International Journal of Science and Engineering Applications
Volume 13-Issue 03, 01 - 04, 2024, ISSN:- 2319 - 7560
DOI:10.7753/IJSEA1303.1001

Photo Face Detection Based on Deep Learning

Qingfeng Yu
School of Electric Information and
Electrical Engineering
Yangtze University
Jingzhou, China

Abstract: Face detection refers to the process of automatically detecting, identifying and locating faces in images or videos by using computer vision technology. The purpose of this technology is to distinguish the face in an image or video from other objects and backgrounds, so as to accurately locate and analyze the face features. Face detection can be achieved in many ways, including feature-based method, image-based method and deep learning-based method. In this paper,MTCNN convolutional neural network method based on deep learning is used, and three cascaded neural networks are constructed by using convolutional neural networks of various scales and small networks with different aspect ratios, and the information such as the position, size and direction of one or more faces is successfully identified and located in the image. At present, MTCNN convolutional neural network method has become one of the most popular technologies because of its high efficiency, accuracy and strong robustness.

Keywords:Deep learning; Face detection; Convolutional neural network; MTCNN neural network

1. INTRODUCTION

Face is a very complicated visual mode, and the human brain can recognize and communicate by observing and recognizing faces. For example, facial expressions can convey emotional states, and different characteristics such as age, gender and race can also be judged by faces. With the upgrading of computer hardware technology and the further development of deep learning algorithm, face detection technology has made remarkable progress. Traditional face detection methods are based on chromaticity, gradient and other characteristics, but they are not ideal in practical application and are easily affected by lighting, occlusion and other factors . The face detection method using deep learning technology has high accuracy and robustness, and can adapt to the change of illumination and angle. Therefore, in many fields, such as identification, security monitoring, medical diagnosis, video phone, video conference, intelligent access control system, face processing and analysis have a very common application .An important task in computer vision is to visually process the face. Now, face recognition technology has been transformed from theoretical progress to practical technology, and has been applied in various fields, such as security monitoring, financial payment, smart home, human-computer interaction and so on[1].

Nowadays, the application scene of face detection is no longer limited to face recognition system, but also has important application value in many other aspects. For example, in the field of intelligent security, with the popularity of security monitoring equipment, face detection technology can help protect security by identifying unfamiliar faces and automatically sending out alarm information. In the field of smart home, face detection technology can realize the functions of face recognition, door opening and intelligent lighting, and improve the quality of life and convenience[2]. In the field of intelligent transportation, face detection technology can realize the functions of vehicle trailing detection, highway entrance management and so on, and improve road traffic safety. These application prospects lead to face detection being widely concerned by researchers .In addition, it has a very wide range of applications, such as facial identification, video surveillance and so on. Ensuring that faces are correctly recognized is the case in our life.Therefore, the development of face detection technology is very necessary and has great potential, which will bring more convenience and safety to our life and work.

2. METHODOLOGY

2.1 Overview of Face Detection

The task of face detection is to judge whether there is a face in a given image, and if there is a face, it will return the position coordinates of the face and its image size. Early face detection methods were mainly based on classical computer vision technologies, such as Haar feature classifier, support vector machine, contour model and so on. These methods usually need to manually extract facial features and then use classifiers for recognition. These methods can realize face detection to a certain extent, but they need a lot of manual intervention and feature extraction, and have low tolerance for noise and occlusion[3] .

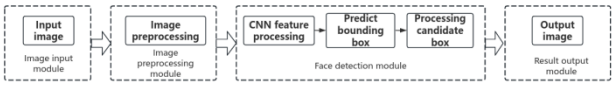

Figure.1 Face detection system

With the rise of deep learning technology, most modern face detection methods adopt neural network model. For example, themethod based on convolutional neural network performs well in face detection. These methods do not need to manually extract features, but directly input the input data (such as images) into the neural network model for feature extraction and classification. The advantage of this method is that it can automatically learn and adapt to various environments and complex scenes, and can deal with various noise and occlusion situations .

Face detection technology is one of the most important research directions in the field of artificial intelligence, and it is a widely studied academic field. In addition, it has a very wide range of applications, such as facial identification, video surveillance and so on. Ensuring that the face is correctly recognized is the basis of so many applications in our life[4]. Therefore, it is very necessary to develop face detection technology, which has great

potential and will bring more convenience and safety to our life and work. In the future, we can look forward to the emergence of more accurate, efficient and intelligent face detection technology, which will bring more convenience and benefits to society.

2.2 Convolutional Neural Network

Convolutional neural network is a kind of deep learning neural etwork specially used for processing data with similar grid structure. CNN is widely used in image recognition, computer vision and natural language processing.The core idea of CNN is to extract the features of input data through convolution layer and pooling layer. The convolution layer convolves the input data through a filter (also called convolution kernel) to capture local features. The results of these convolution operations are fed into the activation function to produce a nonlinear response. Pool layer is used to down-sample and reduce the size of feature map, while retaining the most important features.

CNN is usually composed of multiple convolution layers, activation function layers and pooling layers, and finally connected with the fully connected layer for classification or regression. Through the back propagation algorithm, CNN can automatically learn the characteristics of input data and make corresponding prediction or classification according to the task requirements.Because CNN can effectively process image data and has translation invariance (that is, it is robust to the translation of input data), it has achieved great success in the fields of image recognition and computer vision and has become one of the most important and widely used neural network models in the field of deep learning[5].

2.3 Model Selection

Given an image, the task of face detection is usually completed in three steps:

1. Select an area (usually a rectangular window) on the image as the observation window;

2. Extract some features from the selected window to describe the image area it contains;

3. According to the feature description, judge whether this window just frames a face.

The process of face detection is to repeatedly perform the above three steps until all the windows that need to be observed are traversed. If all the windows are judged not to contain faces, then it is considered that there is no face in the given image, otherwise, the windows judged as faces are further analyzed, so as to give the location and size of faces.

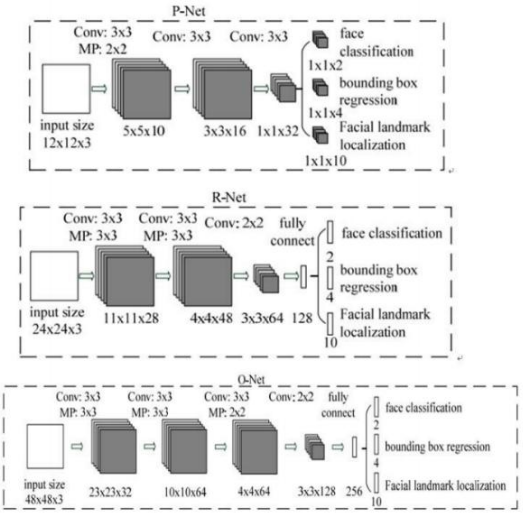

Figure.2Network structure of MTCNN model

In this paper, MTCNN algorithm is used to realize face detection because it is one of the more advanced and effective face detection algorithms at present. MTCNN adopts technologies such as multi-task cascade network structure, feature pyramid and joint training, which can improve the accuracy and efficiency of face detection, and can also deal with detection problems in complex scenes. At the same time, MTCNN can output multiple information such as face frame, key point position and face confidence at the same time, which can meet the needs of many different application scenarios. Therefore, this paper chooses MTCNN algorithm to realize face detection, which can improve the detection efficiency while ensuring the detection accuracy. It is suitable for various scenes and has high practicability and applicability.

3. EXPERIMENTAL PROCESS

3.1 Software Introduction

PyCharm is an integrated development environment (IDE) for Python language developed by JetBrains. It provides functions such as code editing, debugging, error checking, code analysis, version control, testing and document development, which can greatly improve the efficiency of Python development. PyCharm supports a variety of Python frameworks and libraries, including Django, Flask, NumPy, Pyramid, etc. It has powerful functions such as intelligent code prompt, automatic code completion, jumping to definition, code reconstruction, etc., which enables developers to write high-quality Python codes faster.

3.2 Environment Configuration

The overall system construction and environment configuration are as follows.

Table 1.Main configuration module

Operating system	Win11
Development framework	TensorFlow
Language environment	Python 3.6
Compilation environment	Anaconda3-5.2.0
Software	Pycharm community edition

3.3 Face detection process

Face detection is an important task in the field of computer vision, and its goal is to accurately locate and identify the position of human face in images or videos. MTCNN (Multi-task Cascaded Convolutional Neural Network) is a very popular face detection algorithm. It adopts a cascade method to detect faces, including the stages of candidate frame generation, candidate frame refinement and key point location, which can effectively improve accuracy and robustness. After three stages of processing, MTCNN can output the position of the face frame and the position of five key points in the image. This cascade method can gradually improve the accuracy of face detection, and can cope with various complex face detection scenarios[6].

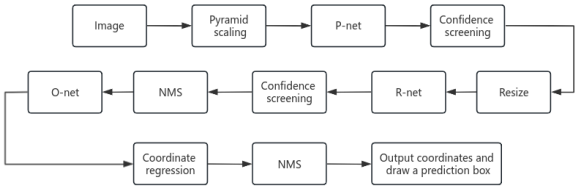

Figure.3MTCNN neural network face detection process

4. EXPERIMENTAL RESULTS

MTCNN uses three cascaded neural networks for face detection, namely P-Net, R-Net and O-Net. They perform rapid screening, accurate positioning and feature extraction for different size face regions[7].

Specifically, P-Net is a shallow convolutional neural network, which is used to quickly screen areas that may contain faces. The output is a prediction result containing multiple bounding boxes, and each candidate box has a corresponding probability value and bounding box coordinates.On the basis of P-Net, R-Net further screens candidate boxes and locates them accurately. It is a deeper convolutional neural network, which can detect and locate face regions more accurately. R-Net also adds a regression predictor to predict the exact position and size of the face frame.O-Net is a deeper network with more convolution layers, which is mainly used to extract more subtle features of human faces. It also contains two independent output layers, one is used to classify the attributes of face, such as age, gender and emotion, and the other is used to regress the exact position and size of the bounding box.

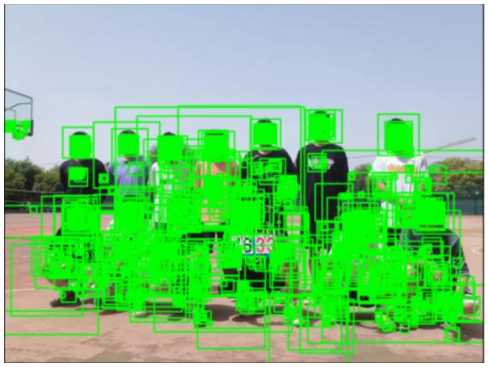

Figure.4P-Net network output results

Figure.5R-Net network output results

Figure.6O-Net network output results

5. CONCLUSION

MTCNN is a face detection algorithm based on deep convolution neural network, which has high accuracy and real-time performance. The algorithm mainly relies on three steps to complete the face detection task: candidate frame generation, feature point location and face classification[9].

In the face classification stage, MTCNN uses a convolutional neural network to judge whether each candidate box is a real face. In this stage, MTCNN adopted the cross entropy loss function of softmaxeffectively sample the data, which improved the classification accuracy of the algorithm.In the stage of candidate frame generation, MTCNN uses a multi-level convolutional neural network, which first extracts image features layer by layer, and then searches for candidate frames by combining scale transformation and sliding window.In the stage of feature point location, MTCNN expands the existing candidate frames through regression, and then uses convolutional neural network to locate feature points, including eyes, nose, mouth and so on.

To sum up, MTCNN algorithm has achieved good results in the field of face detection, and its high accuracy and good real-time performance make it widely used in practical applications. In the future, there is still much room for improvement in MTCNN, such as further improving the detection speed and accuracy of the algorithm, and solving the influence of various factors such as posture, expression and illumination on the algorithm [10].

6. REFERENCES

[1] Wu Jiyun, Chen Shiqin. An improved face detection algorithm for MTCNN [J]. Software Guide, 2019,18(12):78-81.

[2] Chen Yuwei. Face detection and facial key point location based on improved MTCNN model [D]. Donghua University, 2019.

[3] Chen Denggen. Research and Implementation of Face Recognition Based on MTCNN [D]. Anhui Agricultural University, 2021.

[4] Tan Bin. Improvement of MTCNN algorithm for face detection in classroom environment [D]. xihua university, 2019.

[5] Lan Wenfei, Zhang Shenglan, Zhu Rongbo, et al. Face detection algorithm based on improved MTCNN [J]. Journal of South-Central University for Nationalities (Natural Science Edition), 2020,39(06):637-641.

[6] Guan Jiangang. Face detection based on FPGA and MTCNN [D]. South China University of Technology, 2021.

[7] Zhang Yuyun. Application of multi-face recognition based on MTCNN and S-LBPH algorithm in classroom attendance [D]. Southwest Jiaotong University, 2016.

[8] Song mengyuan. research on multi-scale face detection network based on improved Faster RCNN [J]. automation table, doi: 10.16086/j.cnki.issn1000-0380.2022020053.

[9] Rowley,Henry Allan.Neural network-based face detection[D]. ProQuest Dissertations and Theses Full-text Search Platform,1999.

[10] Teli, Mohammad Nayeem.Face detection using correlation filters[D].ProQuest Dissertations and Theses Full-text Search Platform,2013.