Multi-face recognition and dynamic tracking based on reinforcement learning algorithm

Yuxin Li[1]*, Yinggang* Xie[1,2,*] and *Xi* Lu[1]

[1]School of Information & Communication Engineering, Beijing Information Science and Technology University, China
[2]Key Laboratory of the Ministry of Education for Optoelectronic Measurement, Technology and Instrument, Beijing Information Science & Technology University, China

Abstract. Aiming at the problem that the current low accuracy rate of face detection and target tracking, a reinforcement learning algorithm is proposed, which integrates face detection technology and target tracking technology organically, adopts the face detection algorithm based on Multi-Task Convolutional Neural Network (MTCNN) and target tracking algorithm based on Kalman filtering, so as to realize face detection, multiplayer face recognition and dynamic tracking of personnel movement. In this paper, the configuration environment is Anaconda, the operating platform is PyCharm, the video-based face detection and dynamic capture and rapid identification system has been designed and developed. The system consists of two modules: face detection module and target tracking module. The optimized face detection and dynamic capture algorithm improved the detection success rate by about 11.5%, the face detection success rate by about 15.2%, the dynamic capture success rate increased by about 12.0%, and the optimized system has a wider practicality.

1 Introduction

Face detection can indicate the correct position and size of face, most of which is used in the pre-processing part before face recognition. Nowadays, the research on face detection and target tracking at home and abroad is a hot spot. The goal of tracking and identifying faces complements each other. Without finding faces, target monitoring cannot avoid security vulnerabilities, so the collaborative application of two technologies in different areas is significant.

Yang Xiao [1] elaborated carefully face recognition algorithm based on OpenCV and introduced several special algorithms. Shuo Yuan et al. [2] proposed a face tracking algorithm based on deep learning. Jinyi Pan [3] presents a dynamic face tracking algorithm KCFVR that integrates tracking checks and deep learning recognition aids. Weilin Li et al. [4] put forward a face detection and tracking method based on HOG and feature description,

---

* Corresponding author: xieyinggang@bistu.edu.cn

which reduces the interference of multi-face factors. Chengpan Liu et al. [5] introduced Deep-Sort tracking algorithm, which effectively solves the problem of poor real-time face tracking performance.

In the above literature, a variety of methods are proposed for face detection and target tracking. However, the face of the person being measured in the walk is distorted, deformed, partially or completely covered, the above methods are not accurate enough for face detection and tracking. Therefore, this paper proposes a combination of face detection algorithm based on MTCNN and target tracking algorithm based on Kalman filter to detect and track faces. The system block diagram as follows:
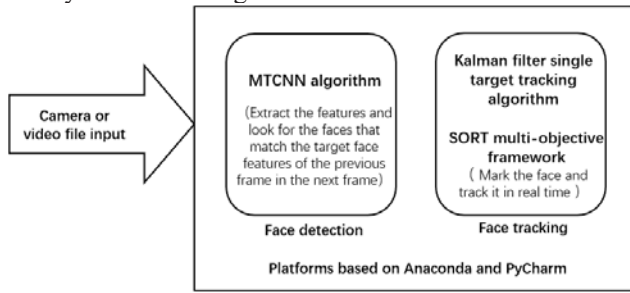

Fig. 1. System block diagram.

2 Face detection based on MTCNN

MTCNN (Multi-task convolutional neural network) [6] adopts the form of multilevel link, and its function is to predict people's facial features. The precision can be improved gradually from rough at first. In the environment affected by various complex factors, it can mark 5 feature points in the block diagram to achieve the face detection with reasonable accuracy, and finally achieve the function of output the correct result of face detection.

MTCNN is mainly divided into three steps: P-NET (Proposal Network), R-NET (Refine Network), O-NET (Output Network). As shown in the figure 2.
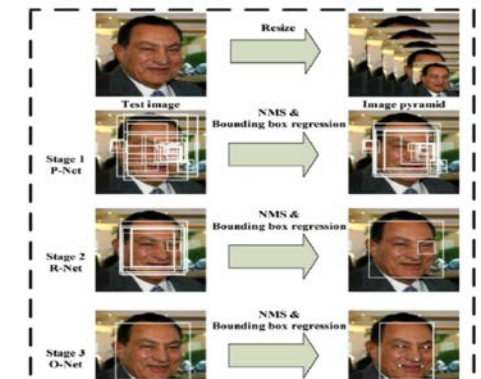

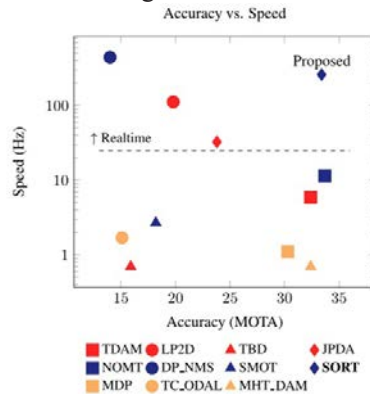**Fig. 2.** Algorithm implementation flow.   **Fig. 3.** Comparison of current tracking methods.

3 Target tracking

Face capture algorithms generally refer to these four situations: local features of face, model, neural network and motion information [7-12].

3.1 Single target dynamic capture algorithm

Kalman filtering [13] is adopted in this single target dynamic capture. Kalman filtering is divided into two steps: state prediction and state update, which will be discussed below.
Here are two important equations (equation of state and observation equation):

$$x_k = A * x_{k-1} + B * u_k + w_{k-1} \tag{1}$$

$$z_k = H * x_k + v_k \tag{2}$$

x_k is the state matrix of the system, z_k is the observed quantity of the state matrix, A is the state transition matrix, B is the control input matrix, H is the state observation matrix, w_{k-1} is the process noise (white noise), v_k is the measurement noise (white noise), the covariance is Q and R, namely, $p(w) \in N(0, Q)$, $p(v) \in N(0, R)$.
The specific process of Kalman filtering state prediction and update is as follows:
a) Time update equation:

$$\hat{x}_k^- = A\hat{x}_{k-1} + Bu_k \tag{8}$$

$$P_k^- = AP_{k-1}A^T + Q \tag{9}$$

b) State update equation:

$$K_k = P_k^- H^T (HP_k^- H^T + R)^{-1} \tag{10}$$

$$\hat{x}_k = \hat{x}_k^- + K_k(z_k - H\hat{x}_k^-) \tag{11}$$

$$P_k = (I - K_k H)P_k^- \tag{12}$$

3.2 Multi-target dynamic capture algorithm

The primary purpose of this algorithm is to generate a set of exclusive sequences from the received image stream, and to search and identify the moving target in this set of sequences.
SORT (Simple Online and Realtime Tracking) algorithm can be applied to achieve dynamic capture in multiple targets. It combines the advantages of Kalman filtering and Hungary algorithm, which can significantly improve the speed and keep the accuracy stable on the excellent line. It is an algorithm with extremely excellent indexes at present.
It can be seen from Figure 3 above that the SORT algorithm selected in this paper achieves a compatibility between accuracy and speed.
In the original algorithm of SORT, the face detection part adopts the Faster-RCNN architecture, and the advantages of the Faster-RCNN are reflected in two aspects. First, because of its advantages as a network structure, it can quickly form corresponding patterns for different design systems, and the performance improvement is visible to the naked eye. Second, the parameters between the two stages can be shared, and then the detection framework can be formed quickly. However, considering the tradeoff between speed and accuracy, this paper has replaced the original Faster-RCNN algorithm in SOFT algorithm with the MTCNN algorithm which is more suitable for requirements of this paper.
This section discusses a linear isokinetic model, because it has no connection with any other moving object in the system, so it is extremely suitable to represent the approximate displacement between the frames of each moving object in the system, that is, to transfer

the data information of the detection target of the current frame to the next frame. The relationship between the target states in the isokinetic model is shown in the formula:

$$x = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T$$

Variables u,v,s and r in the formula respectively represent the position of horizontal pixel in the center of the target under test, the position of vertical pixel in the center of the target under test, the ratio of binding frame and aspect ratio of the target under test.

4 Results analysis

The experimental environment of this system includes 2.4 (GHz) Intel Core i5-6200U, 8192MB memory, Anaconda as configuration environment, PyCharm2019 as the operating platform, and Python 3.6 as the programming language.

4.1 Face detection

In this paper, the method of training samples on WIDER_FACE data set is adopted, and the parameter allocation is as follows:

Fig. 4. Parameters set for training.   **Fig. 5.** Rendering of face detection.

The video stream collected from the Face was transferred to the computer environment, and then the Face-track-Detect environment was configured with Anaconda, and the code of the program was run with PyCharm to achieve the function of face detection.
The actual test results are shown in the figure 5.

4.2 Target dynamic capture

The dynamic target tracking of human face is to mark and predict the moving position of human face with the box in real time.
The process of SORT capture target algorithm and its actual effect diagram are shown in the following figure:
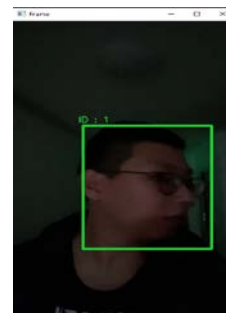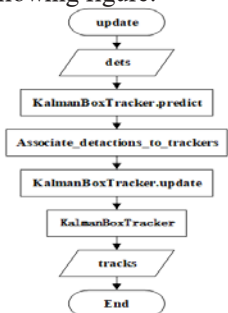

Fig. 6. Algorithm implementation flow.   **Fig. 7.** Full face effect. **Fig. 8.** Renderings of the side face.

Fig. 9. Marching effect (1).   **Fig. 10.** Marching effect (2).   **Fig. 11.** Marching effect (3).
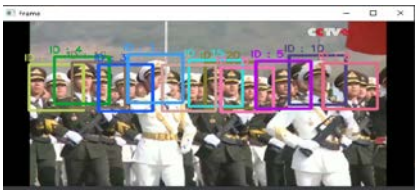

Fig. 12. Multi-target Face detection Effect (1).   **Fig. 13.** Multi-target Face detection Effect (2)

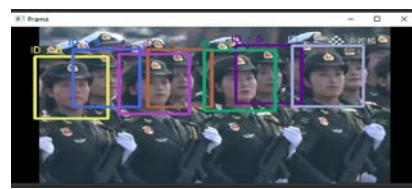**Fig. 14.** Multi-target Face detection Effect (3)   **Fig. 15.** Multi-target Face detection Effect (4)

The face database used for the test is CVPR19-01 of MOT19. The test results are presented in the form of table:

	Total sample size	The number of successful faces detected	Detection success rate	Testing time/ ms	Number of successful captures	Capture success rate
The original algorithm (All parameters are default values)	429	337	78.5%	10500	256	59.6%
Optimized algorithm	429	382	89.0%	8900	307	71.6%

Fig. 15. Comparison of algorithm effects.

Compared with the original algorithm, the method in this paper has a higher success rate of detection and capture. Meanwhile, compared with the previous method, the method in this paper has the shorter running time and better real-time performance.

3 Conclusion

In this experiment, MTCNN algorithm and SORT algorithm are used to work together to process images, extract facial information and further modify it into a face box. SORT algorithm fuses the advantages between Kalman filtering and Hungary algorithm in the realization of multi-objective dynamic capture, and maintains the compatibility of accuracy and speed.

In this study, reinforcement learning algorithm is adopted to realize face detection, multi-face detection and dynamic tracking when people are moving. The framework

proposed in this paper is also applicable to the detection and recognition of other objects. After comparing the basic knowledge and advantages and disadvantages of several common algorithms, the combination of MTCNN face detection algorithm and SORT target tracking algorithm is selected as the algorithm in this paper, which has high accuracy in face detection and target tracking, and also plays a key role in face recognition and other related work.

References

1. Yang Xiao. Research and implementation of face recognition algorithm based on OpenCV [D]. Hubei University of Technology,2018.
2. Shuo Yuan. Face detection and tracking based on HR and Siamese Net [D]. Central China Normal University,2019.
3. Jinyi Pan, Qin Wang, Wei Wang, Yameng Wang, Jianfeng Liu. Dynamic face tracking method based on deep learning assistance [J]. Journal of Zhejiang University of Technology,2018,46(05):481-486.
4. Penglin Li, Jiacheng Zhou, Wei Li. Face detection and tracking based on HOG and Feature descriptor [J]. Journal of Zhejiang University of Technology,2020,48(02):133-140.
5. Chengpan Liu. Dynamic face recognition based on deep learning [D]. Southwest University of Science and Technology,2020.
6. Zhang K , Zhang Z , Li Z , et al. Joint Face Detection and Alignment Using Multitask Cascaded Convolutional Networks[J]. IEEE Signal Processing Letters, 2016, 23(10):1499-1503.
7. Jing Cheng. Camera software system based on face and feature point recognition algorithm [D]. University of Electronic Science and Technology,2007.
8. Zhihua Jia. Research on face Recognition Technology of Intelligent Security System in Urban Rail Transit [J]. Communication World,2019,26(08):9-10.
9. Ping Cheng, Leiqiang Wang, Dongwei Ni. Design and Implementation of moving face capture system based on embedded platform [J]. Journal of PLA University of Science and Technology (Natural Science edition), 2012, 13(006):630-634.
10. Tingting Xu. Research on face detection Algorithm based on convolutional neural network [D]. 2019.
11. Yirong Zhao, Kunlun Li, Yanbo Li. Face Recognition based on Self-training Large Interval Near neighbor [J]. Digital Technology and Application,2015(02):67-68.
12. Xiaobo Liao. Research on target tracking algorithm based on Kalman filter in orbit [D]. University of Chinese Academy of Sciences (Institute of Optoelectronics Technology, Chinese Academy of Sciences),2017.
13. Zhihong Cheng, Liqin Huang. On-line multi-target tracking algorithm based on Kalman filter and multi-information fusion [J]. Information and communication,2019(03):35-38.
14. Tianping Li,Pingping Zhou,Hui Liu.Multiple features fusion based video facetracking [J].Multimedia Tools and Applications,2019,78(15).

15. Jiaxing Huang, Zhengnan Yuan, Xuan Zhou. A Learning Framework for Target Detection and Human Face Recognition in Real Time[J]. International Journal of Technology and Human Interaction (IJTHI),2019,15(3).

16. Wei Zeng, Lidong Jiang, Yueyuan Liu, Seng Yu. Design of multi-task Cascading Convolution face Tracking System [J]. Microcontroller and embedded system application,2020,20(07):47-50+55.

17. Jiebing Sun. Research and development of multi-person identification system based on Adaboost face detection algorithm [D]. 2020.

18. Lin Zhang, Jin Hou. A Face tracking algorithm in low illumination scene [J]. Automation and Instrumentation,2020(11):5-9.

19. Bernardo Lagos-Álvarez,Leonardo Padilla,Jorge Mateu,Guillermo Ferreira.A Kalman filter method for estimation and prediction of space–time data with an autoregressive structure[J]. Journal of Statistical Planning and Inference,2019,203.

20. Hazhar Sufi Karimi,Balasubramaniam Natarajan. Kalman filtered compressive sensing with intermittent observations[J]. Signal Processing,2019,163.