Face Recognition and Model Analysis Using CNN Based on Self-Collected Information

Lanxin Zhang *

Grinnell College, Grinnell, 50112, USA

* Corresponding Author Email: zhanglan@grinnell.edu

Abstract. Face recognition helps people to determine the features and characteristics of a face, which in turn helps the police to find criminals, reduce the crime rate and ensure the safety of the society. In order to further improve the performance of face recognition, the main objective of this research is to collect a large amount of face information using Convolutional Neural Networks (CNN) to achieve effective recognition. Specifically, first, a code is used in the annotation to define the coordinates of the face. This function converts the classification labels into numbers that can be understood by the model. Second, generate targets in order to parse the annotation file and obtain objects from it. Third, a CNN model is constructed. The features in the data are captured using the adaptive weights of the CNN. In addition, the predictive performance of different models is analysed and compared. In addition, the objective function will transform the face coordinates of the file attached to the boxed list, which helps to acquire the commanded faces in the image file. It is shown through experiments that the proposed model can perform face recognition effectively. The proposed program in this paper can provide valuable thoughts and insights to the field.

Keywords: Face recognition, Convolutional Neural Networks, Face information, Classification.

1. Introduction

The COVID-19 pandemic has been crossing the whole world for several years. According to the data from WHO, this global epidemic has caused over 6.95 million death cases and 7.69 hundred million confirmed cases [1]. To prevent the spread of the epidemic and thus protect their lives for this reason in public places, people must wear masks. However, this paper found that there are few people who refuse to wear face masks in public, which is dangerous for other people's health and lives and creates chaos for the society and country.

Therefore, to prevent this kind of accident from happening again remind people to wear face masks and check whether they wear face masks correctly or not in public. This paper would like to design a kind of "machine" in front of public places, such as schools, workplaces, restaurants and cafes etc [2]. In the past, there were some volunteers who used fever thermometers in front of the gates of public places. They reminded people to wear their masks and measured their body temperature [3]. Also, some policemen and policewomen used monitors to detect someone who did not wear their masks. These artificial methods not only need a lot of people to work hard but also have low efficiency [4]. Therefore, it needs to design a kind of code by computer language to detect people who do not wear their masks or wear their masks incorrectly. This paper chose to use Python as computer language because this language is popular and powerful for a variety of fields in society, such as exploiting websites, and designing computer games and AI software. Now this paper would like to use this powerful computer language to design detective codes to monitor people wearing their masks in the best way in public spaces.

The main objective of this research is to perform effective recognition by collecting a large amount of face information using Convolutional Neural Networks (CNN). Specifically, first, a code is used in the annotation to define the coordinates of the face. This function converts the classification labels into numbers that are understandable by the model. Second, targets are generated in order to parse the annotation file and obtain objects from it. Third, a CNN model is constructed. The features in the data are captured by utilizing the adaptive weights of the CNN. Additionally, the predictive performance of different models is analysed and compared. In addition, the objective function will

transform the face coordinates of the files attached to the boxed list, which helps to acquire the commanded faces in the image files. The experimental results show that the proposed model can perform face recognition effectively. Face recognition can help people to determine the features and characteristics of faces, which can help the police to find criminals, reduce the crime rate, and ensure the safety of the society.

2. Methodology

2.1. Dataset description and preprocessing

In this paper, a large number of face images are collected and translated into files [5,6]. Then this paper finds the file name and predicts the face by finding it in the provided annotations. Then coordinate the faces and generate the objective function. Parse the annotated file and get the object from it. For this function, the annotation file will be read to get the objects under the object tag in the annotated XML file. The annotation file (XML file) will be scraped and all the objects in the annotation file being called by the function will be found as a list of strings.

2.2. Proposed approach

The focus of this paper is on face recognition using CNN. Image features are captured by introducing CNN. Face feature capture and recognition is accomplished by integrating and reorganizing the inputs. Specifically, first, the model builds diverse images. In terms of algorithms, there are shared parameters between the convolutional layers of the CNN. CNN is used to transfer images from data and analyze them, and then learn them through multiple hidden layers. If someone does not comply with the rule of wearing masks in public places, this information is passed through the police and community networks and then the authority gets and records the information about that person. Deep learning framework to build a neural network model for face recognition, this approach is based on CNN. The model will then discriminate between particular facial images that are different from others and design a function that allows the monitor to alert the police. Design a network that is integrated with multilevel image identity recognition. Figure 1 shows the architecture of the system.

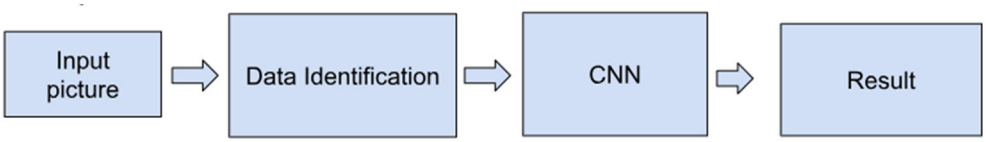


Figure 1. The pipeline of the study

2.2.1. CNN

CNN is the representational algorithm for deep learning, it can classify the input layers by shift-invariant classification [7]. Input, Convolutional, Pooling and Fully Connected layers are composed and widely used in the field of computer vision with the following characteristics: convolutional layer: a CNN uses a convolutional layer to extract features from the input data, and filters the input data and extracts local features through convolutional operations. This enables the CNN to efficiently process the spatial structure of the image. Pooling Layer: CNNs reduce the dimensionality of the data by downsampling through a pooling layer. Common pooling operations include Maximum Pooling and Average Pooling, which help to reduce the model's sensitivity to location while improving computational efficiency [8]. Fully Connected Layers: After multiple convolution and pooling layers, CNNs typically have one or more fully connected layers for mapping high-dimensional features to predictions in the target class. The output of the fully connected layers can be fed into an activation function such as softmax to get the final classification result. The general CNN process includes input, convolution, activation, pooling, and output. Input the raw image into the CNN model. The features of the image are extracted by convolution operation [9]. The output of the convolutional layer is nonlinearized to increase the expressive power of the model. Downsampling the feature map to reduce the data dimension. Mapping the high dimensional features to the category prediction results through

multiple fully connected layers. The results are activated using functions such as softmax to get the final classification results. Through the training process, the CNN model automatically learns the representation of the image features so that it can accurately classify and predict the imageThe CNN imitates the principle of the creature's visual perception. In the module of face recognition, the diverse images are inputted to the input layer, the convolutional layer represents the components obtained by filtering the input images [5]. The polling layer reduces the number of output data by applying the operations [6]. The last part of the full connection transmits signals only to the other full connection layers.

2.2.2. The construction of the model

Model construction involves data selection, data acquisition, data screening and model adaptation to data process. Specifically, firstly, it will collect all the photos of human faces, and will use cameras in public places if a creature enters a public place for collecting faces. To realize this function, design a physical device that can help the monitor to detect the change of heat energy. If the heat in the room would increase incredibly, the camera would take pictures. These photos are then collected at the input. Secondly, the data after acquisition is based on the style and visualization ability of the data. Some weak samples are screened out by manually determining whether they will interfere with the model. Of course, this paper also adds some error cases appropriately to help the model to improve the generalization ability. Third, the model is constructed by introducing CNN to build convolutional layer, pooling layer and fully connected layer. Maximum pooling is used in the pooling stage to capture the key targets in the visual area as much as possible.

2.2.3. Data Identification

This module distinguishes special facial images that are different from other images. First, a generating box function is created to harmonize the faces in the provided annotations for prediction. It then translates the annotations into a list of strings. The different boxes are then converted and labeled as acceptable models.

2.2.4. Loss Function

For the loss function in this human face recognition, the Categorical Cross-entropy loss function is used for configuring the learning process, then it further configures optimizer, and the function will iterate over the data in batches, Finally, the function evaluates model performance. The loss function is as follows [10],

$$I(W_i) = -\sum_{n=1}^{N} \sum_{j=1}^{M} 1\{yn = j\} \log \frac{\exp\left(W_{i,j}^{T} \cdot x_n\right)}{\sum_{m=1}^{M} \exp\left(W_{i,m}^{T} \cdot x_n\right)} \quad (1)$$

The number of samples and subjects are the parameters N and M, x_n is the fused feature image and y_n is the x_n identifier. The parameter W_i represents $[W_{i,1}, W_{i,2}, W_{i,3}, \cdots W_{i,M}]$ coefficient of learning. The parameter T means the transposition of the coefficient of learning.

Nearest Neighbor Algorithm is employed for performing image classification. First, by comparing the features extracted from the query image and the gallery image. Second, the images are classified based on the similarity of the features extracted from the query image with the features extracted from the stored gallery image. In addition, the nearest neighbor algorithm is simple to use. Third, the module finds the nearest matching features required to classify the image through multiple iterations. The loss function chosen for the experiments is the SoftMax function. However, the processing time is increased due to the need for multiple iterations to classify the image. Training the system using the SoftMax loss function and iterating over the samples to recognize individuals based on probability scores achieves higher accuracy and shorter processing time.

2.3. Implemented details

By running the model, several crucial parts are underscored. Firstly, the photos to coordinate parameters; Then the model will locate the human face and the organs on the face. If the result is generated near to "1", the person has a high possibility of wearing the face mask; if the result is generated near to "0", the person has a low possibility of not wearing the face mask. The "pass" rate is almost 93% for the given data.

3. Result and discussion

In the conducted study, a complex model that consists of a variety of images of human faces, including different races and ages. As shown in the Figure 2, it finds that there are almost 93% of visitors who entered public places wore their face masks. However, there were seldom people who did not follow the rule to wear their masks when they were in public places. That depends on the validation accuracy of the codes, since the value of accuracy of the images was close to but less than 1, there exists an error rate for the functions.



Figure 2. The performance curve of the model

By calculating the loss function, there is still some uncertainty between the data and the results of the calculation. Error reduction is used as an optimization objective to reduce the empirical risk of the model through the joint action of the loss function and the optimization algorithm. By binarizing the loss and accuracy data and analyzing it in multiple dimensions, this paper can interpret and analyze the data to perform preliminary validation on the trained model based on validation data (given the complexity of the data and the time required for lengthy processing).

Furthermore, as shown in Table 1, by analyzing the last five calendar elements it can be noticed that the classification loss is decreasing due to the increasing number of training calendar elements. The experimental results demonstrate that the model avoids both overfitting and underfitting of the data. This also verifies that the proposed model can be appropriately generalized to classify new test data. It subsequently achieves an accuracy of 93% in testing. The final overall accuracy of the system was approximately 94% and maintained a processing time of 0.2389 seconds per input.

Table 1. The Performance of the Model

Epoch	Loss	Accuracy	Val_loss	Val_accuracy
16	0.1969	0.9411	0.2389	0.9301
17	0.2047	0.9389	0.2114	0.9276
18	0.1949	0.9402	0.2125	0.9288
19	0.2024	0.9395	0.2169	0.9288
20	0.2056	0.9374	0.2089	0.9301

4. Conclusion

This research project proposes a method to analyze face coordinate data and establish face spatial analysis based on CNN. In order to evaluate the proposed method a large number of experiments are conducted and show that the proposed model can effectively analyze faces and perform accurate recognition. In the future, the performance of the model will be across diverse facial micro-features as the next phase of research. The research will focus on recognizing the complexity of face images for face recognition. Although CNN-based face recognition has some drawbacks such as low speed of processing datasets and outputs, which are common to CNN models. There is also a lack of validation datasets to improve the accuracy of the predictions. For future iterations, the trend of iterations is related to the advantages of the current industry trends and the development prospects are related to controlling the number of iterations, which means reducing the risk of overfitting or underfitting as well as the loss of data. Therefore, the model needs to be refined to minimize overfitting or underfitting by adding preprocessing layers, which include convolutional layers, pooling layers, etc.

References

[1] Vadlapati J Senthil V and Ewin V 2021 Facial recognition using the OpenCV Libraries of Python for the pictures of human faces wearing face masks during the COVID-19 pandemic 12th International Conference on Computing Communication and Networking Technologies (ICCCNT).

[2] Wang J and Li Z H 2018 Research on face recognition based on CNN IOP Conference Series: Earth and Environmental Science 170 (3).

[3] Kaur G et al. 2022 Face mask recognition system using CNN model Neuroscience Informatics 2 (3) p: 100035.

[4] Xie Z M Li J J and Shi H 2019 A Face Recognition Method Based on CNN Journal of Physics: Conference Series 1395 (1).

[5] Younis K S and Abdullah a A 2017 A new implementation of deep neural networks for optical character recognition and face recognition Proceedings of the new trends in information technology pp: 157 - 162.

[6] Silwal R Alsadoon a Prasad P W C Alsadoon O H and Al-Qaraghuli A 2020 A novel deep learning system for facial feature extraction by fusing CNN and MB-LBP and using enhanced loss function Multimedia Tools and Applications 79 pp: 31027 - 31047.

[7] Li Y Wang Z Li Y Zhao X and Huang H 2020 Design of face recognition system based on CNN In Journal of Physics: Conference Series 1601 (5) p: 052011.

[8] Rakhra M Singh D Singh A et al. 2022 Face Recognition with Smart Security System 10th International Conference on Reliability Infocom Technologies and Optimization (Trends and Future Directions) pp: 1 - 6.

[9] Ning X Xu S Nan F et al. 2022 Face editing based on facial recognition features IEEE Transactions on Cognitive and Developmental Systems.

[10] Li Y Lu Z Li J and Deng Y 2018 Improving deep learning feature with facial texture feature for face recognition Wireless Personal Communications 103 pp: 1195 - 1206.