Research on Care Workers Physical Exercise Video Tracking based on Deep Neural Network

Yanli Dou

Department of Physical Education, Lvliang University, lvliang, China

Abstract—**Most nursing workers have a positive cognitive attitude towards physical exercise, but their exercise behavior lags behind. There are significant differences in the frequency, time, experience and load of physical exercise among nurses of different ages. Care workers' union organizations should try their best to provide health assistance to medical staff to meet their needs for physical exercise, so as to ensure their physical and mental health. Aiming at the problem of target tracking in motion video, this paper proposes a method of tracking motion video of nursing staff based on deep neural network (DNN). The effectiveness and adaptability of this method are verified by an example. This method can not only track and estimate the position of the target effectively, but also describe the shape of the target well, thus solving the problem that the shape of the target is complex and difficult to track in motion video.**

Keywords—**DNN, Care workers, Physical exercise, Goal tracking**

I. INTRODUCTION

As the main body, Care workers are the basic force to complete medical tasks. Their health status not only directly affects the completion of medical and health care tasks, but also maintains the prosperity of China's health industry [1]. Healthy physique and abundant physical strength are the material basis for Care workers to work smoothly, as well as the guarantee of improving work efficiency, extending the working life, and healthy and longevity [2]. What about the health status and physical exercise status of care workers? There is not much attention and research, let alone the great attention of the society [3]. As care workers who treat patients and save people, they often contact with patients and should actively participate in physical exercise, which not only helps to enhance their own resistance and reduce the chance of being infected by diseases, but also helps to set an example for patients and achieve the demonstration effect [4].

In order to fully understand the real situation of Care workers' participation in physical exercise, and to deeply investigate and study their awareness of physical fitness and the present situation of participating in physical

exercise, it is of great significance to advocate a healthy "sports lifestyle" and improve the quality of life and health level [5]. As a group, Care workers are engaged in occupations with high knowledge density, high risk and high intensity of labor, which lead to less leisure time, irregular physical exercise and poor health [6]. In recent years, reports of untimely deaths have been common, which has sounded a wake-up call for us. Advocating lifelong physical education, physical exercise and fitness activities among the frontline Care workers in clinical practice, enhancing the physical and mental health of Care workers, allowing them to better engage in work with a strong physique and a happy mood, which will contribute to the construction of a harmonious society. Both have extremely important practical significance [7].

Deep learning techniques have been widely used in human behavior recognition problems, but the number of parameters of deep learning models is huge, and the use of computers with strong computational power is required during both training and inference [8]. This mode increases the burden of the background server, puts high demands on the concurrency of the background server, and increases the difficulty of developing the corresponding software. Under the condition of ensuring certain accuracy, it is of great significance to study the behavior recognition model with fast calculation speed and few parameters [9]. Therefore, this paper puts forward the research on the video tracking of the physical exercise of Care workers based on DNN.

II. RELATED WORK

Physical exercise is the most effective method and means to enhance physical fitness and promote health. Caring for and understanding the health of Care workers, especially their attitude, motivation, interest, and need for physical exercise, is the most effective means to promote their health. Literature [10] stated that the attitudes, motivations, interests, needs of care workers towards physical exercise, with the aim of attracting attention from all levels of functional departments to their health status, health environment, healthy life, better

helping them to improve their physical and mental health, and pertinently giving assistance in physical fitness and promoting their physical and mental health. Literature [11-12] indicates that people's physical exercise behavior should be dominated and controlled by physical exercise consciousness, which is manifested in the cognitive attitude of Care workers towards physical exercise. Since the literature [13] won the championship in ILSVRC in 2012, DNN has been quickly applied to various fields. Literature [14] researched and proposed two network architectures of single-stream network and dual-stream network. The difference between these two different methods is how to integrate the spatio-temporal information of the video. Literature [15] studies suggest on the basis of the felzenszwalb mannequin, which proposes to dynamically track the moving human body by learning an appearance model that detects a target.The targets were traced using a planar anthropomorphic morphology model in [16] to guarantee that the tracking effect was more precise.

III. VIDEO TRACKING OF CARE WORKERS' PHYSICAL EXERCISE

As a group of intellectuals, Care workers have high work intensity, high risks, and excessive psychological pressure [17]. It's all for the patients. They forget to take a rest and their leisure time is replaced by work. The status of Care workers participating in physical exercise is not optimistic, especially those under the age of 30 have the highest percentage of participating in exercise 0 times a week. The reason why Care workers cannot participate in physical exercise or does not participate in physical exercise is mainly due to the particularity of Care workers [18]. Table 1 shows the situation of Care workers participating in physical exercise activities.

According to table 2, among care workers who regularly participate in physical exercise, the purpose of physical exercise is ranked in the top 4 categories: for strengthening fitness and fitness, for dissipating the mind, dissipating entertainment, for losing weight, bodybuilding, for resilience and improvement of one's own spiritual emotions [19].

With the formation and establishment of the social competition mechanism, the pressure of care work has become more prominent. This pressure mainly acts on the spirit of Care workers [20]. Because they bear the burden of treating diseases and saving people, they are facing urgent and critical patients and surgeries every day. A high degree of concentration, a little carelessness will lead to medical accidents, so they are thinking about how to perform surgery and treatment anytime and anywhere, especially clinicians, they claim that it is almost difficult to completely relax to rest or exercise, because the heart There is always concern [21]. From this point of view, work occupies a large amount of time outside of 8 hours, causing Care workers to be busy all day long and cannot get the rest and physical exercises they deserve.

Interest in sports always points to certain sports content, practice methods and composition forms. The statistical results can truly reflect the sports interests of Care workers, whose sports interests focus on fitness, entertainment and bodybuilding in turn. The root of this phenomenon lies in the high cultural level of clinical Care workers, whose intellectual development is manifested in individual activities, who like to move from group to independence, and can rationally and objectively understand their physical conditions, age characteristics and environmental factors [22]. When participating in sports activities, they generally avoid competitive events with certain collision and danger, and choose activities that are within their ability, have high practical value and benefit for life. Target tracking algorithms have evolved over the years, with a plethora of excellent tracking algorithms emerging.The workflow of the conventional target tracking algorithm is shown in Figure 1. In the field of computer vision, object tracking technology is widely used in various fields, so that computers can achieve the purpose of intelligence, high efficiency and strong interactive ability. According to Gavila, the application of tracking technology is divided into four aspects: motion recognition, virtual reality, interaction between human and computer and intelligent monitor [23]. Bayesian state filtering has two theoretical tools: Kalman filter and particle filter [24].

IV. DNN BASED VIDEO TRACKING OF PHYSICAL EXERCISE AMONG CARE WORKERS

A. Simple Recurrent Neural Networks

We summarize typical algorithms for each class of tracking model since its application from the beginning of deep learning techniques to the targeted tracking field, as shown in Table 3. An expanded analysis of each class of tracking model is presented below.

In 2006, inspired by the visual mechanism of human brain, Professor Hin-ton of University of Toronto put forward the concept of deep learning [25]. Deep learning is a new field developed from artificial neural network in machine learning, and its structure diagram is shown in Figure 2.

Considering that the tracking results of a single tracker may be unstable and unreliable, by combining deep learning and traditional ensemble learning methods, multiple weak classifiers are combined into a strong classifier in an adaptive weighted combination to improve the discrimination of the model Capability and stability are also current research hotspots in the tracking field. The basic framework based on deep integrated learning is shown in Figure 3.

With the rapid development of deep learning in the field of computer vision, researchers have found that the supervised way convolutional neural network (CNN) structure has better image abstraction ability relative to other network structures, and the convolutional network structure diagram is shown in Figure 4.

In his book Statistical Learning Methods, Li Hang called sequence modeling labeling problem, which maps an input or observation sequence to an output or marker sequence, while video behavior recognition problem corresponds to the former. In traditional machine learning methods, researchers often use hidden Markov models and conditional random fields to complete sequence modeling. However, in recent years, deep

learning methods, such as recurrent neural networks (RNN), have outstanding performance in sequence modeling due to their powerful characterization capabilities [26].

Compared with the classic fully connected network, RNN has many advantages, such as parameter sharing and can characterize long-term dependencies. The cyclic network will have a memory mechanism to provide information about the previous time step for the prediction of the current time step. Figure 5 shows the classic RNN architecture. The network shown in the figure has cyclic connections between hidden layers.

The calculation formula after unfolding this network in time steps is as follows

$$
\begin{aligned}
a_t &= b + W h_{t-1} + U x_t \\
h_t &= \tanh(a_t) \\
o_t &= c + V h_t \\
\hat{y}_t &= soft\max(o_t)
\end{aligned}
\tag{1}
$$

Where x and y represent data points with corresponding annotations, h is the loop unit or loop body, respectively, and τ is the loss value between the predicted value and the annotated value. The above formula demonstrates the calculation method at time step T, by which the activation value at the current time step is calculated from the input data x_t and the previous hidden state h_{t-1}, and subsequently the current hidden state h_t is calculated by which the log probability o_t of the output at the current time step can be calculated, Model loss can finally be calculated from this output.

B. Long Short-term Memory Network

The Video Attn model proposed in this paper uses LSTM to model the timing information of the video. LSTM is an advanced version of the classic RNN, which to a certain extent alleviates the problem of gradient disappearance or gradient explosion in traditional RNN, and at the same time for serialized data Relevant information with long distances has good characterization ability. LSTM solves the problems encountered in simple RNNs by explicitly introducing memory units. Figure 6 shows the structure of a neuron in an LSTM network.

A first look at the structure of the LSTM neuron in the upper image would seem very complex and we could ignore its internal structure and look at its inputs and outputs first when performing the analysis. The network receives three inputs, x_t is the input for the current time step, h_{t-1} is the output of the last LSTM neuron, c_{t-1} is a memory of the last neuron and the most important component of LSTM neurons, which can also be referred to as the cell state of that neuron. The network has two outputs, h_t is the output of the current neuron and c_t is the memory state of the current cell. Therefore, an LSTM neuron makes decisions through the input of the current time

step, the output of the previous time step and the memory state, and then outputs the output and memory state of the current time step. The internal structure of LSTM neurons can be divided into three parts, which are the forgetting gate, the input gate and the output gate.

(1) Forgotten door. The calculation formula is as follows

$$
f_t = \sigma\left(W_f h_{t-1} + U_f x_t + b_f\right)
\tag{2}
$$

Among them, W_f and U_f are the weight matrices of the forgetting gate, and b_f is the bias of the forgetting gate.

(2) Input gate. The specific calculation formula is as follows

$$
\begin{aligned}
i_t &= \sigma\left(W_i h_{t-1} + U_i x_i + b_i\right) \\
a_t &= \tanh\left(W_a h_{t-1} + U_a x_t + b_a\right)
\end{aligned}
\tag{3}
$$

Where, W_i and U_i are the weights of the input gate and b_i is the bias of the input gate. W_a and U_a are weights of candidate neural networks, and b_a is bias of candidate neural networks. Eventually, the newly generated memory and the partially forgotten memory converge into the current memory output of that neuron through an operation similar to a three-way valve, updating the cell state at the previous moment to the state at the current moment (new memory), which is calculated specifically as

$$
C_t = C_{t-1}\Theta f_t + i_t\Theta a_t
\tag{4}
$$

The calculation formula is as follows

$$
\begin{aligned}
o_t &= \sigma\left(W_o h_{t-1} + U_o x_t + b_o\right) \\
h_t &= o_t\Theta \tanh(c_t)
\end{aligned}
\tag{5}
$$

Long-term and short-term memory network, a neural network structure based on gating, has internal state to increase memory capacity. Data preprocessing involves transforming the original data into a machine learning or deep learning model, which is also very important for accelerating network training. Common preprocessing methods used in deep learning include normalization and whitening. Video data is a collection of image sequences that need to be pre-processed before they can be sent to the network for training and testing. The processing flow is shown in Figure 7.

The UCF101 motion recognition dataset is videos of human daily activities collected from the YouTube video website, with a total of 13320 videos, divided into 101 human behavior categories. Each of these classes has 25 human movements, each in 4 to 7 groups, with the majority having a video resolution of 320 $\times\times$240. These 101 included 5 major categories of action, human and object interaction, human limb movement, human person interaction, playing music equipment, and all types of exercise.

The videos collected by UCF101 data set are all from real scenes of life, and the data set has great diversity of actions, including great changes in lighting conditions, background, visual scale and corners, camera motion, human body shape and posture, etc. Because of the diversity of videos, this data

set is a very challenging data set.

Experiments will be conducted to verify the effect of the attention mechanism on behavior recognition. At the same time, different hyperparameters will be used to train the model to compare the effects of different hyperparameters on the performance of the model. The UCF101 data set officially split the data set into training set and test set, and the ratio is roughly 2:1. In the training process, the official training set is divided into two parts, 80% of which is used for training, and the remaining 20% is used for verification. Finally, after the training is completed, the official test set is tested. The specific The number of actions in the training set, validation set and test set are shown in Table 4.

In practice, the video format to be processed is not recorded in equal scale, and the human motion region obtained by region segmentation is not square in most cases. The experimental results are shown in fig. 8.

By manually marking the reference frame as the reference true value of each frame of the actual scene video, the distance accuracy curve and success rate curve of tracking the actual scene video can be drawn, as shown in Table 5 and Figure 9. From the experimental results, it can be seen that the video analysis of nurse's physical exercise based on deep neural network has better tracking effect for practical application scenarios.

During the experiment, it is necessary to set the parameters of the model, one of which is the hyperparameters of the neural network, such as the learning rate, the weight decay coefficient, etc., which are related to the convergence rate when the neural network is trained, whether it is overfitting or not.

In order to compare the effects of three frame cutting methods on behavior recognition rate, the video data are processed by three frame cutting methods, and other parameters are consistent. The final experimental results are shown in Table 6.

It can be found that the VideoAttn model proposed in this paper has a better recognition effect for behavior categories that change greatly between video frames, and it is more inclined to learn the characteristics of objects related to actions in the video frames.

C. Suggest

Care workers' union organizations should take various effective measures to encourage care workers to actively participate in physical activity, such as multi organization sports between some departments, especially recreational and ornamental programs.Meanwhile, care workers should enhance the awareness of participating in physical exercise to establish a "" physical lifestyle "", alleviate the stress imposed by the work environment by performing various physical activities, enrich the leisure life, and improve the health level and quality of life. The administrative department should pay attention to the physical exercise of Care workers, arrange the time of work and rest, and effectively improve the fitness environment and conditions of Care workers. According to actual needs, some activity venues, such as table tennis activity rooms, should be built in the limited venues that require small space and are easy to carry out sports. Residential areas should be reserved for sports activities, and venues suitable for group activities should be constructed, such as basketball courts, badminton courts, and venues suitable for walking and aerobics.

V. CONCLUSION

In order to realize the video tracking of physical exercise by medical staff, DNN tracking is proposed. In order to improve the speed of video behavior recognition and reduce the computational complexity of the model, a video Attn model is proposed, which directly uses CNN to model video behavior. After obtaining the video frame features of each time step through the pre-training network, a certain feature fusion method is adopted to superimpose the video frame features of all time steps to form a matrix as the input of subsequent CNN. Experimental results show that the proposed method is effective. This method can not only track and estimate the position of the target effectively, but also describe the shape of the target well, thus solving the problem that the shape of the target is complex and difficult to track in motion video.

REFERENCES

[1] Ouyang Gu, Zhong Bineng, Bai Bing, et al. Application and latest research progress of DNN in target tracking algorithm. Small Microcomputer System, 2018, 39(02):315-323.

[2] Hou Zhiqiang, Dai Bo, Hu Dan, et al. Visual tracking based on perceptual DNN. Journal of Electronics and Information Technology, 2016, 038(007): 1616-1623.

[3] Le Guen M, Cholley B, Fischler M. New Fast-Track Concepts in Thoracic Surgery: Anesthetic Implications. Current Anesthesiology Reports, 2016, 6(2):117-124.

[4] Chaffin D, Heidl R, Hollenbeck J R, et al. The promise and perils of wearable sensors in organizational research. Organizational Research Methods, 2017, 20(1):3-31.

[5] Huang Xiongbo, Hu Yongjian, Wang Yufei. DNN detection method for video motion vector steganography. Journal of South China University of Technology (Natural Science Edition), 2020, v48; No.407(08):5-13.

[6] Snyder B L. Practicing What We Preach: Teaching Psychiatric–Mental Health Student Nurses to Care For Themselves[J]. Journal of psychosocial nursing and mental health services, 2020, 2020, Volume 58, Issue 6, Page 40-45.

[7] Andy ML, Fiona DN, Mildred L, et al. The electronic tracking of referral and attendance at cardiac rehabilitation in Counties Manukau Health: a potential model for New Zealand. The New Zealand medical

journal, 2018, 2016 Volume 129, Issue 1446: Pages 64-71.

[8] Yuan Jiajie, Zhang Ling, Chen Yunhua. DNN image recognition based on attention convolution module. Computer Engineering and Applications, 2019, 55(008): 9-16.

[9] Liu Fan, Liu Pengyuan, Li Bing, et al. Design of deep learning model for video target tracking under TensorFlow platform. Progress in Laser and Optoelectronics, 2017(09):283-291.

[10] He Jiangguo. The effect of different physical exercises on slow-release long-term night shift medical staff. Digest of World Latest Medical Information, 2019, v.19(83):228-229.

[11] Rossetti B J, Dynes T, Brosi B, et al. GRAPHITE: A graphical environment for scalable in situ video tracking of moving insects. Methods in Ecology and Evolution, 2018, 9(4):956-964.

[12] Alkassab A T, Kirchner W H. Assessment of acute sublethal effects of clothianidin on motor function of honeybee workers using video-tracking analysis. Ecotoxicology and Environmental Safety, 2018, 147(jan.):200-205.

[13] Li Jijun. Discussion on several issues related to physical exercise to relieve the psychological pressure of medical staff. Contemporary Sports Science and Technology, 2017, 007(033):201-203.

[14] Li Jijun. A comparative study on the psychological stress of medical staff with different levels of participation in physical exercises: Taking medical staff in Nanchong as an example. Contemporary Sports Science and Technology, 2016, 6(31):153-154.

[15] Zhu L, Chen K, Du Y, et al. Heart Rate Monitoring During Physical Exercise From Photoplethysmography Using Neural Network. IEEE Sensors Letters, 2019, 3(1):1-4.

[16] Anderson L, Orme P, Di Michele R, et al. Quantification of Seasonal Long Physical Load in Soccer Players With Different Starting Status From the English Premier League: Implications for Maintaining Squad Physical Fitness. International Journal of Sports Physiology & Performance, 2016:1-26.

[17] Chen Bianling, Wang Jie. The effect of rehabilitation skills education for Care workers on the ADL scores of patients with stroke sequelae and fractures after surgery. Shanghai Medicine, 2017, 38(010):56-58.

[18] Yang Hui. Research on moving target tracking technology in sports video. Automation and Instrumentation, 2016, 000(006):176-177.

[19] Yu Loucheng. Research on the Tracking Method of Video Moving Targets in Sports. Television Technology, 2018, 042(009):74-79.

[20] Zhao Lei, Xie Ying. Research on Automatic Tracking Method of Sports Image Sequence Marking Points. Automation and Instrumentation, 2019, No.241(11):35-38.

[21] Liu Xiaodong, Chen Xinbao, Yang Guanghui, et al. Exploration of ball sports target tracking methods in sports video. Popular Science and Technology, 2018, 20(01): 114-116.

[22] Chen Xi. Research on the Design and Implementation of Sports Video Analysis System in Sports Training. Electronic Design Engineering, 2018, 26(015):81-84.

[23] Nan Qiuhong. Research on Sports Video Analysis and Recognition in the Process of Sports Training. Modern Electronic Technology, 2017(11):68-71.

[24] Wang Xue. Design and implementation of video analysis system for sports training. Modern Electronic Technology, 2017, 40(015): 29-32.

[25] Killen SS, Ryan C, Williams TD. The Ecology of Exercise: Mechanisms Underlying Individual Variation in Behavior, Activity, and Performance: An Introduction to Symposium. INTEGR COMP BIOL, 2017, 2017,57(2):185-194.

[26] Caldwell H A, Proudfoot N A, King-Dowling S, et al. Tracking of physical activity and fitness during the early years. Applied Physiology Nutrition & Metabolism, 2016:apnm-2015-0338.

APPENDIX

Table 1. Statistics list of nursing staff participating in physical exercise/person (%)

Gender	N	0 times/week	≥3 times/week	≥30min/time	≥Medium intensity	Sports population
Man	257(100)	116(43.8)	80(29.10)	113(42.6)	117(44.2)	56(19.4)
Woman	333(100)	160(47.4)	98(28.1)	110(31.8)	135(39.6)	66(18.2)
Σ	579(100)	265(45.8)	167(28.6)	212(36.5)	241(41.6)	111(18.7)

Table 2. Survey statistics for the purpose of participating in sports activities (/%)

	Man	Woman	Total	Sort
To enhance physical fitness and health	70.5	76.1	73.6	1
In order to relieve boredom, relax and entertain	37.2	39.8	38.6	2
To lose weight and keep fit	17.8	61.1	36.1	3
For self-cultivation and improvement of mental and emotional	26.1	37.5	32.4	4
In order to improve one's athletic ability	42.8	14.7	27.4	5
Developed the habit of exercise in school days	26.1	14.7	19.9	6
In order to increase their chances of social interaction	31.7	7.9	18.6	7
For contact with family	16.8	17.0	17.4	8
In order to communicate with friends and companions	16.8	14.7	16.1	9
For physical therapy and rehabilitation	6.7	0	3.6	10

Table 3. Classification of target tracking algorithms based on deep learning technology

	Algorithm category	Typical algorithm
Classification of target tracking algorithms based on deep learning technology	(1) Tracking model based on online migration and characteristic attribute analysis	DLT, SO-DLT, MDNet, DeepTrack, CNN-SVM, HCFT, FCNT, DeepSRDCF, etc.
	(2) Tracking model based on deep ensemble learning	TCNN, SCTC, HDT, etc.
	(3) Tracking model based on temporal and spatial information	RATM, DeepTracking, ROLO, RTT, SANet, etc.
	(4) Tracking model based on similarity comparison and regression	SiameseFC, SINT, YC-NN, GOTURN, etc.

Table 4. UCF101 data set division

	Cycling	Diving	Typing	Drumming	Pay basketball	Total
Training set	87	95	85	105	90	7596
Validation set	31	32	30	35	31	1963
Test set	49	56	54	56	46	3794

Table 5. Tracking performance of actual scene video

Frame number	Distance accuracy (%)	Success rate (%)
Frame 1	98.45	95.74
Frame 10	99.02	97.25

Table 6. Comparative experiment of different frame cutting methods of UCF 101 data set (%)

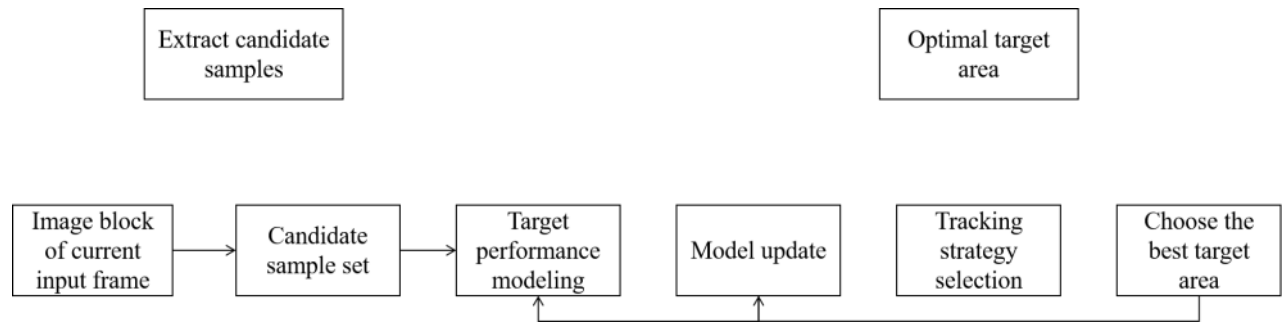| Number of LSTM network layers	Cycling	Diving	Typing	Drumming	Pay basketball	Average recognition rate
Uniform frame cut (reference method)	82.54	90.15	90.79	81.83	86.96	85.98
Randomly cut frames	82.27	83.62	81.63	85.37	90.32	82.78
Randomly cut frames evenly	87.39	92.01	81.65	83.23	84.72	87.74

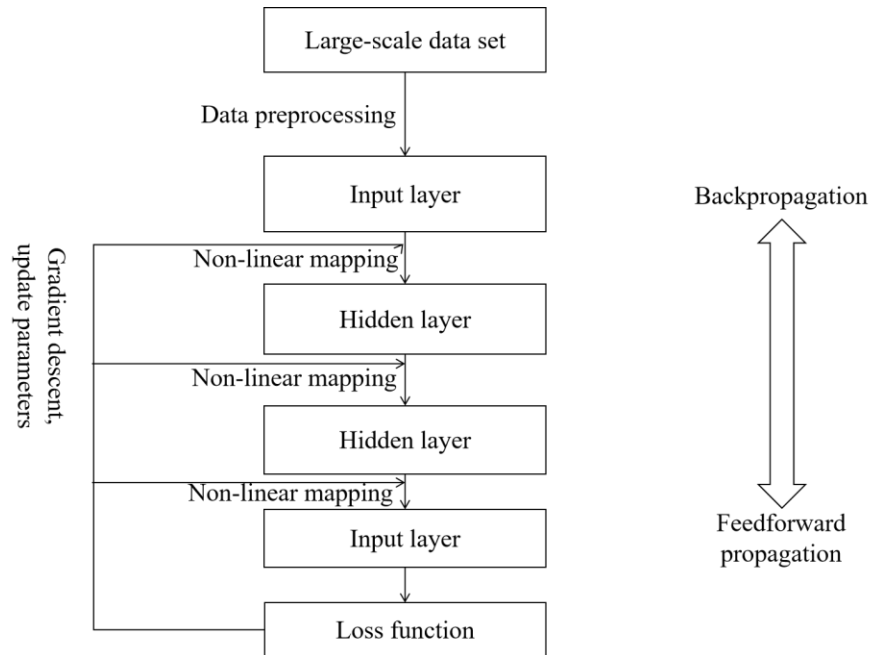Figure 1. Work flow chart of target tracking algorithm

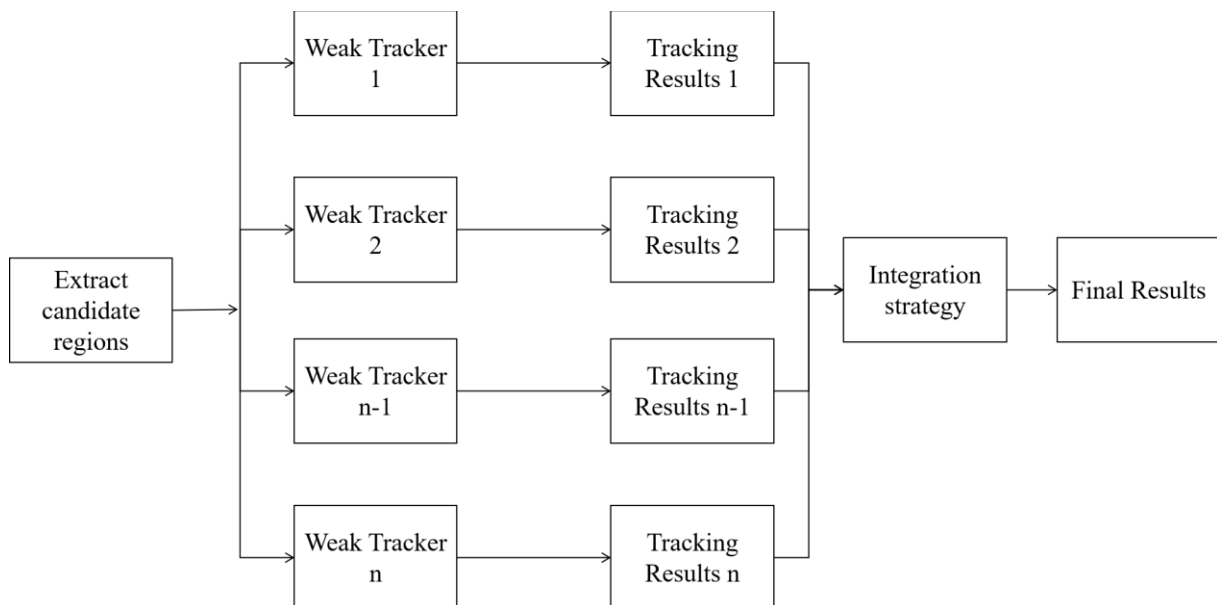Figure 2. The basic framework of deep learning technology

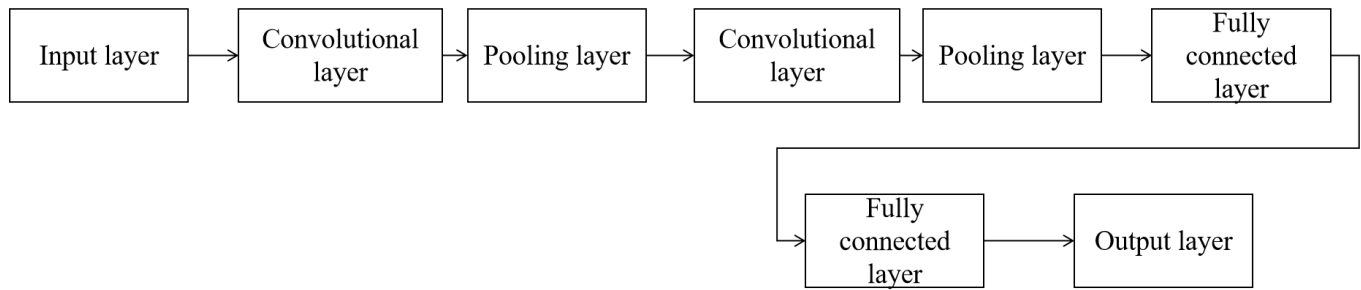Figure 3. Target tracking model based on ensemble learning

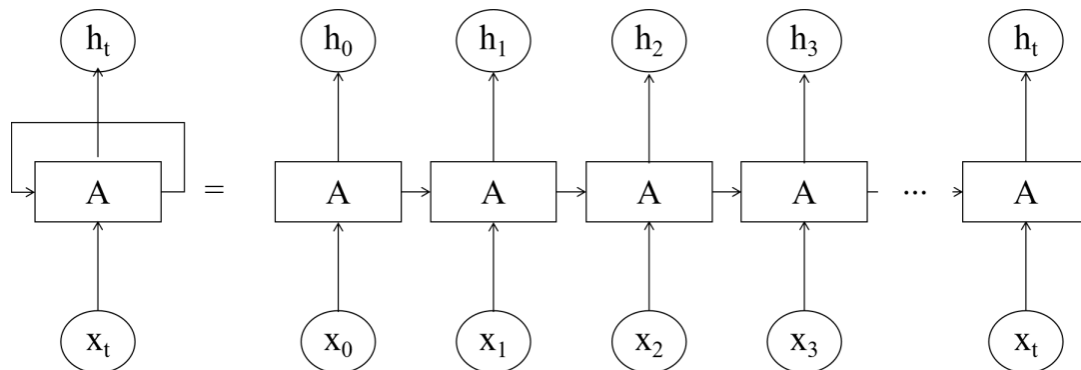Figure 4. Convolutional neural network structure diagram

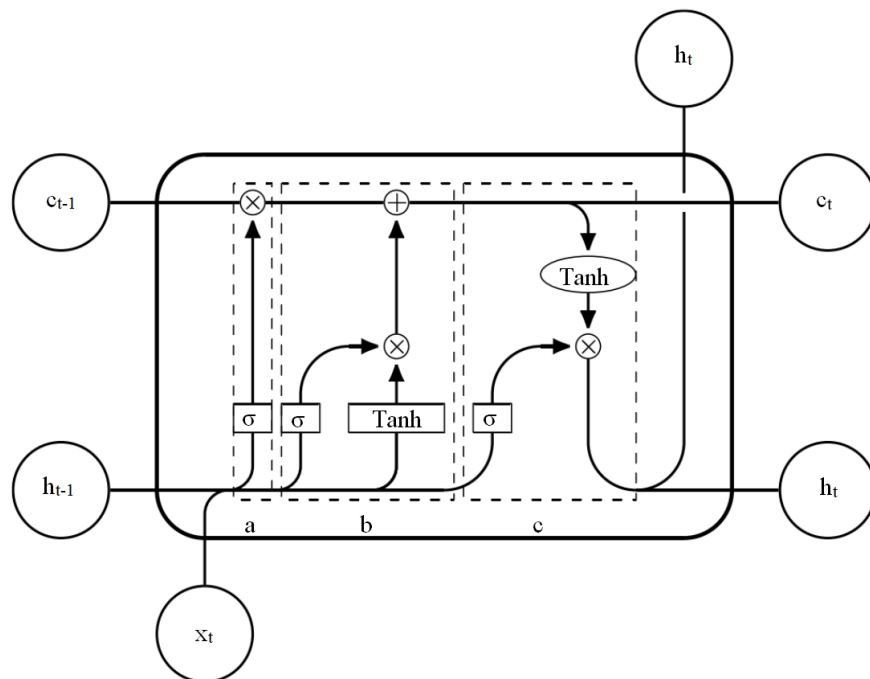Figure 5. Schematic diagram of a simple recurrent neural network
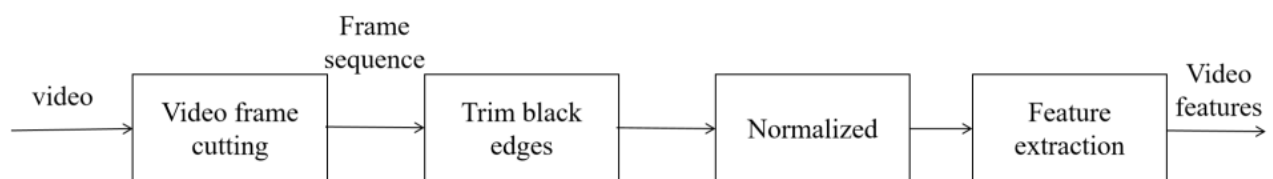

Figure 6. Schematic diagram of LSTM neuron

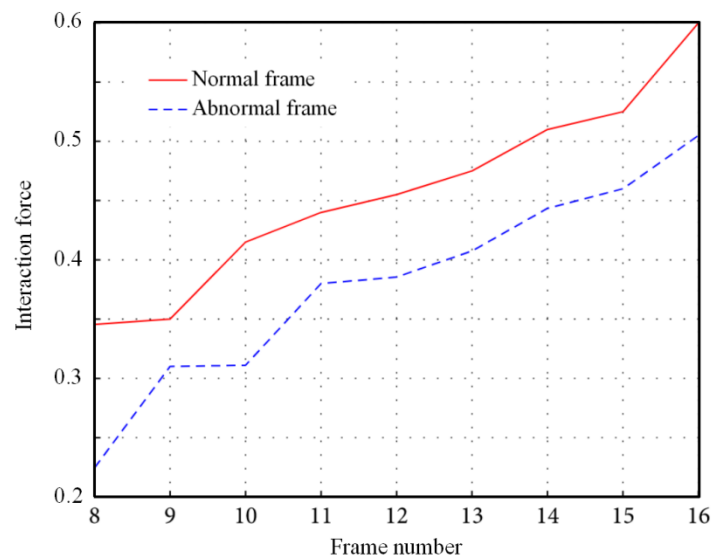

Figure 7. Data preprocessing flow chart
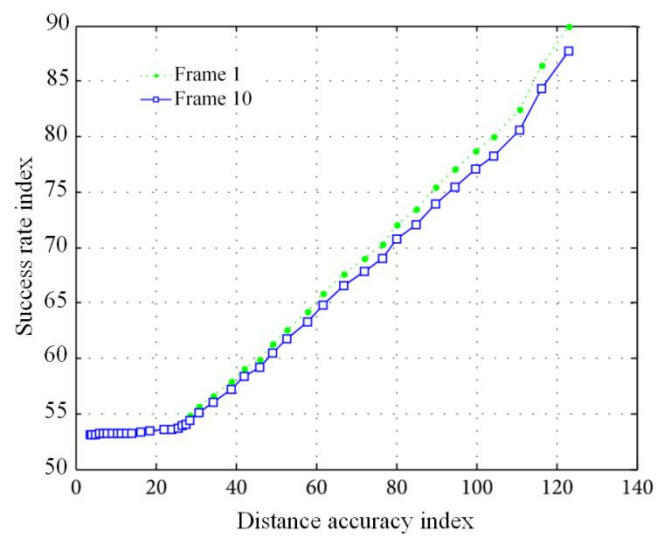
Figure 8. Interaction diagram of normal frame and abnormal frame

Figure 9. Tracking performance of actual scene video