A Survey on Multi-angle Head Pose Classification when Wearing the Mask for Face Recognition under the COVID19 Coronavirus Epidemic

IYRENE P SAM

[1]M. Tech, Department of Computer Science and Engineering, Sree Buddha College of Engineering, Pathanamthitta, India.

Abstract - *Face detection and Face Recognition are often used interchangeably but these prove to be quite different. In fact, Face detection is simply a part of Face Recognition. Face recognition is a method of identifying or verifying the identity of an individual using their face mask. The Face Mask Recognition works in real-time and provides instantaneous warnings of its detection. The use of masks in daily life has become an unavoidable practice in today's world. Masked faces have been common for some or other reasons but the regular and mandatory use of masks commenced on the outbreak of the SARCS-Covid virus and therefore the detection of face masks has also become important. Face recognition with and without occlusions and its various techniques had been an area of research for a long time whereas the multi-angle classification of the face and the detection of the presence of masks is yet another research area.*

Key Words: Face Recognition, Face Detection, Artificial Intelligence, Deep Learning, Convoluted Neural Network.

1. INTRODUCTION

Facial Recognition can be said as an area of study research and development which deals with giving machines the ability and capacity to recognize and verify human faces. Detection of Face mask detection works in real time and gives an instantaneous warning of its detection. And there is no need for the camera to be pointed directly at the face for the detection to be working.

To deal with the current pandemic situation people wear masks on a daily basis and have become a strict regulation to be followed for a long time. The related problems of head pose estimation and facial expression tracking have played an important role over the past 25 years enabling new ways to manipulate multimedia content and interact with users.

Due to the great success of these face detectors, some of them have been integrated into applications so as to facilitate auto-focusing, human-computer interaction, and image database management. The detection of masked faces, which can be very helpful for applications like video surveillance and event analysis, is yet on the path of innovation and improvement.

For a data set to be useful for evaluating face detection, the locations of all faces in these images need to be annotated. In the annotation process, one must ensure that each image contains at least one face occluded by various types of masks, while the six main attributes of each masked face, including locations of faces, eyes and masks, face orientation, occlusion degree, and mask type, are manually annotated and cross-checked by nine subjects. The Survey on face mask recognition will help to know the different techniques used and the prospects of future methods to be implemented.

2. LITERATURE SURVEY

Face Recognition is an area of a recent study which deals with giving machines the ability to recognize and verify human faces. Face Mask Detection works in real-time and gives instantaneous warnings of its detection. It is a study of Artificial Intelligence where it is known as Deep Learning comes into play. It is important to consider the speed of the system so that the real-time object detection can be precise and accurate.

Shuang proposed a system to solve the problem of head pose classification with masks during the COVID-19 coronavirus epidemic. The method uses the colour texture analysis of images and line portraits. Here, the colour space of pictures by researchers to solve the problem of head pose classification is RGB. Compared to RGB, HSV is better at expressing colours. It can very intuitively reflect the hue, vividness, and shading of colours, which is convenient for comparing different colours and therefore the head pose classification with masks based on colour texture analysis, which uses the HSV colour channel to process masked facial images was posed. The dataset used is MAFA. The dataset contains some facial images with and without a mask. There are 23,845 images in total, in which 20,139 images of the training set and 3,706 images in the test set. The information of facial location provided by the MAFA dataset to crop the image is used first. Then, the size of the facial frame is increased to 1.2, 1.4, and 1.6 times the original size for cropping, and normalizes these facial images to 80×80. Then, the h-channel in the HSV colour space of the image is extracted, and normalize to the pixel values of the image in the range of 0-255. The 4×4 kernel is used to filter the extracted image and use the pixel average as the threshold to binarize the image. Finally, the processed image is merged with the portrait image.

Toan Minh Hoang suggested a system that focuses on the challenging tasks of face detection to deal with small size 12 pixels and occlusions thanks to mask or any other. It is an extended feature pyramid network (FPN) module that can detect small faces by expanding the range of P layer, and therefore the network by adding a receptive context module (RCM) after each predicted feature. Based on the Feature Pyramid Network principle, the combination between the low- and high-resolutions are beneficial for object detection especially with objects of different sizes. Furthermore, from the RCM, the method can make use of wide range of context information for small faces. The performance is evaluated by using various public face datasets such as the WIDER Face dataset, the face detection dataset and benchmark (FDDB), and the masked faces (MAFA) datasets having challenging samples such as small face regions and occlusions by hair or other people. The results proved that deface can detect the face region more accurately in comparison to the other state-of-the-art methods while maintaining the processing time. The backbone of defacing are often any conventional deep CNN for feature extraction like VGG-16, RESNET, or SSD.

Fanelli G, Weise T, and Gall J proposed a system for real-time 3D head pose estimation robust to the signal-to-noise ratio of current consumer depth cameras. This is a discriminative random regression forest, that classifies the depth image patches between the highest and thus the rest of head and therefore the remainder of the body and which also performs a regression within the continuous spaces of head positions and orientations. Two existing methods were presented for combining such measures and a 3rd weighting scheme was introduced which favored the regression measure as an exponential function of the node depth. Also compared the proposed methods and observed similar performances in terms of accuracy.

Mukherjee S.S and Robertson N brought a replacement idea of implementing deep head pose classification alongside gaze-direction estimation, especially in multimodal videos. Here, a convolutional neural network (CNN)-model for head pose estimation in low-resolution multi-modal RGB-D data is used. The issue is addressed together of the classifications of human gazing direction and further Fine-tune a regressor supported by the learned deep classifier. Then merge both the models (classification and regression) so on to estimate the approximate regression confidence presented and the state-of-the-art classification leads to datasets which will span the range of high-resolution human-robot interaction data to challenging low-resolution outdoor surveillance data. Further, build and introduce a replacement visual attention model to recover interaction with the environment.

V. Jain and E. Learned-Miller proposed a system during which a replacement dataset of face images with more faces and more accurate annotations for face regions than in previous data sets is presented. Berg created a knowledge set that contains images and associated captions extracted from news articles. The pictures due this collection display large variations in pose, lighting, and background and appearance. Benchmarking face detection algorithms on this dataset hopes to supply good estimates of their expected performance in unconstrained settings.

S. Zhang, C. Chi, Z. Lei, and S. Z. Li proposed a state-of-art single-shot face detector that has been enhanced by the regression and classification ability. Here, the system enlightens the RFE to supply a more diverse receptive field which helps to capture faces better in some extreme poses. Refine Face may be a baseline face detector Retina Net and has five modules. The architecture may be a lightweight architecture that is with the assistance of automatic machine learning (AutoML) methods in order that the system can run in real-time in GPU devices and also in CPU devices within the future.

J. Wang, Y. Yuan, and G. Yu proposed a face detector called face attention network (FAN). This improves the recall of the face detection problem also within the occluded case without compromising the speed. The team comes with new anchor-level attention which can highlight the features from the face and prevents false positives. The datasets used are Wider Face and MAFA. The FAN system when compared to the state-of-art detectors like SFD, SSH, HR, and Scale Face showed that it outperformed state-of-art detectors.

Ge S, Li J, and Ye Q suggested a system that detects masked faces with LLE-CNNS. LLE is a Locally Linear Embedding algorithm. The first module combines two pre-trained CNNs to extract candidate facial regions from the input image and represents them with high dimensional descriptors. The Embedding module is then incorporated to turn such descriptors into a similarity-based descriptor by using the LLE algorithm. In the end the Verification module is incorporated to identify candidate facial regions and refine their positions by jointly performing classification and regression within a CNN which is unified. The dataset preferred here is the MAFA.

Boulkenafet Z was successful in proposing a system that faces spoofs and detects using color texture analysis. The color texture information is obtained from the luminance and the chrominance channels. The databases considered here are the three latest face anti-spoofing databases: CASIA Face Anti-Spoofing Database (CASIA FAS D), Replay-Attack Database, and MSU Mobile Face Spoof Database (MSU MFSD).

Ruiz N, Chong E, and Rehg J M suggested a new system that can determine the pose by training a multi-loss convolutional neural network on 300W-LP, a large synthetically expanded dataset, to predict intrinsic Euler angles (yaw, pitch, and roll) directly from image intensities through pose classification and regression. Synthetic data generation for extreme poses proves to be a better way to improve performance. This method has proven to outperform networks that regress head pose as a sub-goal in detecting landmarks.

3. CONCLUSION

The suggestions proposed by the works of these people provide a deep knowledge into the field of face recognition and detection. Artificial Intelligence and its subsets provide a wide range of improvements and changes in the algorithms and technologies so that it can be used to build an entirely new work in the field of face recognition and face and face mask detection. For example, the emergence of technologies like SSD, Yolo, Dense Net has thrown a light on the improvision of existing algorithms like inception V3 and more.

REFERENCES

1. Toan Minh Hoang, Gi Pyo Nam, Junghyun Cho, And Ig-Jae Kim "Deface: Deep Efficient Face Network For Small Scale Variations".
2. Shuang Li1, Xin Ning, Lina Yu, Liping Zhang, Xiaoli Dong, Yuan Shi , Wei He "Multi-Angle Head Pose Classification When Wearing The Mask For Face Recognition Under The Covid19 Coronavirus Epidemic".
3. Fanelli G, Weise T, Gall J, Et Al., Real Time Head Pose Estimation from Consumer Depth Cameras[J], 2011.
4. Mukherjee S S, Robertson N, Deep Head Pose: Gaze-Direction Estimation in Multimodal Video, IEEE Transactions on Multimedia, 17(11): 2094-2107, 2015.
5. V. Jain and E. Learned-Miller, ''Fddb: A Benchmark For Face Detection Unconstrained Settings,'' Univ. Massachusetts Amherst, Amherst, Ma, Usa, Tech. Rep. Um-Cs-2010-009, 2010, Pp. 1–11.
6. S. Zhang, C. Chi, Z. Lei, and S. Z. Li, ''Refineface: Refinement Neural Network for High Performance Face Detection,'' Ieee Trans. Pattern Anal Mach. Intell., May 2020, Doi: 10.1109/Tpami.2020.2997456.
7. J. Wang, Y. Yuan, And G. Yu, ''Face Attention Network: An Effective Face Detector For The Occluded Faces,'' 2017, Arxiv:1711.07246. [Online].
8. Ge S, Li J, Ye Q, Et Al., Detecting Masked Faces In The Wild With Lle-Cnns, IEEE Conference On Computer Vision & Pattern Recognition, 2017.
9. Boulkenafet Z, Komulainen J, Hadid A, Face Spoofing Detection Using Colour Texture Analysis, IEEE Transactions On Information Forensics & Security, 11(8):1-1, 2016.
10. Ruiz N, Chong E, Rehg J M, Fine-Grained Head Pose Estimation Without Key points, 2017.