IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AN OVERVIEW ON FACE-MASK DETECTION USING DEEP LEARNING

Prof. P.U. Tembhare, Abhishek Chandankhede, Nitin Sonekar , Sarang Rohankar , Sahil Kothekar
Assistant Professor at PCE ,Student at PCE , Student at PCE, Student at PCE , Student at PCE
Department of Computer Technology,
Priyadarshini College of Engineering , Nagpur , India

ABSTRACT

Face mask has become essential thing in our day to day life as the world suffering with corona pandemic. To avoid getting affected by this virus the first and the main step is to wear a face mask. So this paper is all about studying various techniques to build a face mask detector. To detect the mask ,first task is to detect the human face in the image/video. The technique of detecting the human face is consist of basic two steps - finding location and size of human face. Face detection represents the first step in expression recognition , facial features detection i.e. fully automatic face recognition.

1.Introduction

In this review paper , we will discuss about some proposed models that has been made for detecting face, face mask , facial features etc. While discussing about each model , we will see what techniques/algorithms they had used in order to implement the idea . And in the last, we will com pare all techniques by taking look on the methods/algorithms they had used. Papers are discussed in same manner as we wrote in Reference section.

2. Discussion

The Project "Face and facial feature detection" was done by M.V.Choudhar ,M.S.Devi & P.Bajaj. In their project, they used Face detection ,Skin color classification and feature extraction techniques .There are three techniques for face detection - "Knowledge based , Image based and Feature based". Knowledge based technique uses rules about human facial features i.e. 2 eyes symmetrical to each other, nose, mouth. After detection of this features verification process is used to reduce false detection .It is only good for frontal images. In Image based technique , a data is trained with large set of images of human faces and then the input image is compare with this data set so as to detect the face region . This technique has high detection rate but still it is slower than the Feature based technique . In the Feature based technique , we extract the facial features that can not affected by variations in lighting conditions , angle of occlusion and other factors .Among this techniques, "Feature based technique" has highest detection rate. From all the facial features , skin color is an effective feature used to detect the faces as color processing is much faster than processing other features. Their model used following techniques - Face detection , Face detection using YCbCr color space , Eye detecting using Hough Transform. To detect faces in color images with complex background "Feature based" face detection techniques can use skin color information . In this process ,first we have to classify pixels of an image as skin and non-skin pixels . The skin detector detects whether the specific region in the image human skin or not. As people have different color appearance , it is important to select the color space that will be use in modelling skin color . But differences occurs due to changes in luminance not by chromaticity . So the accuracy can be achieve by dropping out this luminance factor . The RGB component converts to YCbCr component .Here in this component luminance information is in Y component & color information in Cb, Cr component . Now the system explicitly separates this both components and uses only color components i.e. Cb,Cr. For the model they collected many samples of human skin with different color and now the Cb Cr values will be calculated . If this values of pixels are 132-173 for Cr and 76-126 for Cb, then this pixels will be skin pixels. Now to detect the eyes ,they used Hough transform technique . Hough transform is use to detect the shapes like line, square , rectangle, etc. But in this method , as the image size increases the processing data of the quality will decrease . So Circular Hough transform is used to detect the circular images.

In the paper “Detecting Masked Faces in the Wild with LLE-CNN” by Shiming Ge, Jia Li, Qiting Ye, Zhao Luo created large data set which they named MAFA (Masked Faces) which contained many masked and non-masked face images by different degree of occlusion such as “Left , Left Front , Front, Right Front, Right”. They define four categories of masks that can be frequently found in Internet images, including: Simple Mask (man-made objects with pure color), Complex Mask (man-made objects with complex textures or logos), Human Body (face covered by hand, hair, etc.) and Hybrid Mask (combinations of at least two of the aforementioned mask types, or one of the aforementioned mask type with eyes occluded by glasses).These paper consist if large dataset Which they named as MAFA .These dataset consist of about 30000 images while training the model .They

proposed LLE-CNNs technique, which consists of three modules Proposal module, Embedding module and Verification module". The proposal module extracts face proposals and characterize them with noisy descriptors. After that, the embedding module transfers the noisy descriptors into similarity-based vectors by embedding the proposals into the feature subspace formed by representative normal faces and non-faces. Finally, the verification module jointly perform regression and classification tasks so as to refine the location and scale of a proposal and identify whether it is a real face .They also have done experiment on existing face detection models that can be roughly grouped into three categories, in which models are based on boosting, Deformable Part Model (DPM) and Convolutional Neural Network (CNN). CNN-based face detectors directly learn face representations from data. It adopts deep learning paradigm to detect the presence of a face in a scanning window. This method demonstrates strong capability in detecting faces with severe occlusion and pose variation. Zhu et al. presented Contextual Multi-Scale Region-based Convolution Neural Network (CMS-RCNN) for face detection in unconstrained conditions. Ranjan et al. presented Hyperface for simultaneous face detection, landmarks localization, pose estimation and gender recognition by using CNNs. From these 3 categories they selected "SURF and NPD models from Boosting based category" , "ZR,HH,HPM models from Deformable Part Model (DPM)" and "MT from CNN based detector ".After this experiment they noticed that from among above techniques , the overall performance of MT is very impressive compared with other approaches.

In the paper "Multi angle head pose Classification of when Wearing the mask for face Recognition under the COVID 19 Epidemic" by Shuang Li ,Xin Ning ,Lina Yu,Xioli Dong ,Yuan Shi , the method of facial landmark detection divides the head pose classification problem into two stages. First is algorithm of facial landmark detection to calibrate the facial image and obtain 68 or 98 feature points of the image and second one is determining the locations of 3D points that are depending on 2D projections in the image. Calculating three Euler angles to get head pose information is the last step. The colour space of pictures used by researchers is generally in RGB. CNNs used to extract features, which are also based on the RBG colour space. Comparing to RGB, HSV is better at expressing colours. So, we use HSV colour channel to process masked facial images. Only the H-channel of the HSV is processed. This are some techniques they followed – HGL: In this method they combined the H-channel of the HSV colour channel with the face portrait and grayscale image, and train the CNN to extract features for classification. Line: Here we generate the image as line portrait and then input it into CNN for training. RGB: The input image in RGB colour space without any processing into the CNN. FSA-Net: This method uses FSA-Net to learn the structure of a single RGB colour space image. As the output of this network is the Euler angle of the head pose for that we have selected set of thresholds that are most effective for pose classification. EPnP-LAB: Here we use the method of combining facial landmark detection with EPnP algorithm. Here we use LAB, LAB algorithm uses 98 points to represent the geometric structure of the face. After use of all methods best accuracy is given by HGL method that is front accuracy 93.64% and side accuracy 87.17%. Reason behind this result is that the images processed through the H-channel can provide focused information to the net. By this step we can ignore the influence of mask and highlight the information of the face itself. After that it combined with the contour information of the face in the portrait and then the pixel contrast information in the grayscale image to better complete the classification task.

In the paper "Masked Face Recognition Using Convolutional Neural Network" by Md. Sabbir Ejaz & Md. Rabiul Islam , the approach for the implementation of project consists of three principle modules - Detecting face from a given image, extract features and finally face from a given image, extract features and recognition. Image acquisition implies the first key stage regarding the face recognition method. They applied data augmentation process on masked and non -masked face images available in the database to enlarge the dataset images so as to get work more reliable and efficient. For real-time performance , MT-CNN is used to to detect and align face from an input image. So they used a pre-trained MTCNN model to detect candidate's masked and non-masked facial area **.** There are three networks cascading in this model. In the process called Image Pyramid, the model first rescaled the image to a certain extent of sizes. Then the candidate's facial area is introduced by the first network i.e. P-Net or Proposal Network. The second network known as R-Net or Refine Network refines the bounding . And in the last, the third network O-Net or Output Network can perform classification of face . O-Net determines facial landmarks from P-Net , R-Net . They extracted the feature using Face Net methodology. FaceNet is the start-of-art for face recognition identification, verification and clustering of neural networks. A large amount of masked and non-masked face images are used to train the FaceNet model. At last the verification process is done to recognize candidate's face by performing the classification task within a unified Support Vector Machine (SVM).

3. Parameter Comaparison

Parameters	**Paper 1**	**Paper 2**	**Paper 3**	**Paper 4**
Data Set	Yes	Yes	Yes	Yes
Hough Transform	Yes	No	No	No
CNN	No	Yes	Yes	Yes
FaceNet	No	No	No	Yes
MT-CNN	No	Yes	No	Yes
Color Extraction techniques (RGB,YCbCr)	Yes	No	Yes	No

4. Conclusion

From the study of all the above papers , we came to the conclusion that there are many techniques to detect faces , like using Hough transform, Color Extraction techniques (RGB,YCbCr). But as we use this methods, complexity increases. To get good accuracy simply we can use large dataset . So as to keep the project simple and efficient ,using MT-CNN with large dataset is better option .

REFERENCES

[1] M.V. Choudhari, M.S. Devin & P. Bajaj Face and Facial Feature Detection https://dl.acm.org/doi/proceedings/10.1145/1980022

[2] Ge, S., Li, J., Ye, Q., & Luo, Z. (2017). Detecting Masked Faces in the Wild with LLE-CNNs. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). doi:10.1109/cvpr.2017.53

[3] Shuang Li ,Xin Ning ,Lina Yu,Xioli Dong ,Yuan Shi. Multi angle head pose Classification of when Wearing the mask for face Recognition under the COVID 19 Epidemic.

[4] Ejaz, M. S., & Islam, M. R. (2019). Masked Face Recognition Using Convolutional Neural Network. 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI).