A Moving Object Detection and Predictive Control Algorithm Based on Deep Learning

Yan He¹, Xiuxian Li*¹ and Hongfei Nie²
1 School of AI, Shanghai Normal University Tianhua College, No. 1661 North Shengxin Road, China
2 Application Engineering Department of Special Products, Shanghai Microelectronics Equipment Group Co., Ltd., China
Email: heyan_886@163.com

Abstract. In traffic scene, a low detection rate of dynamic target is caused by interference features of background area and fast speed of detected moving target. In this paper, an optimal target detection and prediction algorithm is proposed. Firstly, the algorithm of real-time motion parameters (speed, direction, etc.) detection of moving objects (vehicles) based on the depth theory is studied. The original prediction problem is transformed into the problem of automatic updating rate of uncertain parameters and the problem of minimizing the maximum path distance. Secondly, by on-line estimation of the automatic update rate, the proposed trajectory generated is guaranteed to be the minimum length. The allocation strategy minimizes the maximum distance traveled in the collision free. Then, the stability of the estimation error system is guaranteed base on deep learning algorithm. Finally, the simulation shows that the proposed algorithm can be used for moving target detection and path prediction accurately. The performance of the control system is improved.

Keywords: Obstacle avoidance; Optimal path planning; Connectivity; Trajectory tracking

1. Introduction
The frequent traffic accidents are mainly caused by unavoidable human factors. The traffic accident rate will be greatly reduced by the intelligent monitoring and response capabilities of self-driving cars. For driverless car in complex traffic problems, it is necessary to solve the problem of accurately reaching the optimal trajectory and tracking control of unmarked target positions in a chaotic environment. The car based on the driverless technology [1-3] is essentially a mobile intelligent networked robot. The real intelligence and sharing are realized by path prediction and obstacle avoidance detection method.

The most straightforward idea for the detection of moving objects is to analyze the motion field of each point in the image sequence, that is, to find the motion of the corresponding point on the image plane caused by the space motion. However, what is measurable in the image is only the changes in image irradiance, and the motion reflected by the changes in image irradiance is called the apparent motion. Existing target detection and prediction methods include the algorithm based on space-time gradient [4], methods based on correlation [5], methods based on frequency domain analysis [6], and methods based on graph theory [7]. Literature [8] studies the three-level coordinators of multi-driverless cars. Literature [9] uses the concept of sub-dimensional expansion to extend only the state space to high-dimensional space when necessary. There are also some reactive controllers that can control a large number of driverless cars, but do not plan a complete trajectory and cannot solve
the local minimum. The speed of the moving target can be calculated accurately by using the method of tracking based on frequency domain analysis [6]. But this method adopts an iterative method, the calculation time is long, and the real-time tracking cannot be performed, and the method is greatly affected by noise. So it is more suitable for situations where the image noise is relatively small and the target motion speed is not large. Literature [10] uses a probabilistic roadmap and the random priority generation trajectory. It uses a cost function to reflect the maximum time spent on a driverless car or the maximum distance traveled by a driverless car. When multiple driverless cars must have quick access to the first response and search, the optimal trajectory planning for multiple driverless cars that avoid collisions can be performed by the traditional path planner by planning in the joint state space. But the computational complexity usually increases exponentially with the number of driverless cars. To solve this situation, one way is to separate the path plan from the speed of each path. The method based on the graph theory [7], after the motion field is determined, the random noise and some too small motions are removed. It is considered that those regions whose motion vectors are always consistent within a certain range during the test period belong to an object. So each moving target (vehicle)'s motion parameters (speed, direction, etc.) at various times can be determined. Literature [11] studies the research which is based on distance on formation control motion parameters and uses the method based on displacement to control the direction of the car. It also finds the best path and the problem assigned to the target, and minimizing the maximum distance of the car. This proves that the existence of the effective trajectory is known.

This paper aims at the problem of low dynamic target detection rate caused by excessive background area interference characteristics and fast moving speed of detected targets in real traffic scenes. It proposes the optimal target detection and prediction algorithm, and studies the real-time motion parameters (speed, direction, etc.) detection algorithm of moving target (vehicle) based on depth science. It deduces the system error model for the controller design which is suitable for searching the allocation of trajectory and target position at the same time and converts the original control problem into the problem of automatic updating rate of uncertain parameters and the problem of minimizing the maximum distance of path. Secondly, by estimating the automatic update rate online, the trajectory generated by the proposed algorithm is guaranteed to be the minimum length and no collision, and the potential field function is minimized. Then, through the deep learning algorithm, the asymptotic stability of the closed-loop tracking error system and the stability of the estimation error system are guaranteed, and the computational complexity is reduced. Finally, simulations and experiments show that the proposed algorithm can be accurately used for moving target detection and path prediction, and the performance of the control system is improved.

The structure of this paper is as follows, and the concrete problem is described in the first part. The optimal trajectory tracking control algorithm is detailed in the second part. The third part is the simulation result of the algorithm proposed in this paper. The fourth part summarizes the algorithm is proposed in this paper.

2. Deep Convolutional Anti-Generation Network Model

The detection of vehicle moving targets in actual traffic scenes is a difficult problem in the field of driverless car. The research on optimal target detection and prediction algorithms is required because of the problem of low dynamic target detection rate caused by excessive interference characteristics in the background area and fast moving speed of detected objects. First, a three-dimensional model of the target (vehicle) needs to be built, and then the model is projected on a two-dimensional plane, and matched in the image to find the three-dimensional information of the target. Then, according to the change of the posture (position and direction) of the target in each frame image, the parameters (speed, direction, rotation angle, etc.) of the motion model are calculated by an optimized method. The generative confrontation network is a whole architecture. The generation model and the discriminant model can adopt various deep neural networks. The only difference is that it contains two networks. The deep convolutional confrontation generation network model is realized by using the architecture of the traditional convolutional neural network for both the generation model and the discriminant
model. As shown in figure 1, it describe that how to use the convolutional neural network to realize the data’s input-output in the generation model and the discriminant model respectively.

![Deep convolutional anti-network structure.](image)

**Figure 1.** Deep convolutional anti-network structure.

### 2.1. Generate Model Construction
The graph theory is used to indicate the positional relationship between the driverless cars. The position of the driverless car represents the apex of the figure, and the interaction between the driverless cars represents the side of the figure. Therefore, the connected undirected diagram of the driverless car can be expressed as:

$$G = (\partial, \varepsilon)$$  \hspace{1cm} (1)

\(\partial\) is the apex of the graph, that is, the position of the driverless car, and \(\varepsilon\) is the set of edges. Network connectivity utilizes a symmetric adjacency matrix \(A(t)\) composed of time-varying element \(a_{ij}\). In the case of an undirected graph, if \(a_{ij} = a_{ji}\), \(\varepsilon = \{(i, j) | a_{ij} > 0\}\). Since the undirected graph has no self-loop, \((i, j) \in \varepsilon\) means \(i \neq j\). The initial matrix \(A(0)\) can be defined as:

$$a_{ij}(0)\begin{cases} 1, & \|x_{ij}(0)\| \leq R \\ 0, & \text{else} \end{cases}$$ \hspace{1cm} (2)

When \(t > 0\), the dynamic definition of the time-varying matrix is:

$$a_{ij}(0)\begin{cases} 1, & \|x_{ij}(0)\| \leq R \text{ and } a_{ij}(t^-) = 1 \\ 0, & \text{else} \end{cases}$$ \hspace{1cm} (3)

The \(t^-\) is expressed as a moment before time \(t\).

### 2.2. Discriminant Model Construction
The input of the discriminant model is the natural data and the output data of the generated model, and the output is the corresponding class 1 and 0. This part is still implemented by the improved classical convolutional neural network model, all pooling layers are replaced by convolution operations, and all the non-linear functions used by the hidden layer are corrected to Linear unit.

### 3. Prediction Algorithm

#### 3.1. Speed Calculation
Since the position of each target is calculated to ensure integrity is displayed in multiple driverless car settings. The distance from the initial position of the driverless car \(i\) to the target \(j\) position can be defined as:
Where $\| \cdot \|$ is the norm $\ell$, and in the time period $\Delta t$, the minimum acceleration $\mu_{ij}$ of the initial position of the driverless car $i$ to the target $j$ position can be expressed as:

$$
\mu_{ij} = \arg \min_{\mu_{ij}} \left( \frac{2 (L_{ij} - (v_i - v_j) \Delta t)}{\Delta t^2} \right) \text{ s.t. } 0 \leq \| a_i \| \leq R
$$

(5)

The speed curve of the driverless car $i$ is parameterized by a monotonically increasing portion of the trajectory completed at time $t$, where $\theta_i(t) \in [0,1]$. These speed curves need to be calculated for each driverless car, with each driverless car seeking to minimize the time it takes to reach its designated $X$ target while avoiding collisions with driver cars with higher priority. At time $t$, the set of predicted positions of driverless car $i$ is expressed as:

$$
K_i(t) = \bigcup_{\eta \in [0,1]} \mu_{ij} (\eta) \oplus D_k
$$

(6)

Where $D_k$ represents the range of line of sight for each driverless car with R radius, therefore, the speed curve is calculated by minimizing the time that the driverless car reaches its specified target, since the function requires each of the calculated Driverless cars to have higher priority motion trajectories, so the speed curve of the driverless car is calculated in the order in which the sequence $\theta_i(t)$ appears.

### 3.2. Avoid collision Prediction Position Matching Potential Function Generation Function

It is assumed that the minimum distance between the driverless cars is allowed, that is, the driverless car $i$ cannot enter the inner boundary of the driverless car $j$. The two circular areas around each driverless car have a radius of $r_{in}$ and $r_{out}$ ($0 < r_{in} < r_{out}$) respectively, assuming $r_{ij} \leq \min \| a_i \|$ for any driverless cars $i$ and $j$. The collision avoidance area is defined by these two circles as shown in figure 2, and the driverless cars $i$ and $j$ are a pair of driverless cars.

![Figure 2. Schematic diagram of the collision avoidance area.](image)

In order to achieve the separation of driverless vehicles with different detection targets, they gather to their own detection day signs. It is necessary to establish a potential field function with different detection targets. The potential field function requires driverless cars with different targets to move away from each other until there is no interaction between them. When designing the collision avoidance controller, the potential function is satisfied such that the control input becomes zero when $r_{out} \leq \| a_j \|$. So the new potential energy function $\psi_{col}(a_{ij})$ with a finite truncation $r_{out}$ is expressed as:

$$
\psi_{col}(a_{ij}) = \min \| a_{ij} \|
$$
\[
\psi_{\text{col}}(a_y) = \begin{cases} 
\int_{a_y}^{a_y} \psi_{\text{col}}(\alpha) d\alpha, & a_y \in [r_{in}, r_{out}) \\
0 & \text{else} 
\end{cases}
\] (7)

Among them, \( \psi_{\text{col}}(a_y) \) are strictly reduced, and its maximum value is \( \phi_{\text{col}}(u) \) at \( a_y = r_{in} \phi_{\text{col}} \) is expressed as:

\[
\phi_{\text{col}} = -\frac{\|a_y\|}{\|x_r - x_o\|^2} 
\] (8)

Therefore, the collision avoidance control \( F_{i}^{\text{col}} \) of the \( i \)th driverless car is:

\[
F_{i}^{\text{col}} = -\sum_{j \in N} \nabla x_i \psi_{\text{col}}(a_y) 
\] (9)

3.3. Maintain Connectivity

Taking the position error as a consensus variable, \( w_i = x_i - y_i^{*} \), * is the conjugate, so that it works between the two driverless cars \( i \) and \( j \) in the interval \( \|x_j\| \in (r_{out}, R) \), instead of \( x_i \) to maintain connectivity. Therefore, the connectivity constraints of the error term are changed, and the upper and lower bounds of \( \|w_j - w_i\| \) can be obtained using the forward and inverse triangular inequalities:

\[
\|w_j - w_i\| \leq \|x_i - x_j\| + \|y_j^{*} - y_i^{*}\| 
\] (10)

\[
\|w_j - w_i\| \geq \text{abs} \left( \|x_i - x_j\| - \|y_j^{*} - y_i^{*}\| \right) 
\] (11)

When \( \|x_i - x_j\| = R \), \( \text{abs} \left( R - \|y_j^{*} - y_i^{*}\| \right) \leq \|w_j - w_i\| \leq R + \|y_j^{*} - y_i^{*}\| \), \( \text{abs} \) means to go to absolute value.

As can be seen from equation (21), \( \|w_j\| \in (r_{out}^{w}, R^{w}) \) where \( r_{out}^{w} = \text{abs}(r_{out} - \|y_j^{*}\|) \) and \( R^{w} = \text{abs}(R - \|y_j^{*}\|) \).

The potential connectivity function \( \psi_{\text{conn}}(a_y) \) preserved by a finitely truncated connectivity under \( R^{w} \) is expressed as:

\[
\psi_{\text{conn}}(a_y) = \begin{cases} 
\int_{a_y}^{a_y} \psi_{\text{conn}}(\alpha) d\alpha, & a_y \in [r_{out}^{w}, R^{w}) \\
0 & \text{else} 
\end{cases}
\] (12)

Where \( \psi_{\text{conn}}(a_y) \) is a strict increase and the minimum value reached by \( \psi_{\text{conn}}(a_y) \) at \( a_y = r_{out}^{w} \):

\[
\phi_{\text{conn}} = \frac{\|y_j^{*}\|}{\|x_r - R^{w}\|^2 + \frac{1}{Q_{\text{conn}}}} 
\] (13)

Therefore, the connected function \( F_{i}^{\text{conn}} \) can be derived as:

\[
F_{i}^{\text{conn}} = -\sum_{j \in N} \nabla w_i \psi_{\text{conn}}(a_y) 
\] (14)
Where \( N_i/N^-_i = \{ j : j \in N_i \text{ and } j \not\in N^-_i \} \).

Limited potential functionality has been designed to avoid collisions and connection preservation, which means that the control law is bounded. It can be known from equations (19) and (24) that the algorithm control law is expressed:

\[
K_i(t) = -\sum_{j \in N} \nabla x_j \psi_{col} (a_j) - \sum_{j \in N} \nabla w_j \psi_{conn} (a_j)
- \sum_{j \in N} (w_j - w_j^t) - \sum_{j \in N} (v_i - v_j^t) - (w - x^t),
\]

Where \( x \) and \( y \) are the position and velocity of the desired trajectory, respectively. For all \( i \) when \( t \to \infty \), then \( v_i \to v_i^t \) and \( x_i \to x_i^t \), thus forming the desired trajectory afterwards.

4. Simulation

Among the deep learning classification tasks and target detection tasks, the network the Convolutional Neural Network (CNN) is widely used. The 2D scene image library in the KITTI dataset is used as the training set of the network. The main founders of the KITTI dataset are the Karlsruhe Institute of Technology and the American Institute of Technology. They are also the internationally available computer vision algorithm evaluation dataset for autonomous driving scenarios. The test set contains a total of 7481 training images and 7581 test images, with a total of 80256 targets. The selected scenes are divided into urban, rural and speed highways, and the scene is more complete. Each image contains up to 15 cars and 30 pedestrians. In order to improve generalization, the target in the photo contains different levels of occlusion and truncation. The CNN convolutional layer linearly combines the characteristics of the previous layer (speed \( v_i \), acceleration \( \mu_i \), velocity curve \( \theta_i(t) \), collision avoidance control \( F^\text{col}_i \), connectivity function \( F^\text{conn}_i \), algorithm control law \( K_i(t) \)), and the cooperative control input characteristics based on communication topology (speed \( v_i \), acceleration \( \mu_i \), velocity curve \( \theta_i(t) \), collision avoidance control \( F^\text{col}_i \), connectivity function \( F^\text{conn}_i \), algorithm control law \( K_i(t) \)) is established under the condition of given undirected graph \( G \). It enable multiple unmanned systems to meet target detection, collision avoidance control targets and formation control targets, and then nonlinear activation Based on the above theoretical basis, this paper uses a micro-neural network to replace the ordinary convolution process. The micro-neural network is weight-shared for the same feature layer. This improvement can make the neural network learn more complex and useful polymorphic combination features. At the same time, the structure of RCNN target detection is used. In general, Selective Search is used to extract candidate regions. This method takes about 0.5 frames / s, which is not conducive to real-time monitoring. Therefore, the target candidate region is extracted by using BING to lower the detection time target to 3.00 frames / s.

Consider a system of six driverless cars, and the desired formation is a planar regular hexagon, where \( R = 12m \), \( r_m = 4m \), \( r_{out} = 5m \). Initially, driverless cars were randomly placed in an area of \( 10m \times 10m \), and the speed of the car was between 0 and \( 2m/s \). Therefore, it is assumed that the initial errors of the relative positions of two adjacent cars are bounded to ensure that their initial distance remains constant and limited to generate a connected graph of the specified communication range, the initial distance should be greater than \( r_m \). Since the relative speed selection between driverless cars is limited, the difference between the initial speed and the desired speed is limited for all driverless cars.

As figure 3, the driverless car can be formed from the initial position and held at the desired speed.
The minimum and maximum values of the interaction distance between driverless vehicles are plotted in figure 3. The minimum and maximum interaction distances are greater than 4 and less than 12 meters. It confirms collision avoidance and connectivity while ensuring any two adjacent driverless cars’ interaction distance converges to 6 meters, maintaining the stability of the system.

(a) Minimum interactive distance between driverless cars.
(b) Maximum interactive distance between driverless cars.

**Figure 3.** Minimum and maximum interaction distance between driverless cars.

The consensus result of the position error is also proved by plotting the possibility of inconsistent positional errors in the x and y directions in figure 4. The possibility of inconsistent positional errors also tends to zero, thus confirming the positional error to reach a consensus.

(a) Possibility of inconsistent position error in the x-axis direction.
(b) Possibility of inconsistent position error in the y-axis direction.

**Figure 4.** Possibility of position error inconsistency in the X-axis direction and the Y-axis direction of the driverless vehicle.

Figure 5 shows that with the increase of the number of driverless cars, the algorithm in this paper has a low computational complexity.
Figure 5. Comparison of computational complexity of different algorithms.

5. Conclusion
This paper aims at the problem of low dynamic target detection rate caused by excessive background area interference characteristics and fast moving speed of detected targets in real traffic scenes. A moving object detection and predictive control algorithm is proposed in this paper. Firstly, the real-time motion parameter detection algorithm based on depth learning for moving objects (vehicles) is studied, and the original prediction problem is transformed into the problem of automatic update rate of uncertain parameters and minimizing the maximum distance of the path. Secondly, by estimating the automatic update rate online, the trajectory generated by the proposed algorithm is guaranteed to be the minimum length and no collision, and the potential field function is minimized. Then, through the deep learning algorithm, the stability of the estimation error system is guaranteed. Finally, simulations and experiments show that the proposed algorithm can be accurately used for moving target detection and path prediction, and improve the performance of the control system.

Reference
[1] Markwalter B 2017 The path to driverless cars IEEE Consumer Electronics Magazine 6(2): 125-126.
[2] Tariq S, Choi H, Wasiq C M 2016 Controlled parking for self-driving cars IEEE International Conference on Systems, Man, and Cybernetics (SMC)
[3] Greenblatt, Nathan A 2016 Self-driving cars and the law IEEE Spectrum 53(2): 46-51.
[4] Qiu Z R, Xie L H, Hong Y G 2016 Quantized leaderless and leader-following consensus of high-order multi-agent systems with limited data rate IEEE Transactions on Automatic Control 61(9): 2432-2447.
[5] Liao F, Teo R, Wang J L, et al. 2017 Distributed formation and reconfiguration control of VTOL UAVs IEEE Transactions on Control Systems Technology 25: 270-277.
[6] Su H, Zhang N, Chen M Z, et al. 2013 Adaptive flocking with a virtual leader of multiple agents governed by locally Lipschitz nonlinearity Nonlinear Analysis: Real World Appl. 14(1): 798-806.
[7] Cai G B, Yan J, Zhao Y S, et al. 2017 Consensus of multi-vehicle cooperative attack with switching topology and stochastic fault Proceeding of Chinese Automation Congress (CAC 2017).
[8] Wagner G, Choset H M 2011 A complete multirobot path planning algorithm with performance bounds Proceeding of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems: 3260-3267.
[9] Mondal A, Behera L, Sahoo S R, et al. 2017 A novel multi-agent formation control law with collision avoidance IEEE/CAA Journal of Automatica Sinica 4(3): 558-568.
[10] Krupke D, Hemmer M, Mclurkin J, et al. 2015 A parallel distributed strategy for arraying a scattered robot swarm Proceeding of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).
[11] Katsev M, Yu J J, LaValle S M 2013 Efficient formation path planning on large graphs
Proceeding of 2013 IEEE International Conference on Robotics and Automation (ICRA)
Karlsruhe, Germany.