Online Opponent Formation Identification Based on Position Information

Takuya Fukushima[1], Tomoharu Nakashima[1], and Hidehisa Akiyama[2]

[1]Osaka Prefecture University, Osaka, Japan
takuya.fukushima@edu.osakafu-u.ac.jp
tomoharu.nakashima@kis.osakafu-u.ac.jp
[2]Fukuoka University, Fukuoka, Japan
akym@fukuoka-u.ac.jp

Abstract. The aim of this paper is to propose a method for identifying the opponent formation type in an online manner during a game. To do so, opponent teams were clustered according to the position of their players. Each cluster is investigated to determine the difficulty for our team to defeat such a strategy. Then, an identification model is used online to determine if the opponent team adopts such a strategy or not. Furthermore, we also investigate how quickly the opponent formation can be identified. Through a series of computational experiments, it is shown that the model can identify opponent formation quickly and accurately. Therefore, we show the effectiveness of the identification model to switch our strategy.

Keywords: soccer simulation · team strategy · machine learning · log analysis · online identification

1 Introduction

In the domain of RoboCup 2D soccer simulation league, various strategies are implemented by teams to win the competition. For this purpose, it is important for the teams to determine the opponent strategy. Researches are conducted with various objectives such as training optimal decision models for individual players, predicting opponent's behavior to make decisions [1], taking the strategy that is the most suited against opponent teams [2], as well as improving players' behavior like pass and dribble [3][4]. One of the essential tasks in the development of a team in this league is to design an effective strategy. Since there is no perfect strategy, it is difficult to win against all teams with only one strategy. Therefore, to outperform a particular opponent team it is important to adopt the right counter-strategy against it. In order to select the most effective strategy, it is necessary for the team to have a set of strategies. Furthermore, the sooner the opponent team's strategy is identified, the sooner the team can adapt its strategy in order to increase its chance to win the game.

We proposed a model that determines the best player formation for corner-kick situations [2]. The model in [2] consists of two modules, learner and selector.

The learner analyzes opponent team distributions obtained by applying hierarchical clustering and estimates the performance of our strategies against each cluster of opponents by using Bayesian estimator. The selector returns the best strategy to apply according to the estimations done by the learner part. It was shown that it is effective to change strategies by using only positional information of opponents. However, soccer games does not only consist in corner-kicks, thus such a method should be extended to any kind of game mode.

In this paper, we develop a new learner and focus on the method for identifying opponent teams' strategies. The strategy of the opponent team cannot be known right after the game starts. Thus, the identification of the opponent strategy must be done online during the game. This work assumes that strategies are expressed by opponent players' position. First, opponent teams are clustered according to the position of their players and each cluster is investigated to determine the difficulty for our team to defeat such a strategy. Then, an identification model is used online during a game to identify the strategy type (i.e., strategy cluster) of the current opponent team. The effectiveness of the proposed identification model is examined through a series of computational experiments, in order to verify the relationship between elapsed time and identification rate.

2 Related Work

In the research community of the RoboCup Soccer Simulation 2D, various works that focus on the analysis of opponent teams by using a coach agent have been proposed. For example, Gregory *et al.*[5] developed a coach agent that learns offensive and defensive advices by using decision trees. Ramin *et al.*[6] researched the coach development by using rule-based expert systems and decision-making trees. The coach learns to predict agent behavior and automatically generates advices to improve team's performance. Mazda *et al.*[7] also researched opponent modeling for prediction two-layered case based reasoning.

Additionally, the relationship between strategy and positioning was well investigated in [8]. For the strategy representation, some methods that focus on player positioning have been proposed. For example, Visser *et al.*[9] proposed a system for recognizing opponent's formations and then applying a counter formation. Luis *et al.*[10] proposed a method to change the players positioning according to the strategy of the team. In addition, Akiyama and Noda [11] proposed a player positioning method based on Delaunay Triangulation built from a set of ball coordinates. Other analysis of opponent team's strategy was proposed as Riley and Veloso [12] where they proposed a method for identifying the opponent team by recording the positions, passes and dribbles of opponent players. Faria *et al.*[13] proposed a formation classification method by using players and ball coordinates. However, the proposed method is not accurate enough and requires a large computation time amount. Therefore, a light and highly accurate method is required to perform strategy adaptation during a game.

3 Formation Identification

In this paper, we define a team strategy as the player's positioning during a game. Is is assumed in this paper that opponent teams do not change their strategy during a game. In order to identify the strategy of the opponent team, it is necessary to recognize the features of the formation. In this section, we first describe how the formation's features are extracted. Then, we describe the method used to identify the strategy type of the opponent team.

3.1 Data Extraction Based on Opponent Position Information

To use players' position as the inputs of the learning model requires to consider their uniform numbers. Thus, the order of the players would be a problem in the construction of the model. To cope with this issue, an opponent formation is numerically expressed by discretizing the soccer field by a grid as shown in Fig.1. Then, the number of players present in each cell is counted. The value of each cell is used as the input of the learning model. The value in each cell shows the number of opponent players at a certain cycle, and the results are integrated. Then, the average value is computed by dividing the integration obtained so far by the number of observed cycles. This set of the average values is used as input data of our identification model. For example, if the field is discretized by a grid of size 6×4, opponent formation is expressed by a 24-dimensional vector.

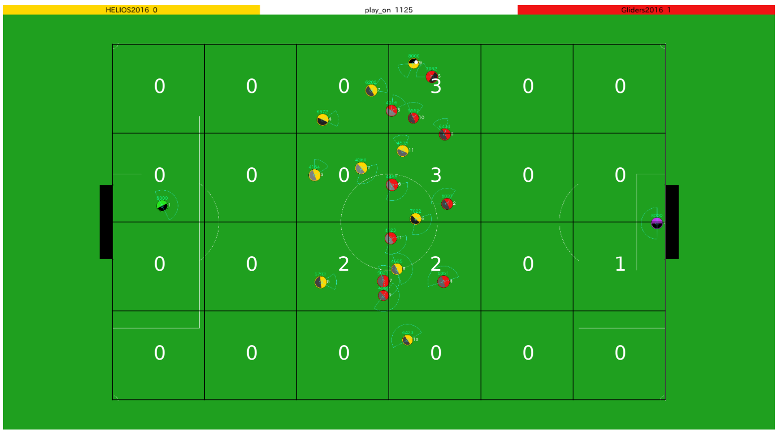

Fig. 1. Discretization of the soccer field by a grid of size 6×4.

3.2 Formation Identification Model

To investigate the effectiveness of the identification model, we compare the performance of three different supervised-learning-based classification methods:

Neural Network (NN), Support Vector Machine (SVM) and Random Forest (RF). Regardless the model used, the vector obtained in the previous section is used as input of the classification and each input is labeled by the corresponding opponent's formation. The methods used for labeling and generation of training data are explained in the following section.

4 Labeling Training Data

Opponent teams are assumed to try to counter our basic gameplay. Opponent teams' strategies are categorized according to the weaknesses of our formations. Training data are labeled according to this categorization. In this section, we first explain the criterion used to make the decision of changing our strategy or not. Secondly, we describe the opponent's strategy classification method and the labeling of training data.

4.1 Weaknesses Identification

In this work, our current strategy is considered as weak if our team has the ball for a large amount of consecutive steps during the game but fails to score a goal. In such a case, out team should change its strategy.

In order to identify such a situation, we define a weakness indicator. It is calculated from a lot of game logs for each opponent teams. Opponent teams for which the average value is equal to or larger than a particular threshold are considered as difficult to defeat for our team. In this case, we should adapt our strategy against this set of opponents. The method used to compute such a value is detailed in Section 5.

4.2 Strategy-Type Labeling Based on Opponent Positioning

In order to investigate the typical defensive formations, opponents teams determined from the previous section were clustered. To do so, a Gaussian mixture distribution is used, and its hyper-parameters are optimized using the Expectation Maximization (EM) algorithm. The optimal number of clusters is determined by using Calinski-Harabasz index [14]. The index is defined as the ratio of the within-cluster dispersion to the between-cluster dispersion.

Training data are labeled according to the different resulting clusters. These data are used to train the classification model where the number of labels conforms to the optimal number of clusters.

5 Experiments

Training data are labeled according to the different resulting clusters as described in Section 4. These data are used to train the classification model where the number of labels conforms to the optimal number of clusters. At first, we investigate the optimal number of clusters and classify the typical defensive formations of opponent teams. In a second experiment we identify the formation used by each opponent teams according to the locations of their players, in order to determine the most appropriate strategy in each situation. At the same time, we verify how many cycles are enough to identify the opponents' formations. For this purpose, we experiment with various numbers of elapsed steps in order to check the accuracy rate. The effect of the grid size on the accuracy rate is also investigated.

5.1 Clustering Process

In this first experiment, we classify the formation of opponent teams considered as difficult to defeat for our team and for which we should reconsider our strategy. We define a weakness indicator as in (1), where *p(k)* is the ball possession time of our team and *g(k)* is the score at the game *k*.

$$value(k) = d(g(k)) \cdot p(k), \tag{1}$$

$$d(x) = \begin{cases} 1 & (x = 0), \\ 0 & (x \geq 1). \end{cases} \tag{2}$$

The ball possession rate is estimated from the game logs. The team is considered to possess the ball if two consecutive kicks are done by players of the same team.

By determining the optimal number of clusters, we examine the number of distinct formation. Game logs were generated by making our team playing 200 games against various opponent teams. In the following, we made our team, HELIOS [15], playing against eleven opponent teams: CYRUS2014 [16], InfoGraphics [17], HERMES2015, Gliders2016 [18], FURY [19], HERMES2016 [20], MarliK2016 [21], Ziziphus [22], FRA-UNIted [23], WrightEagle [24] and Ri-one [25] that participated in the RoboCup competitions between 2014 and 2016. These teams are selected by using the weakness indicator. In this experiment, only the game logs of the first half were used. We tried to determine various different numbers of clusters varying between 2 and 10. The number of clusters with the highest Calinski-Harabasz index was selected as the optimal number of clusters. The formations of the opponent teams were expressed by discretizing the field by a grid of size 30×20. Thus, an opponent formation is represented by a 600-dimensional vector.

The experimental results are shown in Fig.2. From the result, three or four clusters seem to be the optimal number. When the number of clusters is set to three, as shown in the Fig.3, teams are categorized into two typical defensive strategies plus one strategy considered as normal. Such Fig.3 above teams (like CYRUS2014, HERMES2015, FURY and Ziziphus) whose the defensive strategy consists in gathering all the players in front of the goal. The second defensive strategy we determined is employed by teams like InfoGraphics(Fig.3 below) that aligns vertically its players to prevent the other team to perform passes going through the defense.

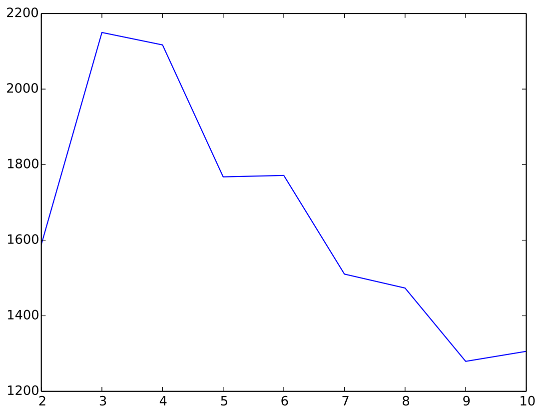

Fig. 2. Variation of the Calinsky-Harabasz index according to the number of clusters.

When the number of clusters is set to four, a new type of defensive strategy (Gliders2016) looking likes the straight line strategy explained previously in addition to the results observed from the three clusters. The difference between InfoGraphics and Gliders2016 is that Gliders2016 is aggressive when holding balls, whereas InfoGraphics is passive. Therefore, InfoGraphics and Gliders2016 are treated as the same formation, and the optimum number of clusters is 3.

Let us call the formation depicted in the top of Fig.3 as "wall", the formation depicted in the bottom of Fig.3 as "line", and the other formation as "normal". We decide to change our strategy against "wall" and "line" formation teams. It should be noted that WrightEagle is classified as "normal". The aim of the proposed method are to find "wall" defense teams, thus the result that WrightEagle is classified as "normal" corresponds to our expectations.

5.2 Formation Identification

In this section, according to the results obtained in the previous experiment, the type of opponent teams' formations are identified according to their player formations. In order to verify if the proposed method can be used online, we investigated the relationship between amount of time spent to observe the opponent team and the accuracy rate of the classification model. We also investigated the influence of the grid size on the accuracy rate.

Based on the results of the first experiment, we trained different models used to recognize three classes. Fields are are used five different discretization of the soccer field: 6×4, 12×8, 15×10, 24×16, 30×20. We made our team, Helios, play 19 games against teams who participated in RoboCup 2016, CYRUS2014, InfoGraphics, HERMES2015 and WrightEagle. We focused only on the first half

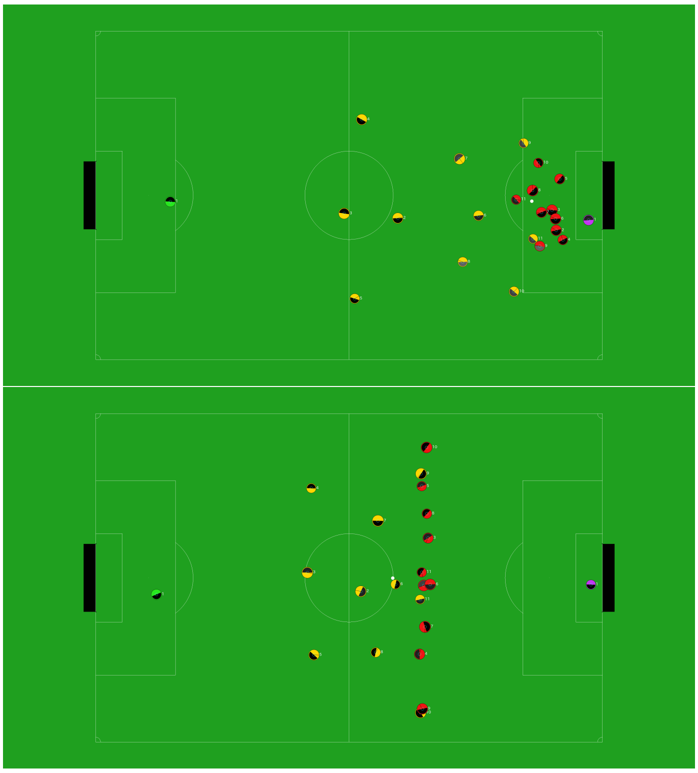

Fig. 3. Typical defensive formations. Top: "wall", bottom: "line".

of each game log. CYRUS2014 and InfoGraphics participated in RoboCup2014. HERMES2015 and WrightEagle participated in RoboCup 2015. Among the wall teams, CYRUS2014 and FURY change strategies according to goal difference. For this reason, we only used game data in which Helios did not score any goal against these two teams. For each class, we extracted position information from about 3000 logs and treated them as training data. Table1 summarizes the resulting labels. To do so, we performed the classification by testing three different classification models, a NN, a SVM and a RF. The hyper-parameters we employed for each model are summarized in Table2.

Table 1. Opponent teams' formation labels.

Opponent Team	label
CYRUS2014	wall
HERMES2015	wall
FURY	wall
Ziziphus	wall
InfoGraphics	line
Gliders2016	line
others	normal

Table 2. Hyper-parameters used for the classifiers.

Classifier	Parameter	Setting
NN	Activation function	Logistic function
	Optimization algorithm	L-BFGS method
	Structure	3 layers
	Number of neurons in Input layer	The number of grid
	Number of neurons in Output layer	3 neurons
	L2 penalty	0.0001
	Tolerance	0.0001
SVM	Kernel	Linear
	Penalty	1.0
	Tolerance	0.0001
RF	Criterion	Gini index
	Number of trees	10
	Sampling	bootstrap

The experimental results are depicted in the Fig.4-8. Each figure depicts the results according to a particular discretization of the soccer field. As shown in Fig.4-8 the models accuracy's convergence start from around 1500 cycles regardless the grid size. When the number of cycles is sufficiently large, the accuracy rates of each classifier are similar. Regarding NN and SVM, the accuracy rate is proportional to the number of grids, and it is high even for short amount of time spent to analyze the opponent formation. From this results, we can state that the larger the grid is, the sooner the opponent formation can be identified.

On the other hand, with RF, when the number of cycles is small, the accuracy rate is low compared with NN and SVM.

6 Conclusion

In this paper, we proposed a method to identify opponent strategies online according to their player formations. The proposed method decides whether our current strategy should be changed or not by using opponent position information. If the opponent strategy can be figured out, it is possible to employ a more suited strategy to increase the chance to outperform the opponent. For this purpose, opponent formation was classified, and three identification models were constructed by using machine learning methods. Future works include surveying the classification rate for unknown teams. We will also investigate which strategy is effective against opponent teams.

References

1. Thomas Gabel, Martin Riedmiller and Florian Trost. A Case Study on Improving Defense Behavior in Soccer Simulation 2D: The NeuroHassle Approach. *Proc. of the 12th RoboCup Symposium*, pages 61–72, 2008.
2. Jordan Henrio, Thomas Henn, Tomoharu Nakashima, and Hidehisa Akiyama. Selecting the Best Player Formation for Corner-Kick Situations Based on Bayes' Estimation. *Proc. of the 20th RoboCup Symposium*, 12 pages, 2016.
3. Luiz A. Celiberto Jr., Carlos H. C. Ribeiro, Anna Helena Reali Costa, and Reinaldo A. C. Bianchi. Heuristic Reinforcement Learning applied to RoboCup Simulation Agents. *Proc. of the 11th RoboCup Symposium*, pages 220–227, 2007.
4. Aijun Bai, Feng Wu, and Xiaoping Chen. Towards a Principled Solution to Simulated Robot Soccer. *Proc. of the 16th RoboCup Symposium*, pages 141–153, 2012.
5. Gregory Kuhlmann, Peter Stone, and Justin Lallinger. The UT Austin Villa 2003 Champion Simulator Coach: A Machine Learning Approach. *RoboCup 2004: RoboCup 2004: Robot Soccer World Cup VIII*, pages 636–644, 2005.
6. Ramin Fathzadeh, Vahid Mokhtari, Morteza Mousakhani, and Alizera Mohammad Shahri. Coaching with Expert System Towards RoboCup Soccer Coach Simulation. *Proc. of the 10th RoboCup Symposium*, pages 488–495, 2006.
7. Mazda Ahmadi, Abolfazl Keighobadi Lamjiri, Mayssam M. Nevisi, Jafar Habibi, and Kambiz Badie. Using a Two-Layered Case-Based Reasoning for Prediction in Soccer Coach. *Proc. of the International Conference on Machine Learning; Models, Technologies and Applications. MLMTA'03*, pages 181–185, 2003.
8. Shokkofeh Pourmehr, and Chitra Dadkhahand. An Overview on Opponent Modeling in RoboCup Soccer Simulation 2D. *Proc. of the 15th RoboCup Symposium*, pages 402–414, 2011.
9. Ubbo Visser, Christian Drücker, Sebastian Hübner, Esko Schmidt, and Hans-Georg Weland. Recognizing formations in opponent teams. *RoboCup 2000: Robot Soccer World Cup IV*, pages 391–396, 2001.
10. Luis Paulo Reis, Nuno Lau, and Eugnio Oliveira. Situation Based Strategic Positioning for Coordinating a Simulated RoboSoccer Team. *Balancing Reactivity and Social Deliberation in MAS*, Vol. 2103, pages 175–197, 2001.
11. Hidehisa Akiyama, and Itsuki Noda. Multi-Agent Positioning Mechanism in the Dynamic Environment. *Proc. of the 12th RoboCup Symposium*, pages 377–384, 2008.
12. Patrick Riley, and Manuela Veloso. On Behavior Classification in Adversarial Environments. *Proc. of the 5th Distributed Autonomous Robotic Systems*, pages 371–380, 2000.
13. Brgida Mnica Faria, Lus Paulo Reis, Nuno Lau, and Gladys Castillo. Machine Learning Algorithms applied to the Classification of Robotic Soccer Formations and Opponent Teams. *Proc. of the IEEE Cybernetics and Intelligent Systems*, pages 344–349, 2010.

14. Calinski, T., and J. Harabasz. A dendrite method for cluster analysis. *Communications in Statistics*, Vol. 3, No. 1, pages 1–27, 1974.
15. Hidehisa Akiyama, Tomoharu Nakashima, Jordan Henrio, Thomas Henn, Sho Tanaka, Tomonari Nakade, and Takuya Fukushima. HELIOS2016: Team Description Paper. *RoboCup2016 Leipzig, Germany*, Total 6 pages, 2016.
16. Rauf Khayami, Nader Zare, Maryam Karimi, Payman Mahor, Ardavan Afshar, Mohammad Sadegh Najafi, Mahsa Asadi, Fatemeh Tekrar, Ehsan Asali, Ashkan Keshavarzi. CYRUS 2D simulation team description paper 2014. *RoboCup2014 Joo Pessoa, Brazil*, Total 6 pages, 2014.
17. Siddharth Pritam " Infographics "team: Selecting Control Parameters via Maximal Fisher Information. *RoboCup2014 Joo Pessoa, Brazil*, Total 4 pages, 2014.
18. Mikhail Prokopenko, Peter Wang, Oliver Obst, and Victor Jauregui. Gliders2016: Integrating multi-agent approaches to tactical diversity. *RoboCup2016 Leipzig, Germany*, Total 6 pages, 2016.
19. Amir Darijani, Aria Mostaejeran, Mohammad Reza Jamali, Aref Sayareh, Mohammad Javad Salehi, Borna Barahimi. FURY 2D Simulation Team Description Paper 2016. *RoboCup2016 Leipzig, Germany*, Total 6 pages, 2016.
20. Mohammadreza Javan, Amirreza Ramezanzadeh, Mohammadjavad Kashi, SeyedEmad Mohamadi, Amirabbas Pashaeehir. HERMES: Soccer 2D Simulation Team Description Paper 2016. *RoboCup2016 Leipzig, Germany*, Total 6 pages, 2016.
21. Nima Nozari, Omid Hanifezade Masouleh, Sasan Jarah Fazel, Samaneh Nazari, Ahmad Eskandarzadeh, Fatemeh Jahandari, Milad Mofidiyan, Sajjad Parvin. MarliK 2016 Team Description Paper *RoboCup2016 Leipzig, Germany*, Total 5 pages, 2016.
22. Mohsen Sadeghi Pour, Azin Ghasemi, Kamyar Kalajooyeranjbar, Khashayar Kalajooyeranjbar, Mohammad Chaposhloo, Mohsen Firoozbakht. Ziziphus Team Description Paper RoboCup2016 *RoboCup2016 Leipzig, Germany*, Total 5 pages, 2016.
23. Mohammadreza Javan, Amirreza Ramezanzadeh, Mohammadjavad Kashi, SeyedEmad Mohamadi, Amirabbas Pashaeehir. FRA-UNIted — Team Description 2016. *RoboCup2016 Leipzig, Germany*, Total 6 pages, 2016.
24. Xiao Li, Rongya Chen, and Xiaoping Chen. WrightEagle 2D Soccer Simulation Team Description 2015. *RoboCup2015 Hefei, China*, Total 5 pages, 2015.
25. Kota Asai, Yuki Katsumata, Takuro Shibayama, Hiroki Nomura, Ryo Kondo, Hayato Tanaka, Koichi Uchinishi, Masataka Mizumoto, Tsubasa Fuzimitsu, Makoto Sei, Yuri Tani, Shiori Kubo, Yume Matsushita. RoboCup 2016 - 2D Soccer Simulation League Team Description Ri-one (Japan). *RoboCup2016 Leipzig, Germany*, Total 6 pages, 2016.

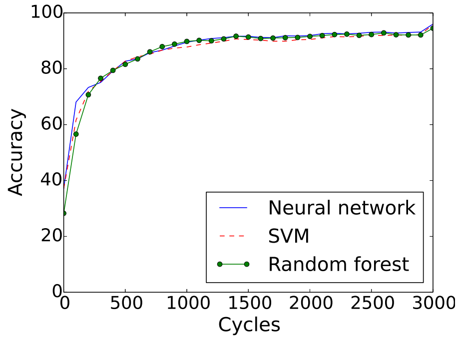

Fig. 4. Accuracy rates of the three models according to field discretized with a grid of size 6 × 4.

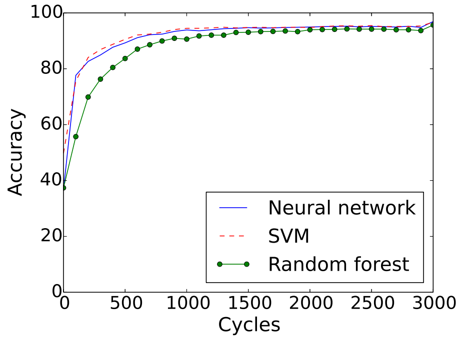

Fig. 5. Accuracy rates of the three models according to field discretized with a grid of size 12 × 8.

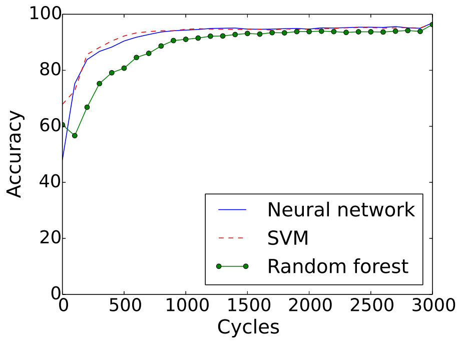

Fig. 6. Accuracy rates of the three models according to field discretized with a grid of size 15 × 10.

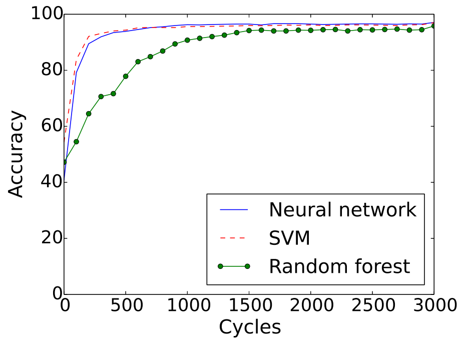

Fig. 7. Accuracy rates of the three models according to field discretized with a grid of size 24 × 16.

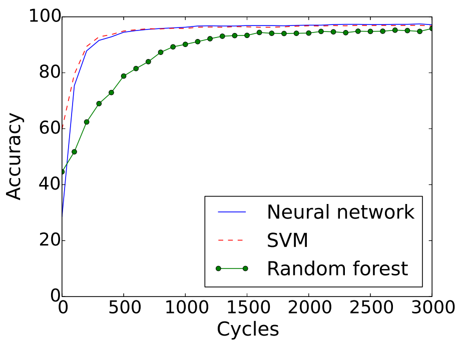

Fig. 8. Accuracy rates of the three models according to field discretized with a grid of size 30 × 20.