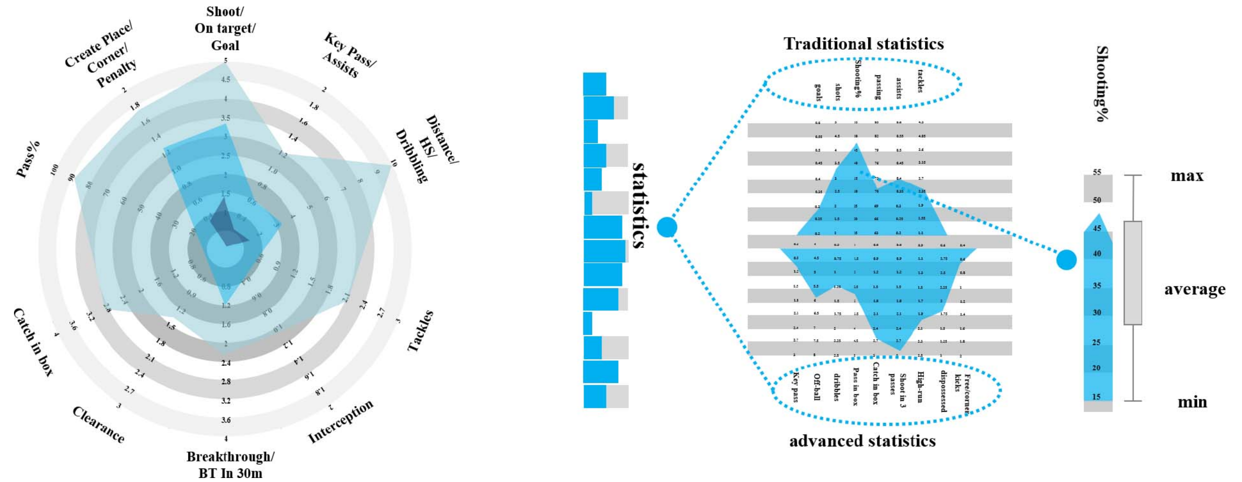


Fig. 11. Radar and Iceberg: we can have a visual evaluation of a player's comprehensive strength. If we have a clear aim, we can classify the data by skills. In this figure, we choose traditional and advanced statistics. Coaches can use this to find some useful players who do not have good-looking statistics.

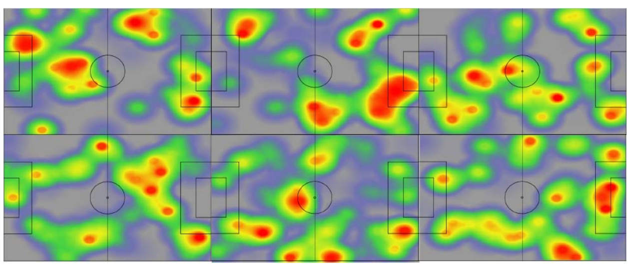


Fig. 12. Heat map of six periods in a game. A game will be divided into six 15-min parts to study mainly due to the physical and mental changes. Proceed from the running condition, coaches can find the weak period of the team.

TABLE I
DETECTION ACCURACY

Method	Accuracy (%)	Training Time (s)
DBN [10]	98.05	5017.62
SAE [11]	97.45	3420.37
SDA [12]	97.64	3489.73
DBM [45]	98.43	11652.20
HELM [46]	98.87	32.14
MLP [47]	96.63	2009.41
CNN [4]	94.92	2137.79
BLS	97.74	8.62
RDBLS	98.26	6.78

important tactic of the Chinese team. It is effective when we have a height advantage. When our players have a disadvantage in basic skills, which always means a lower possession percentage, the place kick is the best way to attack. Thus, we distinguish the ground ball and the air ball. We marked the route of threatening round in different colors, which reminds coaches the key part of a game. If a coach is interested in pass statistics between two players, he or she can use GreenSea to obtain all pass information, including time, location, result, and period. He or she can continue to obtain the entire period pass process, an animation, and the origin video.

V. Experimental Results

A. Quantitative Comparisons

In this section, experimental results are given to verify the proposed system. To confirm the effectiveness of the proposed system, detection experiments are applied on the 13th Chinese national games data as well as the training matches data of the Shanghai U20 soccer team. To prove the effectiveness of RDBLS, we will compare the ability of our method with the existing mainstream methods, including DBNs [10], SAEs [11], another version of SDA [12], deep Boltzmann machines (DBM) [45], an extreme learning machine (ELM)-based multilayer structure HELM [46], and multilayer perceptron-based methods (MLP) [47]. Although very deep structure CNN [3] can achieve very good results even on the ImageNet dataset, yet we only compare with the original deep CNN (LeNet-5) [4] for fair comparison as RDBLS only adopts linear feature mapping. Table I gives quantitative comparisons of different methods. In Table I, we find that our approach can achieve much less training time, while the accuracy is still maintained at a similar level as other methods. More results with different mapped features are given in Table II. To evaluate the effectiveness of our marking system, we compute the root-mean-square error (RMSE). From the result shown in Table III, when compared with either the traditional methods of elastic net regression (ENR), the lowest RMSE value and the fastest runtime can always be obtained by our proposed RDBLS. However, when comparing our RDBLS with the popular LSTM [48], we can realize a similar RMSE value. But the runtime of our RDBLS is faster than LSTM. The detailed experimental parameters could be checked in [46]. The values of RMSE and runtime in Table III show the effectiveness of our RDBLS-based method.

In Table IV, we test three features: 1) low-frequency Fourier transform feature (FFT); 2) Gabor magnitude; and 3) local binary pattern (LBP) instead of using the original unaligned images. They are used to perform the recognition, with the information of former sequences, we test the find-back success rate when the kernelized correlation filter (KCF) tracker [49] loses its target. The key point of KCF is that KCF uses a circulant matrix, and the surrounding area of the target to generate the positive and negative samples, which largely

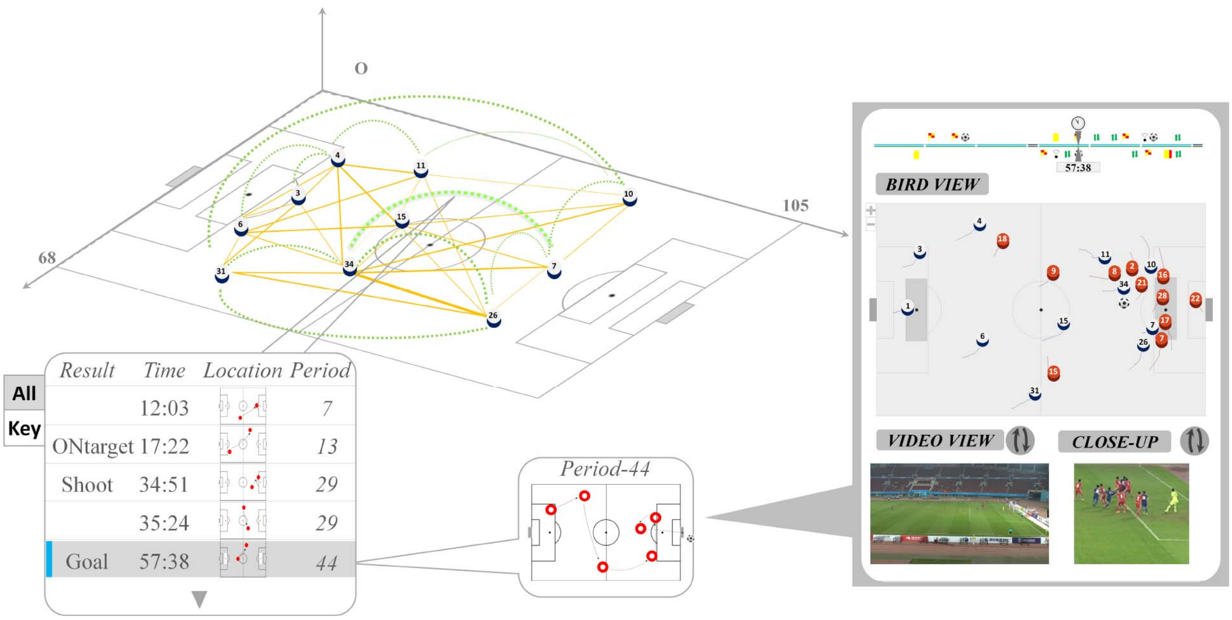


Fig. 13. 3-D pass figure: the traditional passing figure can only be used to visualize the numbers. We distinguish air pass from the ground pass as it is good statistics to see the style of a team or a player. According to the breadth of the lines, we can find the tactic core of a team. We can also find the main offensive routine.

TABLE II
DIFFERENT NUMBERS OF ENHANCED NODES

Feature Nodes	Enhanced Nodes	Accuracy (%)	Training Time (s)
100	4000	97.28	23.62
100	6000	97.49	33.34
100	8000	97.56	44.76
100	9500	97.68	52.32
100	10500	97.70	60.83
100	11000	97.70	63.29
100	14000	97.71	70.85
200	8000	97.65	56.44
200	11000	98.18	70.17
200	14000	98.33	76.12
400	14000	98.37	97.58
400	16000	98.66	109.69
400	18000	98.78	125.35
600	14000	98.82	108.22
600	18000	98.87	141.27
800	18000	98.88	173.82

TABLE III
SKILL MARKING

Video	ENR		LSTM [48]		RDBLS	
	RMSE	Runtime (s)	RMSE	Runtime (s)	RMSE	Runtime (s)
game1	3.35	603.27	2.43	801.22	2.60	7.65
game2	4.34	561.42	3.36	623.43	3.58	5.94
game3	3.76	542.03	1.87	707.16	2.23	6.32
game4	3.88	589.78	2.61	763.29	3.18	7.41
game5	3.96	567.42	1.82	721.32	3.15	8.54
game6	4.24	597.35	2.19	840.14	4.23	6.27
game7	4.03	539.21	1.54	761.33	3.97	8.61
game8	4.16	614.54	2.94	770.57	4.98	6.26

augments the number of training samples. As a result, KCF has largely increased the accuracy of traditional correlation-filter-based trackers. Besides, the processing speed of KCF can reach 172 fps, while over 30 fps could be called real-time performance. In conclusion, our RDBLS method performs with acceptable accuracy and a much better running time. Table V shows the effect of the increment of input patterns and enhancement nodes. In addition, to prove the effectiveness of our system for dealing with the special situation, we have still conducted some experiments on different multiview tracking methods. In this experiment, our system has an outstanding performance compared with the other multiview tracking algorithms [50]–[52]. We can conclude from Fig. 14 that our RDBLS-based multiview tracking method works well under the condition: small tracking target, similar features, and large background.

RDBLS
Xu et al.
Zhang et al.
Fleuret et al.

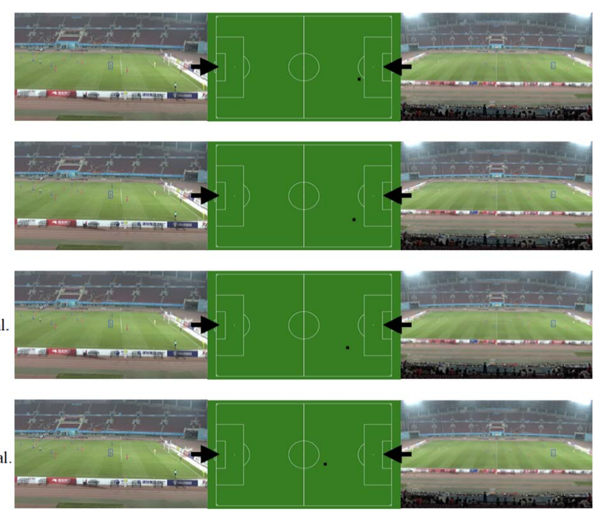


Fig. 14. Results in comparison with other multiview tracking methods, that is, Xu *et al.* [50], Zhang *et al.* [51], and Fleuret *et al.* [52]. The left and right columns show the match videos shot from different angles. The middle column presents the tracking results from the top view.

B. Study I (Selection and Training)

Teenage training is the foundation of the development of the soccer of a country. However, there are problems. Compared

TABLE IV
FIND BACK SUCCESS RATE AND RUNNING TIME USING THREE FEATURES AND FORMER POSITION DATA

Method	FFT		Gabor		LBP		Average
	Rate (%)	Runtime (s)	Rate (%)	Runtime (s)	Rate (%)	Runtime (s)	
SVM	15.8	22.6	49.8	16.4	27.7	20.36	31.1
SRC	37.2	3680.93	65.7	3722.18	58.6	3608.21	53.9
LRC	20.3	31.72	26.4	30.02	28.4	33.79	25.0
CRC_RLS	20.1	19.31	23.3	20.01	27.5	19.54	23.6
LCCR	27.8	20.09	58.6	21.65	64.2	21.23	50.2
GELM	41.2	12.33	64.3	14.46	54.6	14.55	53.3
RDBLS	45.7	0.58	70.3	0.75	69.3	0.67	61.8

TABLE V
TRACKING RESULTS ON A SOCCER GAME VIA INCREMENTAL LEARNING: INCREMENT OF INPUT PATTERNS AND ENHANCEMENT NODES

Number of Feature Nodes	Number of Enhancement Nodes	Number of Input Patterns	Testing Accuracy (%)	Each Additional Training Time (s)	Accumulative Training Time (s)	Each Additional Testing Time (s)	Accumulative Testing Time (s)
100	3000	10000	95.3	1.4	1.4	0.4	0.4
100	3000→4600	10000→ 20000	96.0	3.6	6.1	0.2	0.8
100	4600→6200	20000→30000	96.1	7.5	13.2	0.1	1.0
100	6200→7800	30000→40000	96.2	11.2	27.7	0.3	1.3
100	7800→9400	40000→50000	96.2	21.4	52.3	0.1	1.7
100	9400→11000	50000→60000	96.3	30.0	85.4	0.1	1.9

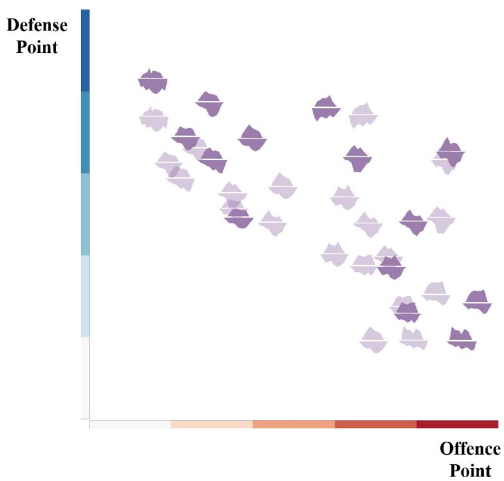


Fig. 15. Coaches can select two intriguing classifications. In this figure, coaches can concern about players of different emphasis: offense or defense. Each iceberg represents a player.

to the lack of money or equipment, the lack of experienced coaches and teachers is more serious. In one word, we do not have a long-term project for teenage soccer players to grow up. We do not have a standard for coaches to refer as well. Therefore, we cooperated with a coach who has a great deal of experience in teenage training. He has proposed two problems. First, coaches can only rely on their experience to select and train young trainees, which may lead to an omission or a misdirection of talents. Sometimes, a real genius may not have good physical quality. Second, even a club has the chance to go abroad and to have a game with strong soccer teams, they do not have a system to measure and compare their abilities. Coaches do not know the soccer level that the trainees of any age should reach, so they cannot find the real deficiency to take specialized training to remedy. GreenSea can help the coaches build a database from six-year-old children to adult players. Not only trainees but also young coaches would progress because of GreenSea.

1) Distribution Figure: Normally, there are four main player properties which coaches concern about: 1) offense ability; 2) defensive ability; 3) organizing ability; and 4) physical power. With the application of the marking system, we can obtain four points of players. GreenSea can use three or two properties as coordinates and fill the iceberg or radar figure of every player, as can be seen in Fig. 15. By accumulating data from different countries and all ages, we have a map of all levels. When a coach wants to select talented trainees, he could choose to show the player's radar figure of the same age only and judge from the position on the map. Using data spanning months, he can see the change in the batch of trainees. Who makes the most progress, whether trainees progress as the training project planned, and how far before our trainees can catch up with even-aged trainees from soccer power, are all clearly and obviously presented on the map.

2) Determining Position: There is a problem that will perplex every soccer player when he was young: which position to play. Years of experience can be referred. For example, a goalkeeper should be tall, a center-back is always strong and imperturbable and a forward ought to be good at shooting and stopping under the defense. However, there are also many famous superstars who do not satisfy regularity. Though it is the fact that huge change often happens on young players, it is meaningful to choose a future template for a child and continuing to keep a watchful eye on which position he has the highest score. By comparing the time-varying iceberg figure in Fig. 16, we can find a superstar whose iceberg at the appointed age is semblable and has a similar change. Coaches can let trainees play every position and find their best score and figure.

C. Study II (Tactics Analysis)

We were invited by the coaching staff of the Shanghai U20 soccer team to try out GreenSea in the 13th National Games. They expect us to focus on two problems. First, U20 players are not older than 20. Some of them are chosen at the age of 18. They are not experienced and may be physically down in a game. Coaches want a visualization of it so that young players can learn from it in training. In an actual game, to

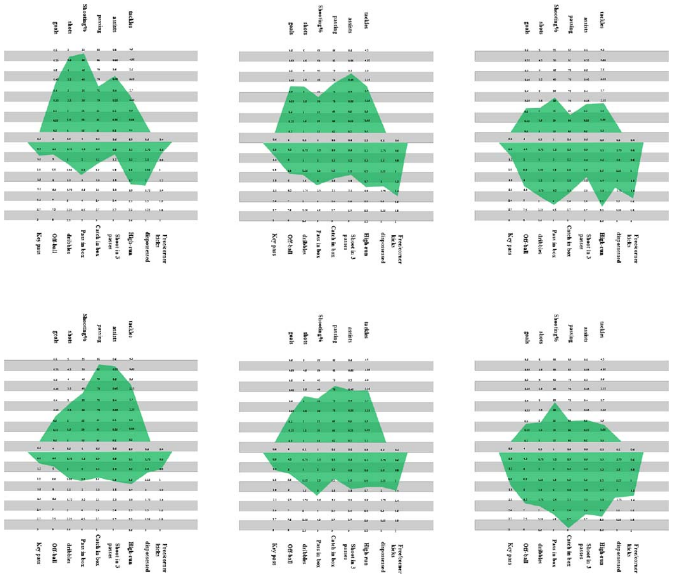


Fig. 16. Comparison of different positions and water surface.

analyze opponents' physical strength arrangement and to make the adjustment will help our players win the game. Second, the schedule is very tight. Teams did not have a long time to gather and train. There are only one or two gap days between games. Team Zhejiang is the final opponent of the team Shanghai. We use GreenSea to study the semifinal between Team Zhejiang and Team Hubei and develop a corresponding strategy.

1) Possession Change: Possession of football is the foundation of any tactic. Generally, the team that has a high percentage of possession has the initiative. They will set the pace of the game, making the opponent passive and have more chances to shoot. Relatively, the team which has a low percentage of possession also has its opportunity: counter attacking. Even so, a confident team will try everything to raise the percentage, which is why we attach importance to the possession change. As shown in Fig. 17, GreenSea counted the amounts, location, time, and types of losing possession of team Zhejiang. We have a conclusion: first, most of the possession changes happen due to being tackles other than an interception and interceptions mostly belong to blocks, which conforms to the fact that the U20 players are not experienced and are likely to make mistakes. Second, most possession changes happen at the side place. Third, if we consider time, the number of misplaced passes increases in the second half, which can hardly prevent us from associating the physical power problem.

2) Offence Mode: Different teams have different styles. The so-called style can be understood as a way that can maximize the advantages. Teams that have players with better basic skills are more possible to play pass-and-control. Teams having tall players prefer lofted ball. The style is especially obvious in the offense. GreenSea counted and drew the route of every shoot and two passes before it. According to Fig. 18, we can easily find that team Zhejiang tends to use passes from side to center, especially when it is close to the baseline. Most threatening attacks start from an air pass. A wide range of pass is not rare.

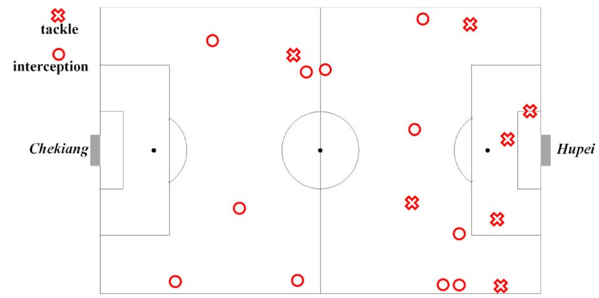


Fig. 17. GreenSea judges the possession of the ball from the position of players and football. There are two types of possession change: interception and tackle, which also can be automatically detected and then confirmed by a human.

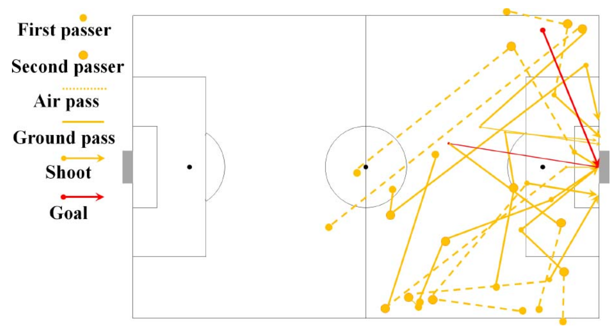


Fig. 18. GreenSea records and shows each shoot and two passes before. Coaches can see offensive mode of the opponent.

The header ability is not outstanding while their long shoot is extremely powerful.

3) Designing Tactics: After the visual analysis, we had a discussion with coaches and provided them with our advice. First, defense tactic, use close-marking defense other than blocking the pass route. When it comes to the second half, on the contrary, better to use block and create an interception. Second, pay attention to the side other than the center, no matter in offense or defense. Third, the main offense tactic should be created and seize the opportunity of a place kick, including a corner kick and a free kick.

VI. Conclusion

Soccer statistics' auto calculation and visual analytics are the tendencies in modern soccer. It helps establish an integral soccer training and assessment system. Although statistics and various kinds of charts are widely used, researches on how to make quantitative statistics more meaningful and designing suitable charts for soccer are rare. We present a long-term recurrent discriminative tracking system, which is able to largely reduce the running time with relatively high accuracy on our dataset. A marking system is also proposed, our system focuses on the comprehensive ability of players rather than only focusing on the goals. We visualize the statistics by using the radar figure, the iceberg figure, and the heat map. Each type of figure has a different emphasis, which makes the numerical assessment of players clearer for the coaches. The radar figure presents an overall evaluation for a player or a team, the iceberg figure focuses on the potential abilities of the players, and the heat map illustrates players' positions and key areas during one match. From the results and feedback, we are actually progressing in the right direction. In this article, we introduce our system GeenSea, and then we perform two experiments. In the first study, we prove our ability to help youth coaches train youth players. In the second study, we give GreeenSea practical use in formal games. Though we have obtained promising performance and enthusiastic feedback from the coaches, the short schedule and the strength gap increase the uncertainty of the outcome. In the future, we will obtain and apply more data. The machine-learning method is also planned to use to quantify the parameters of the marking system. By accumulating data, we hope the standard itself could become an achievement. We will also utilize the close shot to detect the actions for more comprehensive information.

REFERENCES

[1] (2019). *Opta Sports*. [Online]. Available: http://www.optasports.com

[2] (2019). *WhoScored*. [Online]. Available: http://www.whoscored.com

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Neural Inf. Process. Syst. (NIPS)*, 2012, pp. 1106–1114.

[4] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.

[5] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[6] C. L. P. Chen and Z. Liu, "Broad learning system: An effective and efficient incremental learning system without the need for deep architecture," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 1, pp. 10–24, Jan. 2018.

[7] C. L. P. Chen and Z. Liu, "Broad learning system: A new learning paradigm and system without going deep," in *Proc. 32nd Youth Acad. Annu. Conf. Chin. Assoc. Autom. (YAC)*, 2017, pp. 1271–1276.

[8] Z. Liu and C. L. P. Chen, "Broad learning system: Structural extensions on single-layer and multi-layer neural networks," in *Proc. SPAC*, 2017, pp. 136–141.

[9] Z. Liu, J. Zhou, and C. L. P. Chen, "Broad learning system: Feature extraction based on K-means clustering algorithm," in *Proc. ICCSS*, 2017, pp. 93–100.

[10] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006.

[11] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science*, vol. 313, no. 5786, pp. 504–507, 2006.

[12] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, "Extracting and composing robust features with denoising autoencoders," in *Proc. Int. Conf. Mach. Learn. (ICML)*, 2008, pp. 1096–1103.

[13] Y. Wu, J. Lim, and M.-H. Yang, "Online object tracking: A benchmark," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2013, pp. 2411–2418.

[14] S. Zhao, S. Zhang, and L. Zhang, "Towards occlusion handling: Object tracking with background estimation," *IEEE Trans. Cybern.*, vol. 48, no. 7, pp. 2086–2100, Jul. 2018.

[15] T. Zhou, F. Liu, H. Bhaskar, and J. Yang, "Robust visual tracking via online discriminative and low-rank dictionary learning," *IEEE Trans. Cybern.*, vol. 48, no. 9, pp. 2643–2655, Sep. 2018.

[16] J. Xiao, R. Stolkin, Y. Gao, and A. Leonardis, "Robust fusion of color and depth data for RGB-D target tracking using adaptive range-invariant depth models and spatio-temporal consistency constraints," *IEEE Trans. Cybern.*, vol. 48, no. 8, pp. 2485–2499, Aug. 2018.

[17] J. Shen, Z. Liang, J. Liu, H. Sun, L. Shao, and D. Tao, "Multiobject tracking by submodular optimization," *IEEE Trans. Cybern.*, vol. 49, no. 6, pp. 1990–2001, Jun. 2019.

[18] K. Bernardin and R. Stiefelhagen, "Evaluating multiple object tracking performance: The CLEAR MOT metrics," *EURASIP J. Image Video Process.*, vol. 2008, no. 1, pp. 1–10, 2008.

[19] A. Milan, S. Roth, and K. Schindler, "Continuous energy minimization for multitarget tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 58–72, Jan. 2014.

[20] J. Xing, H. Ai, and S. Lao, "Multi-object tracking through occlusions by local tracklets filtering and global tracklets association with detection responses," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2009, pp. 1200–1207.

[21] B. Yang and R. Nevatia, "An online learned CRF model for multi-target tracking," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2012, pp. 2034–2041.

[22] G. Shu, A. Dehghan, O. Oreifej, E. Hand, and M. Shah, "Part-based multiple-person tracking with partial occlusion handling," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2012, pp. 1815–1821.

[23] Y. Xiang, A. Alahi, and S. Savarese, "Learning to track: Online multi-object tracking by decision making," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015, pp. 4705–4713.

[24] M. Keuper, S. Tang, Z. Yu, B. Andres, T. Brox, and B. Schiele, "A multi-cut formulation for joint segmentation and tracking of multiple objects," 2016. [Online]. Available: arxiv.abs/1607.06317.

[25] (2019). *Squawka*. [Online]. Available: http://www.squawka.com

[26] (2019). *Bloomberg*. [Online]. Available: http://www.bloomberg.com

[27] (2019). *Champion*. [Online]. Available: http://data.champdas.com

[28] (2019). *Sky Sports*. [Online]. Available: http://www.skysports.com

[29] (2019). *Simi Scout*. [Online]. Available: http://www.simi.com/en/products/behavior-and-tactical-analysis/simi-scout.html

[30] (2019). *FourFourTwo*. [Online]. Available: http://www.fourfourtwo.com

[31] (2019). *Wyscout*. [Online]. Available: http://wyscout.com

[32] C. Perin, R. Vuillemot, and J.-D. Fekete, "SoccerStories: A kick-off for visual soccer analysis," *IEEE Trans. Vis. Comput. Graphics*, vol. 19, no. 12, pp. 2506–2515, Dec. 2013.

[33] D. Sacha *et al.*, "Dynamic visual abstraction of soccer movement," *Comput. Graph. Forum*, vol. 36, no. 3, pp. 305–315, 2017.

[34] M. Stein *et al.*, "Bring it to the pitch: Combining video and movement data to enhance team sport analysis," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 13–22, Jan. 2018.

[35] M. Stein *et al.*, "How to make sense of team sport data: From acquisition to data modeling and research aspects," *Data*, vol. 2, no. 1, pp. 1–23, 2017.

[36] H. Janetzko, M. Stein, D. Sacha, and T. Schreck, "Enhancing parallel coordinates: Statistical visualizations for analyzing soccer data," *Electron. Imag.*, vol. 2016, no. 1, pp. 1–8, 2016.

[37] M. Stein *et al.*, "Visual soccer analytics: Understanding the characteristics of collective team movement based on feature-driven analysis and abstraction," *ISPRS Int. J. Geo Inf.*, vol. 4, no. 4, pp. 2159–2184, 2015.

[38] A. Cox and J. Stasko, "SportVis: Discovering meaning in sports statistics through information visualization," in *Proc. InfoVis Poster*, 2006, pp. 114–115.

[39] N. Henry, J.-D. Fekete, and M. J. McGuffin, "NodeTrix: A hybrid visualization of social networks," *IEEE Trans. Vis. Comput. Graphics*, vol. 13, no. 6, pp. 1302–1309, Nov./Dec. 2007.

[40] N. Adrienko and G. Adrienko, "Spatial generalization and aggregation of massive movement data," *IEEE Trans. Vis. Comput. Graphics*, vol. 17, no. 2, pp. 205–219, Feb. 2011.

[41] Y. Wu *et al.*, "iTTVis: Interactive visualization of table tennis data," *IEEE Trans. Vis. Comput. Graphics*, vol. 24, no. 1, pp. 709–718, Jan. 2018.

[42] T. Polk, J. Yang, Y. Hu, and Y. Zhao, "TenniVis: Visualization for tennis match analysis," *IEEE Trans. Vis. Comput. Graphics*, vol. 20, no. 12, pp. 2339–2348, Dec. 2014.

[43] H. Pileggi, C. D. Stolper, J. M. Boyle, and J. T. Stasko, "SnapShot: Visualization to propel ice hockey analytics," *IEEE Trans. Vis. Comput. Graphics*, vol. 18, no. 12, pp. 2819–2828, Dec. 2012.

[44] Y. Onoue, N. Kukimoto, N. Sakamoto, and K. Koyamada, "Minimizing the number of edges via edge concentration in dense layered graphs," *IEEE Trans. Vis. Comput. Graphics*, vol. 22, no. 6, pp. 1652–1661, Jun. 2016.

[45] R. Salakhutdinov and G. Hinton, "Deep Boltzmann machines," in *Proc. AISTATS*, vol. 5, 2009, pp. 448–455.

[46] J. Tang, C. Deng, and G.-B. Huang, "Extreme learning machine for multilayer perceptron," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 4, pp. 809–821, Apr. 2016.

[47] C. M. Bishop, *Pattern Recognition and Machine Learning* (Information Science and Statistics). Berlin, Germany: Springer-Verlag, 2006.

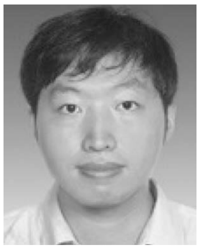
[48] Q. Li, X. Zhao, and K. Huang, "Learning temporally correlated representations using LSTMs for visual tracking," in *Proc. IEEE ICIP*, 2016, pp. 1614–1618.

[49] J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 3, pp. 583–596, Mar. 2015.

[50] Y. Xu, X. Liu, Y. Liu, and S.-C. Zhu, "Multi-view people tracking via hierarchical trajectory composition," in *Proc. IEEE Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 4256–4265.

[51] S. Zhang, X. Yu, Y. Sui, S. Zhao, and L. Zhang, "Object tracking with multi-view support vector machines," *IEEE Trans. Multimedia*, vol. 17, no. 3, pp. 265–278, Mar. 2015.

[52] F. Fleuret, J. Berclaz, R. Lengagne, and P. Fua, "Multicamera people tracking with a probabilistic occupancy map," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 267–282, Feb. 2008.



Bin Sheng received the B.A. degree in English and the B.Eng. degree in computer science from the Huazhong University of Science and Technology, Wuhan, China, in 2004, the M.Sc. degree in software engineering from the University of Macau, Macau, China, in 2007, and the Ph.D. degree in computer science and engineering from the Chinese University of Hong Kong, Hong Kong, in 2011.

He is currently an Associate Professor with the Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai, China. His current research interests include image/video processing, virtual reality, machine learning, and computer graphics.

Prof. Sheng is an Associate Editor of the IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY.

Ping Li (Member, IEEE) received the Ph.D. degree in computer science and engineering from the Chinese University of Hong Kong, Hong Kong, in 2013.

He is currently a Research Assistant Professor with the Hong Kong Polytechnic University, Hong Kong. He has one image/video processing national invention patent, and has excellent research project reported worldwide by *ACM TechNews*. His current research interests include image/video stylization, artistic rendering and synthesis, and creative media.

Yuhan Zhang received the B.A. degree in French and the B.Eng. degree in information engineering from Shanghai Jiao Tong University, Shanghai, China, in 2019, where she is currently pursuing the M.Eng. degree in information engineering with the SJTU-ParisTech Elite Institute of Technology.

Her current research interests include image/video processing, machine learning, and computer vision.

Lijuan Mao received the Ph.D. degree in sports science from East China Normal University, Shanghai, China, in 2002.

She is currently a Deputy Director of the Shanghai Municipal Education Commission, Shanghai. She is also a Full Professor and the Former Vice President with the Shanghai University of Sports, Shanghai. Her current research interests include sports science and sports kinesiology.

C. L. Philip Chen (Fellow, IEEE) received the M.S. degree in electrical engineering from the University of Michigan at Ann Arbor, Ann Arbor, MI, USA, in 1985, and the Ph.D. degree in electrical engineering from Purdue University, West Lafayette, IN, USA, in 1988.

He is currently a Chair Professor and the Dean of the School of Computer Science and Engineering, South China University of Technology, Guangzhou, China. Being a Program Evaluator of the Accreditation Board of Engineering and Technology Education in the USA, for computer engineering, electrical engineering, and software engineering programs, he successfully architects the University of Macau's Engineering and Computer Science programs receiving accreditations from Washington/Seoul Accord through the Hong Kong Institute of Engineers (HKIE), of which is considered as his utmost contribution in engineering/computer science education for Macau as the former Dean of the Faculty of Science and Technology. His current research interests include systems, cybernetics, and computational intelligence.

Dr. Chen received the IEEE Norbert Wiener Award in 2018 for his contribution in systems and cybernetics, and machine learnings. He is also a 2018 Highly Cited Researcher in Computer Science by Clarivate Analytics. He was a recipient of the 2016 Outstanding Electrical and Computer Engineers Award from his alma mater, Purdue University. He was the IEEE Systems, Man, and Cybernetics Society President from 2012 to 2013. He is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS: SYSTEMS, and an Associate Editor of the IEEE TRANSACTIONS ON FUZZY SYSTEMS and the IEEE TRANSACTIONS ON CYBERNETICS. He is currently a Vice President of the Chinese Association of Automation (CAA). He is a member of the Academia Europaea, the European Academy of Sciences and Arts, and the International Academy of Systems and Cybernetics Science. He is a Fellow of AAAS, IAPR, CAA, and HKIE.