# *Wind farm water area path planning algorithm based on A* and reinforcement learning*

Tianqi Zha
*Wuhan University of Technology*
*Intelligent Transport Systems Research Centre*
Wuhan, China
zhatianqi@126.com

Lei Xie *
*Wuhan University of Technology*
*Intelligent Transport Systems Research Centre*
Wuhan, China
xielei@whut.edu.cn

Jiliang Chang
*Wuhan University of Technology*
*Intelligent Transport Systems Research Centre*
Wuhan, China
9212554662@qq.com

*Corresponding author

***Abstract*—In recent years, the scale of offshore wind farms is increasing because of the high efficiency and pollution-free wind power resources. However, the introduction of many facilities in the corresponding wind farm sea area has led to the increasing difficulty of ship navigation. Therefore, it is very important to plan safe and efficient driving path according to the corresponding starting and ending points for the navigation of ships in the increasing wind farm area. In this paper, a path planning algorithm based on the hybrid method of A* algorithm and reinforcement learning is proposed, which can plan an effective collision avoidance path for the sea area of wind farm. Then the method is used to simulate the ship's path planning in a wind farm, which proves the feasibility of the method. Finally, it shows that the method has universal reference significance for ship navigation in the wind farm waters.**

***Keywords—ship, plan, path, hybrid, algorithm***

## I. INTRODUCTION

The problem of whether the ship can safely pass through the target waters to reach the destination has always been widely concerned in the process of navigation. Correct path planning for the target waters plays an important role in solving this problem. Path planning can be regarded as a process of collision-free optimal path planning from the starting point to the target point, which integrates driving path and environmental constraints. Its essence is a constrained optimization problem, which is involved in the navigation and control of agents.[1-3] In order to solve the problem of ship's path planning in water navigation, many related methods have been proposed by scholars. Common path planning algorithms include Dijkstra algorithm[4], A* algorithm, artificial potential field method[5], grid method[6], ant colony algorithm[7], genetic algorithm[8] and so on. A* algorithm is very effective in solving the shortest path in static road network by constructing heuristic function. [9] It can plan the route with arbitrary shape obstacles[10] and theoretically guarantee the convergence of the global optimal solution[11,12].

However, in some application scenarios, the general A* algorithm may generate many nodes, take up a large amount of computing time, and be inefficient. Other scholars have improved the A* algorithm and achieved some results. Meng Zhongjie et al. proposed a sparse A* algorithm with variable step size, which effectively solved the contradiction between the extended step size, search time and planning results, and had good stability in the simulation results.[13] Tian Kuo et al. improved the A* algorithm by expanding the number of search neighborhoods and introducing the minimum "bend" estimation cost function to plan a smoother and better trajectory.[14] By analyzing and using obstacle information, Xiao Zibing et al. formulated a reasonable node expansion strategy and introduced a positive route guidance mechanism, which effectively improved the efficiency of path search. [15] In order to further improve the A* algorithm and solve the problem of path planning in specific scenarios, this paper takes the wind farm area as an example, and proposes a hybrid path planning method based on A* algorithm and reinforcement learning algorithm, which can search safe and efficient routes in complex environment of wind farm waters and ensure the safety of navigation of ships.

## II. A* ALGORITHM FOR WIND FARM WATERS

The complex environmental conditions in the water area of the wind farm pose a great threat to the safety of all kinds of ships sailing in it. Therefore, an effective method is needed to explore the path and complete the path planning to guide ships sailing in this special sea area. With the great improvement of computer computing ability in recent years, a large number of computing bottlenecks have been overcome, and various intelligent algorithms have emerged, enriching the path planning algorithm. For example, greedy algorithm, heuristic algorithm, ergodic algorithm and other types of algorithms are constantly applied in path planning. Generally speaking, greedy algorithm always makes the best choice in the current situation and ignores the global effect; ergodic algorithm is inefficient because of the huge amount of computation, and it takes a long time to get the results. For the environment of sailing ships in wind farm waters, greedy algorithm is easy to ignore the complexity of the waters and cause potential safety hazards, while the vast waters environment brings tremendous computational pressure on the application of ergodic algorithm. Therefore, greedy algorithm and ergodic algorithm are inappropriate as path planning algorithm for wind farm waters.

Heuristic A* algorithm is to explore the feasible path in the space, so that the required path can be found quickly. This method does not traverse the environment, greatly simplifying the calculation. Moreover, because this method is an exploration of the environment, it does not ignore the situation of the water area. Therefore, it is appropriate to choose the method related to the heuristic A* algorithm for the path planning of ships sailing in the water area of the

The research was supported by the National Key Technologies Research & Development Program (Grant No. 2017YFC0804900, 2017YFC0804904), the National Science Foundation of China (Grant No. 51679181).



978-1-7281-0489-8/19/$31.00 ©2019 IEEE

wind farm.

### A. Basic Principles of A* Algorithms

A* algorithm is a heuristic algorithm which can search effectively. When using this method for path planning, the environment is preprocessed by the grid method, and the planning area is divided into grid points. Select the current grid location as the starting point, the required grid location as the end point, and prepare two empty lists for storage: one is to record the current location around the environmental value of open list, so that each step can select the best path location in this list; the other is to record the final planning path of closed list. In the n-step planning process of A* algorithm, the current location is added to the closed list, and then the adjacent accessible grid location is added to the open list. The calculation and value F(n) can be expressed as:

$$F(n) = G(n) + H(n) \tag{1}$$

In the formula above, G (n) denotes the movement of the current step from the starting point, and H (n) denotes the estimated distance of the current step from the end point. F(n) is obtained from the sum of current movement quantity G(n) and the estimated distance from the current grid point to the target value H(n) of each grid point in the open list, and an optimal grid point is selected and moved according to the sum value until the end point is reached. The closed list storing the final path is obtained through A* algorithm, and the searched path is obtained by tracing the stored planning points from the closed list. For the small-scale regional path planning problem, the solution process is fast and concise, and the target result with high efficiency can be obtained.

### B. Improvement of A* algorithm for wind farm waters

Wind farm covers a large area of water. In this given environment, using traditional A* algorithm for path planning of starting and ending points will lead to a sharp increase in computation and reduce planning efficiency. For wind farm waters requiring large-scale path planning, this paper proposes an improved method of traditional A* algorithm, which combines reinforcement learning method to carry out a hybrid algorithm for path planning of wind farm covered waters, and improves the effect of traditional A* algorithm on path planning in this special waters. For wind farm waters requiring large-scale path planning, this paper proposes an improved method of traditional A* algorithm, which combines reinforcement learning method to carry out a hybrid algorithm for path planning of wind farm covered waters, and improves the effect of traditional A* algorithm on path planning in this special waters.

Considering the wide coverage of wind farm waters, using traversal algorithm to accurately plan the ship's trajectory in this special waters requires fine gridding of huge charts, so the number of grids generated is difficult to count. Although heuristic A* algorithm avoids traversing a large number of grid maps, it still needs to refine the path search according to the grid points of the map. The huge number of gird points that need to be processed in the path search process also limits the efficiency of the planning process. Therefore, the idea of A* algorithm to search for the optimal planning path in the refinement of raster points is changed. Change the map to rough rasterization operation to reduce the calculation dimension of A* algorithm when searching the path and improve the efficiency of path planning. Generally speaking, if there are obstacles in a grid point of rough grid, the information of obstacles in the grid will be generalized, and the nodes involved in obstacles will be set as impassable nodes. This method sacrifices the resolution of the map, so that the map information represented by each grid point has been greatly expanded. Using this method to guide ships sailing in wind farm waters can not guarantee its accuracy. Therefore, it is necessary to carry out reasonable rough gridding treatment combined with maps. When marking unreachable gridding points, attention should be paid to making rough gridding points not cover waterways, so as to avoid shielding possible planning paths.

When A* algorithm is used in the rough raster map improved according to the situation of the sea area, in order to ensure the safety of route planning in complex waters, the judgment part of variable moving step is added to improve the performance of collision avoidance in the process of route planning. A fixed step size can be determined in the planning process to efficiently obtain the required path in open and barrier-free waters. However, once approaching the complex waters with obstacle grid points, compared with the planning strategy of small step length, detour will be carried out in advance when moving with fixed stride. In this way, compared with the path with small step length to improve the resolution, the use of fixed step size strategy will result in a greater deviation in the local path selection for detour, so that the detour path length with long stride is necessarily longer than the obstacle avoidance path planned by small step size. This makes the path chosen with a fixed stride length move strategy worse than with a small step length move strategy.

A* algorithm with variable step size is selected for the sea area near obstacle grid points. The steps can be expressed as follows: The large step L used in the open sea area is taken as the original step, and the barrier grid points in the sea area ahead of the ship are detected at the same time; if there are obstacles, start the variable step size strategy, the step size is continuously reduced to half of the original step size until the safe step size is determined. The next step of path planning is carried out recursively until the barrier-free existence around the ship is detected.

A* algorithm is used to simulate the sea area with barrier grid points, as shown in Fig. 1. The left one is obtained by using variable step size strategy, while the right one is obtained by using fixed step size strategy, the yellow grid represents the starting point, the blue grid represents the termination point, the white grid represents the obstacle point, and the green grid represents the search path. Taking the grid edge length as the measurement reference, we can easily calculate the length of the two strategy search paths in the grid map. Therefore, by taking the side length of a grid as a reference, the path planned by the variable step size strategy is 7 grids in length, while the nodes searched by the fixed stride size strategy are reduced, but they pass 8.75 grids in length to avoid obstacles. It can be seen that the use of variable step strategy in obstacle waters increases the search node to some extent, but it can effectively improve the accuracy and efficiency of path planning.



978-1-7281-0489-8/19/$31.00 ©2019 IEEE

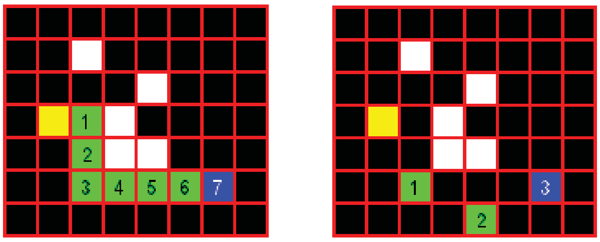



Fig. 1. Path planning diagram of obstacle area.

## III. Hybrid Path Planning Method

The method described above uses a small step A* algorithm to search the path in obstacle waters. In the original generalized grid map, if there are obstacles in the grid point, the grid point is set as the barrier grid point. In this case, even if the non-obstacle part of the grid point has the condition of ship passage, the whole grid point will be expanded to obstacle grid point because of the obstacle. The hidden path in the barrier grid point will be automatically ignored by using A* method. In order to find a more suitable planning path in obstacle waters, the reinforcement learning method is used to search the path near the obstacle grid point, which makes the mobile path of the obstacle grid point visible. Combining with A* algorithm, the hybrid method is used to plan the path, and finally a feasible path with local modification is obtained.

### A. *Reinforcement Learning Method*

Reinforcement learning refers to the method by which an agent continuously explores the behavior of obtaining the best reward through reward, so as to obtain the best strategy. In the standard reinforcement learning framework, besides the agent itself, there are four basic elements: strategy, reward and punishment function value, guidance function and environment model[16].

The reinforcement learning method can be naturally introduced into the path planning by taking the final goal point of the path as the final reward goal. The starting point, target point and obstacle are obtained as environmental models, and the reward function is used to update the optimal path needed for path planning in continuous exploration.

Reinforcement learning is a kind of algorithm. There are many specific algorithms in this kind of algorithm. For example, Q-learning, sarsa and other value-based algorithms, policy gradients and other direct selection behavior algorithms, as well as model-based reinforcement learning algorithm. The direct selection algorithm is usually inefficient, and the result may be the local optimal path. However, the model-based reinforcement learning algorithm has a large amount of work to model the water environment of the wind farm and the effect of the model is poor. Therefore, it is appropriate for us to choose a model-free and function-based reinforcement learning algorithm for path planning.

### B. *Introducing sarsa algorithm*

Q-learning algorithm is an offline algorithm in the model-free reinforcement learning algorithm based on function values. It chooses the direction of maximizing the reward value each time it updates, and then chooses the action again in the next state. This method is reckless and greedy. It is not suitable to apply it to the path planning in the water environment. Rummery and Niranjan proposed a model-based algorithm in 1994, initially called an improved Q-learning algorithm, later called Sarsa learning algorithm [17]. Sarsa algorithm is a relatively conservative algorithm, which is sensitive to obstacles and can effectively avoid adverse factors in the planning process, so that the security and reliability of the planning path can be guaranteed.

Q-learning algorithm uses the maximum value of the value function to iterate. The update of Q-value depends on various hypothetical actions (essentially a different strategy algorithm), while sarsa uses the actual Q-value to iterate. It updates the value function strictly according to the experience gained from implementing a strategy[18,19]. In Q-learning algorithm, the behavior selection strategy and the iteration of the value function of the learning system are independent of each other, while the sarsa learning algorithm realizes the iteration of the behavior value function in the form of strict time difference learning, that is, the behavior decision is consistent with the iteration of the value function.

The steps of the hybrid improved algorithm by introducing the sarsa algorithm are as follows:

- Large step L is used to plan the whole path, and the whole path with lower resolution is obtained.
- Change the big step to the small step near the obstacle node, and modify the local path.
- Sarsa algorithm is used to refine the search in obstacle nodes to get the available route hidden in obstacle grids.
- Comparing the obstacle avoidance part of local path with the available route hidden in barrier grid, the efficient navigation path is chosen as the final result of path planning.

## IV. Application of Mixed Method

Through the above analysis, it can be seen that the hybrid method proposed in this paper is reasonable to some extent, and it has certain reference value for the path planning of the increasing water area of wind farm. In order to verify whether this method can ensure that the shortest path with safety guarantee is planned for ships between the starting and stopping points in the wind farm water area, we can build a model of the sea area of the wind farm, carry out simulation of the path planning for the modeling sea area, and judge the advantages and disadvantages of this method through simulation results.

### A. *Rasterize*

As mentioned above, it is necessary to rasterize the map of the water environment of the wind farm before the path planning. When the traditional method is used to rasterize the wind farm water area map with the same grid size, it will cause many disadvantages because the grid points are too rough. However, if variable step size method is used in path planning, the errors caused by rough raster map with equal size division on path planning can be reduced. In order to simplify the operation of variable step-size planning of ships, the distance of passing through each raster point is taken as the basic step size of ship movement. As long as the raster points are continuously refined, the raster point size of the



978-1-7281-0489-8/19/$31.00 ©2019 IEEE

map can be changed. At the same time, the ship moves a grid point differently in the rough large grid point than in the refined small grid point, thus obtaining the prototype of variable step size movement.

In this paper, the idea of quadtree is used to rasterize the map in order to make the raster map achieve the effect of variable step size beside obstacles and generate the required refined raster. The basic idea of quadtree rasterization is to divide the rasterized map recursively into equal structures to form four equal rasterized spaces. This recursion continues until the grid reaches its maximum resolution or the grid points to be further subdivided are completely occupied by obstacles. The information of the final map space is stored on the leaf node, and when using the path planning algorithm, we only need to search the path in the leaf node. Rasterize the partial water area map near the wind farm as described above, so as to obtain the raster map with variable step-size pathfinding based on raster points, as shown in Fig2.

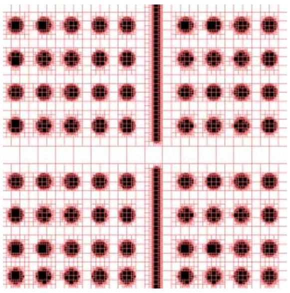


Fig. 2. Rasterized model of wind farm water area with variable step size.

In the figure, the black circle represents the fan, the black strip represents the bridge built in the water area, the white area represents the water area for the required path planning, and the numerous red grid areas of different sizes indicate that the map area for the required path planning has been rasterized as described above.

### B. *Global path planning*

A* algorithm is used for path planning for the same starting and ending points in the areas of rough rasterized map and fine rasterized map respectively. Simulation results are shown in Fig3. It can be observed that there are significant differences between the path planned by the left striding length A* algorithm and the right variable step size A* algorithm.

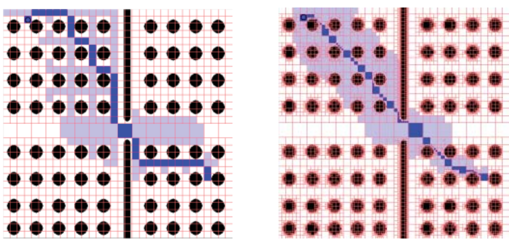


Fig. 3. Simulation diagram of overall path planning.

For the rough raster map, there must be raster points with obstacles. At this time, for the sake of safety, the raster points with obstacles must be abandoned when using the A* algorithm to find the road, which will inevitably cause the consumption of roads. In order to exactly compare the different paths planned by the two methods, the related parameters of the above two planning paths are compared in the list, as shown in table 1.

TABLE I. PATH PARAMETERS PLANNED BY DIFFERENT METHODS

| **Step Length Selection** | ***The Path Cost(m)*** | ***Number of path nodes*** |
|---|---|---|
| Large step | 6809 | 38 |
| Variable step size | 5654 | 71 |

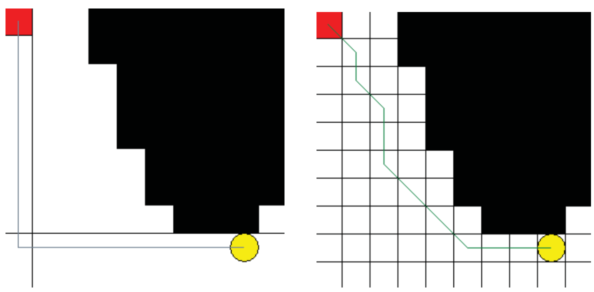
As can be seen from the above table, compared with the variable step size A* algorithm, the large stride length A* algorithm has less path nodes but more paths. In this way, if the ship uses the large stride length A* method, the planning path obtained will inevitably reduce the sailing efficiency of the ship between the starting and stopping points. In order to obtain A planning path with more heading efficiency in wind farm waters without collision, A* algorithm of variable step strategy is advantageous. In order to obtain a more efficient planning path in the field of wind farm without collision, A* algorithm using variable step size strategy is advantageous.

### C. *Local correction*

For the raster map, it is inevitable that the raster points reaching the highest resolution are still mixed raster points of obstacles and water environment. In this case, when using A* algorithm for path planning, the possibility paths in obstacle raster points will be ignored, which will be corrected by using reinforcement learning method.

Assuming that the maximum resolution raster points are eight times the length of the ship, the most detailed raster points can still contain a lot of map information. If safe detour is conducted around the most refined obstacle grid points according to A* algorithm, navigable channel information hidden in the grid points is inevitably ignored. However, if all maps are divided more finely, it will increase the calculation and affect the planning efficiency. In this way, the semi-obstacle raster points around the overall planning path can be divided into more detailed raster points, and sarsa reinforcement learning algorithm can be used to learn in the further refined raster areas. Finally, the obtained strategy is compared with the algorithm planned by A* algorithm to obtain the optimal path.

Path correction is carried out for the waters near an obstacle, as shown in Fig. 4. The black grid area represents unnavigable waters, the white grid area represents navigable waters, the red grid area represents the starting point, and the black grid area represents the ending point. The left figure is the overall planning path obtained by using A* algorithm in the general raster map. In the right figure, the waters around the mixed raster points are further divided, and the revised strategy and planning path are obtained by using reinforcement learning method in the new refined raster map.





978-1-7281-0489-8/19/$31.00 ©2019 IEEE

Fig. 4. Overall planning and local correction contrast path diagram.

As assumed above, the right side length of each grid in Fig. 4 represents the size of the planned ship, which is used as the metric. It can be easily seen from the simulation situation near obstacles in the water area of the wind farm that both methods can reach the destination without collision. However, the overall path planning method moves 16 grid lengths around the mixed grid points, but it only takes 13.07 grids to reach the target grid points after local correction with reinforcement learning method. It can be clearly seen from the moving distance that this correction method has a further optimization effect on the planned path in the water area of the wind farm.

Obviously, the hybrid path planning method based on A* and reinforcement learning proposed in this paper can finally obtain a feasible path planning scheme in wind farm waters with obstacles, which indicates that the hybrid method proposed in this paper provides a feasible solution to the problem of wind farm waters path planning.

## V. Conclusion

In this paper, a hybrid method based on A* and reinforcement learning algorithm is proposed for path planning of wind farm waters. Through analysis and simulation, it can be seen that this method has certain reference value for ship navigation in wind farm waters. In the future research, the method provided in this paper can further improve the accuracy and efficiency, enhance the interaction with the environment, make it more real-time and accurate path search while having the real-time planning ability in dynamic environment, which still needs further discussion.

## Acknowledgment

On the completion of this paper, I would like to express my sincerest thanks to the teachers and classmates who have helped me. Everyone's perseverance in scientific research and realistic working attitude have always inspired me to constantly learn knowledge and pursue the realm of life. At the same time, thank the National Key Technologies R&D Program(2017YFC08049002017YFC0804904), so that I can complete this paper. I sincerely thank all the people who care and help me. Thank you.

## References

[1] Zhou Mingxiu, Cheng Ke, Wang Zhengxia. Improved ant colony algorithm in dynamic path planning. Computer Science, 2013, 40(1): 314-316.

[2] Luo Yuan,Shao Shuai Zhang Yi.Location. path planning of mobile robots based on information fusion. Computer Applications, 2010, 30(11): 3091-3093.

[3] Zhang S,Deng W,Zhao Q,et al.Dynamic trajectory planning for vehicle autonomous driving[C]. Pittsburgh:IEEE International Conference on Systems, Man and Cybernetics,IEEE Computer Society,2013.K. Elissa, "Title of paper if known," unpublished.

[4] Kang Wenxiong, Xu Yaozhao. Hierarchical Dijkstra algorithm for node constrained shortest path. Journal of south China university of technology (natural science edition), 2017, 45 (01): 66-73.

[5] Liu Yanju, Dai Tao, Song Jianhui. Improved path planning algorithm of artificial potential field methodResearch. Journal of Shenyang University of Technology, 2017, 36 (01): 61-65+76.

[6] Zhu Aibin, Liu Yangyang, He Dayong, He Shengli. Potential field grid method for solving local minimal problems in path planning. Mechanical design and research, 2017, 33 (05): 46-50.

[7] Wang Zhizhong. Path planning of mobile robots based on improved ant colony algorithm. Mechanical design and manufacturing, 2018, (1): 242-244.

[8] Liu Erhui, Yao Xifan. Path Planning and Implementation Platform of AGV Based on Improved Genetic Algorithms. Computer Integrated Manufacturing System, 2017, 23(03): 465-472.

[9] Woo-Jin Seo, Seung-Ho Ok, Jin-Ho Ahn, Sungho Kang, Byungin Moon. An efficient hardware architecture of the A- star algorithm for the shortest path search engine. 2009 Fifth International Joint Conference on INC, IMS and IDC. Piscataway: IEEE Computer Society, 2009: 1499- 1502.

[10] Wei Ruixuan, Xu Zhuofan, Wang Shulei, Lu Minghai. A self-optimizing A-Star UAV route planning algorithm based on Laguerre diagram. Systems Engineering and Electronic Technology, 2015, 37 (03): 577-582.

[11] Li Ji, Sun Xiuxia. Research on UAV path planning algorithm based on improved A-Star algorithm. Journal of Military Engineering, 2008 (07): 788-792.

[12] Qi Wei, Li Xia, Cai Wanyong, Lu Qianhong. Position selection and route planning of early warning aircraft in air attack. Firepower and command and control, 2016, 41 (12): 64-68.

[13] Meng Zhongjie, Huang Panfeng and Yan Jie. Track planning technology of hypersonic vehicle based on improved sparse A~* algorithm. Journal of Northwest Polytechnic University, 2010, 28 (02): 182-186.

[14] Tian Kuo, Liu Xu. Unmanned aerial vehicle dynamic path planning based on multi-strategy SSO and improved A* algorithm. Electro-optic and control, 2017, 24 (11): 31-37.

[15] Xiao Zibing, Qu Yaohong, Yuan Dongli. Efficient A-Star route planning algorithm based on obstacle information. Firepower and command control, 2018, 43 (09): 71-75..

[16] Yang Yi, Pang Yongjie, Li Hongwei, Zhang Rubo. Research on USV Local Path Planning Method Based on Reinforcement Learning under Complex Sea Conditions (English). Journal of Marine Science and Application, 2014, 13 (03): 333-339.

[17] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 2012, 25(2).

[18] Boser B E, Guyon I M, Vapnik V N. A training algorithm for optimal margin classifier. Proceedings of Annual Acm Workshop on Computational Learning Theory, 1996: 144-152.

[19] Donahue J, Jia Y, Vinyals O, et al. DeCAF:A deep convolutional activation feature for generic visual recognition. Computer Science, 2013, 50(1): 815-830.



978-1-7281-0489-8/19/$31.00 ©2019 IEEE