Review Article

A New Gesture Recognition Technology Used in Intelligent Inspection Drones

Jun Zhu[1], Zhi-Zhuang Duan[2], Ding-Ding Yu[3*]

[1]Zhejiang Normal University Xingzhi College, Zhejiang, China
[2]Doctor of Law, Associate Professor of Zhejiang Normal University Xingzhi College, Master Tutor, China
[3]Zhejiang Normal University Xingzhi College, Zhejiang, China

*Corresponding author: Ding-Ding Yu
Zhejiang Normal University Xingzhi College, Zhejiang, China

Abstract

The present invention relates to an alarm method for recognizing gestures based on a camera, comprising: setting a gesture data set; establishing a frame by the camera to recognize the trend of human activities; recognizing human gestures by using a convolutional neural network; comparing a captured binary gesture image with a security gesture, and setting a counter parameter with a value of 1; if the gesture image is successfully compared with the security gesture, setting the value of the parameter to plus 1; if comparison between the gesture image and the security gesture fails, setting the value of the parameter to 0; and when the value of the parameter is greater than three, establishing a signal for the camera and the receiving device. The camera of the present invention uses the convolutional neural network to recognize human gesture, processes the human gesture and compares a gesture image with a security gesture. After the conditions are met, the position is sent to a receiving device of the police officer on duty to call the police in time, thereby reducing the frequency of crime. In the whole process, the camera is used for the capturing, and a person injured can effectively prevent more serious persecution without shouting.

Keywords: Recognizing gestures, Capturing, Compared.

INTRODUCTION

The present invention relates to the field of computer vision and human-computer interaction, in particular to an alarm method for recognizing gestures based on a camera.

With the continuous growth of the urban population, the social relations of people have been relatively complex, and the difficulty of controlling the lives of residents has been increased. However, the security system for smart city still needs to be continuously enhanced and improved. At present, in some alleys in the city, the patrol of police officers or security personnel is also limited, but there are a certain number of security tools such as cameras.

However, for persons who have been persecuted, the camera can only play a role in capturing criminal records. If the persecuted persons shout loudly, it will inevitably lead to more serious persecution.

Therefore, the society urgently needs to capture other effective information of the victim while retaining the original function of the camera to call to the police to further reduce the frequency of crimes in the alleys.

The present invention mainly invents an alarm method for recognizing gestures based on a camera with respect to the above problems. The camera uses a convolutional neural network to recognize human gesture, and obtains a gray image after background removal and binarization. The recognized binary gesture image is repeatedly compared with a security gesture, and after the comparison conditions are met, the information of the position can be sent to the police officer on duty or a receiving device of a police station to capture the gesture information of a person injured to call the police in time.

An alarm method for recognizing gestures based on a camera, which adopts a camera with a built-in high-definition camera module or an infrared template, and a receiving device for receiving signals, and comprises the following steps:

Firstly, it will setting a gesture data set; Furthermore establishing a frame by the camera to recognize the trend of human activities and recognizing human gestures by using a convolutional neural network. It is comparing a captured binary gesture image with a security gesture, and setting a counter parameter counter

Gesture with a value of 1. If the gesture image is successfully compared with the security gesture, setting the value of counter Gesture to plus 1. If comparison between the gesture image and the security gesture fails, setting the value of counter Gesture to 0. When the calculated value of counter Gesture is greater than three, establishing a signal for the camera and the receiving device. If the camera and receiver complete communication, the camera will send the coordinate position that meets the conditions of the gesture alarm to the receiver.

The alarm method based on camera recognition gesture mainly has the following great points, the setting of gesture data set, the camera establishing a frame to identify human activity trends, the use of convolutional neural network to identify human gestures, the camera and the receiving device establishes communication.

Setting the gesture data set specifically comprises:
First of all photographing below the high-definition camera to obtain security gesture images. Then converting each image to a YCbCr color space, removing a background through a GMM algorithm and resizing the images to 28×28 pixels of gray or binary images. For each gesture type, 600 and 200 images were taken from the subjects and the CNNs were trained and tested, respectively. Obtaining the binary images that are similar in gesture types, and obtaining image details with the gray images and finger positions.
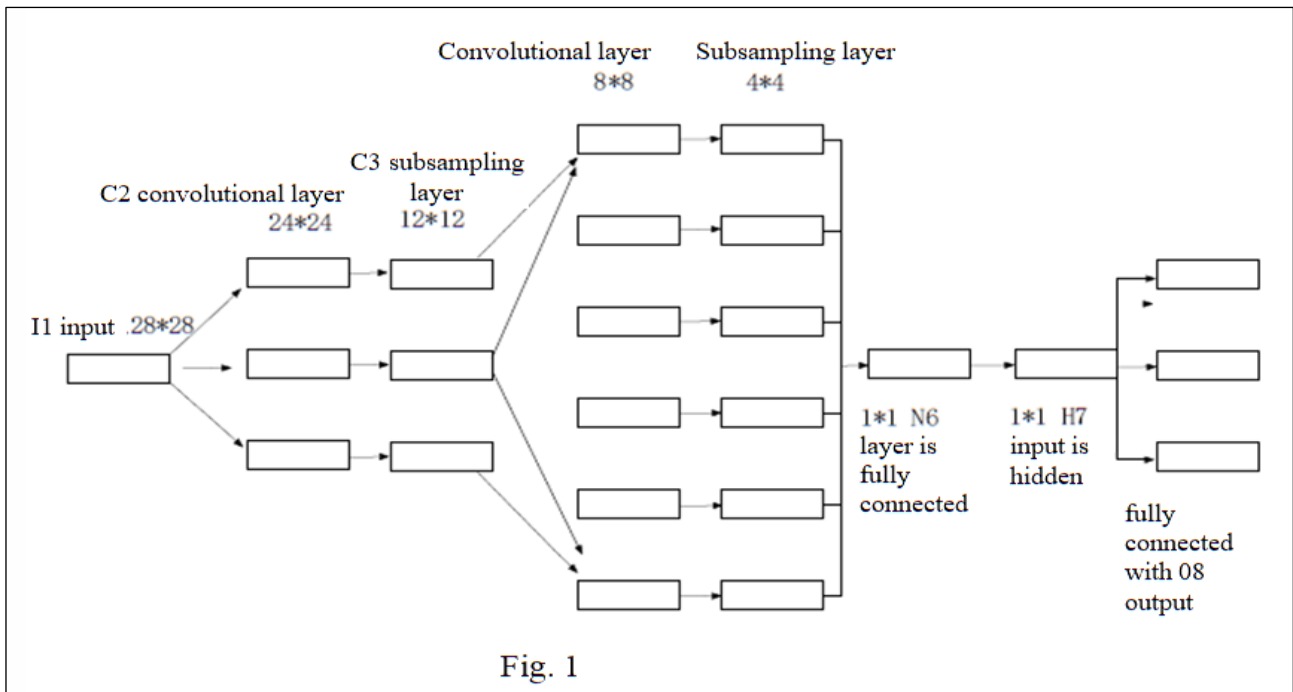

Fig. 1

Recognizing human gestures by using the convolutional neural network specifically as follows.

First of all it is building gesture models by using fingertip and hand contours as extraction features. Then

it will using the YCbCr color space less sensitive to illumination conditions than the RGB color space, and determining the pixel value (B) of the binary image by a YCbCr threshold, expressed as:

$$B = \begin{cases} 255, & \text{if } 50 \leq Y \leq 255; \\ & 90 \leq Cb \leq 155; \\ & 135 \leq Y \leq 180; \\ 0, & \text{otherwise.} \end{cases}$$

Classifying some pixels on a ring and a little finger as non-skin pixels, solving the problem by a Gaussian mixture model (GMM), and representing a GMM by K Gaussian components as:

$$P(x) = \sum_{k=1}^{K} P(k)P(x|k),$$

Where in P(k) is a prior probability (π) and P(x|k) is the conditional probability expressed as Gaussian distribution with mean (μ) and covariance (Σ);

Calibrating the image of the hand according to the position and the direction, deriving the binary image and calculating the center ($\bar{x}, \bar{y}$) of the hand as follows:

$$M_{i,j} = \sum_{x} \sum_{y} x^i y^j f(x,y)$$

$$\bar{x} = \frac{M_{1,0}}{M_{0,0}}, \quad \bar{y} = \frac{M_{0,1}}{M_{0,0}}$$

Where in x and y are coordinates of the skin pixel with the pixel value f(x,y) set to 1, and vice versa.

CONCLUSION
With this alarm technology based on camera recognition gestures, it can greatly improve the recognition ability of intelligent inspection drones, realize timely alarms, and reduce the frequency of crimes. And the whole process is captured by a camera, so the victim does not need to shout loudly, which can effectively prevent more serious persecution and maintain social order.

REFERENCES
1. Yang, K., Wang, Y., Tiw, P. J., Wang, C., Zou, X., Yuan, R., ... & Yang, Y. (2024). High-order sensory processing nanocircuit based on coupled VO2 oscillators. *Nature Communications*, *15*(1), 1693.
2. Tanusha, D., & Badhulika, S. (2024). Comparative Analysis of Micro Patterned PDMS-Based Piezoresistive Pressure Sensors with Multifunctional Strain and Health Monitoring Applications. *Sensors and Actuators A: Physical*, 115139.
3. Wang, Z., Wan, H., Meng, L., Zeng, Z., Akay, M., Chen, C., & Chen, W. (2024). Optimization of inter-subject sEMG-based hand gesture recognition tasks using unsupervised domain adaptation techniques. *Biomedical Signal Processing and Control*, *92*, 106086.
4. Dutta, A., Niu, Z., Abdullah, A. M., Tiwari, N., Biswas, M. A. S., Li, B., ... & Cheng, H. (2024). Closely Packed Stretchable Ultrasound Array Fabricated with Surface Charge Engineering for Contactless Gesture and Materials Detection. *Advanced Science*, 2303403.
5. Tuazon, R., & Mortezavi, S. (2024). Automatic labeling of facial zones for digital clinical application: An ensemble of semantic segmentation models. *Skin Research and Technology*, *30*(2), e13625.