MASKED FACE RECOGNITION USING CONVOLUTIONAL NEURAL NETWORK

The focus of this work rests upon developing face recognition accuracy within different types of masks. Many of the governments across the world are also interested in the face recognition system to secure public places such as parks, airports, bus stations, and railway stations, etc. In our daily life activities like passport checking, smart door, access control, voter verification, a criminal investigation, and many other purposes face recognition is widely used to authenticate a person correctly and automatically. Face recognition has gained much attention as a unique, reliable biometric recognition technology that makes it most popular than any other biometric technique likes password, pin, fingerprint, etc. During current years, deep learning has obtained numerous breakthroughs in several computer vision areas, such as object detection, object classification, object segmentation, and of course, face detection and verification. That's why the masked face is being one of the majors concerning factors within the domain of face recognition.

Their approach consists of three principal modules: Detecting a face from a given image, extract features, and finally recognition. With the scarcity of a large number of images, we apply the data augmentation process on masked and non-masked face images available in the database to enlarge the dataset images, so that our work is more reliable and efficient. Non-Maximum Suppression (NMS) is applied in the MTCNN model to refine the candidate bounding boxes suggested by the first stage P-Net, the second stage R-Net and the third stage O-Net network before providing output. Qiao proposed Multi-task Cascaded Convolutional Neural Network (MTCNN) to detect and align a face from an input image which can outperform many face-detection benchmarks while holding real-time performance. After face detection, crop and resizing methods are applied to input images. The bounding box founded in the face detection level is produced by the MTCNN model. The bounding box is then used to crop the face portion from the input images. Lastly, this verification process is consolidated to recognize candidates face by performing the classification task within a unified Support Vector Machine (SVM). CNN can obtain different local essential features from the data, can select global training components, and has been successfully implemented in many disciplines of pattern recognition applications. Deep CNN networks are typically trained on large labeled datasets like ImageNet to pick out general features which are appropriate for several detections and recognition jobs like image classification and verification, object detection, segmentation to texture identification

In later work, it is our importance to enhance and enlarge our work to address different extreme masks condition of face recognition. So FaceNet model trained on masked and non-masked images gives better accuracy for simple masked face recognition. In this work, the FaceNet pre-trained model has been used for improving masked face recognition. At the same time Industry, 4.0 and/or Sustainable Technology will try to enhance computer adoption and mechanization with autonomous and intelligent systems fed by data and machine learning.

In this work different combinations of masked and non-masked face images are made to find out the best recognition accuracy. The larger area under the curve shows both a high precision rate and a high recall rate. In the ROC curve, the X-coordinate defines the false positive rate, and the Y- coordinate defines the true positive rate. The ROC curve presents the changes of recall vs precision correlation with the variation of the threshold value for defining a positive in our model. For various thresholds, the precision-recall curve has been shown different changes among precision and recall Receiver Operating Characteristic (ROC) curve is another visualization method which shows the actual performance of a classification model.

The primary concern to this work is about facial masks, and especially to enhance the recognition accuracy of different masked faces. An abundant number of researches works have been performed for recognizing faces under different conditions like changing pose or illumination, degraded images, etc. Besides, its performance has been also evaluated within excessive facial masks and found attractive outcomes. Experiments signify that this mentioned approach gives a remarkable performance on masked face recognition.

Face Detection and Recognition Based on Visual Attention Mechanism Guidance Model in Unrestricted Posture

Since the traditional method needs to design different artificial features for different tasks, such as grayscale features, contour features, and HOG features, these features are easily affected by the imaging angle, and the generalization ability is poor. There are still some challenging problems in face detection and recognition technology mainly due to the nonrigid features and the influence of complex background. Compared with detection and recognition algorithms for shallow learning models such as boosting, decision trees, and neural networks, deep learning represented by convolutional neural networks implements the deep nonlinear network structures through operations such as local receptive fields and weight sharing. By directly extracting features from the detection area and then using machine learning algorithms to classify and recognize, the accuracy of the model classification can be improved to a certain extent, but the characterization ability of features directly affects the recognition accuracy of the system. The face detection and recognition model based on machine learning is a popular research direction in the field of computer vision.

To solve this problem, YOLO-V2 introduces the idea of the anchor mechanism in the Faster R-CNN network and uses the k-means clustering method or fuzzy c-means method to generate a suitable prior bounding box. Multiple bounding boxes are predicted in each grid, and each predicted bounding box is scored to demonstrate that the bounding box completely contains the confidence of the object, which is defined as follows: C ? Pr(object) × IoUtruth pred, Pr(object ∈ { } 0, 1), where Pr(object) indicates the probability of the object contained in the bounding box. The YOLO-V3 network is a better deep learning model in the field of object recognition, the network evolved from the YOLO and YOLO-V2 networks. Besides, YOLO-V2 also introduces batch normalization, dimensional clustering, fine-grained features, multiscale training, and other strategies; compared with YOLO, YOLO-V2 greatly improves the detection accuracy. By using multiscale prediction to detect the 2 Scientific Programming final object, its network structure is more complicated than YOLO-V2.

The input image is extracted by the feature extraction network to extract high-level semantic features, and the feature-guided attention module is used for feature fusion; the face analysis network predicts the face position, height, and offset heat map based on the obtained high-level semantic feature and obtains the face boundary box. The process can be described as follows: firstly, the number of C3 and C4 feature channels can be reduced to 256 by the convolutional layer with the size of the 1 × 1 convolution kernel, which can reduce the amount of calculation; then, the backbone network feature maps (namely, P3 and P4) after bilinear interpolation and upsampling 2 times are, respectively, input into the guided attention module to feature fusion. Finally, the multitask loss function weighted optimization can be adopted to train our proposed network, where the weighted loss function can be denoted as follows: F′ C H W Average pooling Max pooling Fully connected layer Attention vector Figure 4: Feature channel attention module structure.

A large number of simulation experiment results show that our proposed method is superior to other comparison algorithms for the accuracy of occlusion face detection and recognition on the face database. Since the useful information for partial-occluding face objects is usually obscured by the background, the network also needs to determine the spatial location of the useful information while enhancing the feature expression of the occlusion object through the channel attention module. Compared with the baseline model, the face detector's missed detection (MR) of the occlusion image has a significant decrease after adding the attention module, indicating that our proposed attention mechanism can effectively guide the detector to focus on the occlusion object. To qualitatively and quantitatively analyze the accuracy of the proposed algorithm for occluded face detection, this paper selects the comparison algorithm due to the improvement of the face detection accuracy of the attention perception fusion module and the use of the multiscale pyramid pooling layer to capture high-level semantic features.

To improve the accuracy of face detection and recognition, a visual attention mechanism guidance model is proposed in this paper, which uses the visual attention mechanism to guide the model to highlight the visible area of the occluded face.YOLO-V3 can predict the bounding boxes of different scales, which can detect small objects more effectively than YOLO-V2, but there are still missing detections for partially occluded face objects. Since the useful information for partial-occluding face objects is usually obscured by the background, the network also needs to determine the spatial location of the useful information while enhancing the feature expression of the occlusion object through the channel attention module.