Facial Mask Detection Based on Template Matching and Semantic Segmentation

[1]Darshna Dalvadi

[1]Research Scholar
[1]Computer Science,
[1]C. U. Shah University , Wadhawan, India

Abstract: Facial mask detection can be used for ATM centers where it is prohibited to wear a mask, in any situation where there is a need to identify the facial mask this proposed method can be used. Proposed method uses template matching algorithm to detect the facial mask. Many times when person are visiting hospital it is made compulsory to wear a mask to protect themselves from any contagious disease. In such scenario template matching based facial mask detection algorithm can be used. This paper propose the algorithm and also provides the review of various techniques.

***IndexTerms* – Image Restoration, filter, machine learning, deep learning, fuzzy**

I. INTRODUCTION

Face recognition has emerged as a fascinating subject in image processing and computer vision. It has a wide range of applications, from facial motion capture to face identification, which requires the face to be recognized with high accuracy at first. Face detection is becoming increasingly important today because it is utilized not just on photos but also in video applications such as real-time surveillance and face detection in videos. With the developments of Convolutional networks, high accuracy picture classification is now attainable. After face detection, pixel level information is frequently required, which most face detection algorithms cannot offer[1].

Template Matching is a high-level machine vision algorithm that detects image portions that match a predetermined template. Advanced template matching algorithms enable the detection of template occurrences regardless of orientation or local brightness. Template Matching approaches are versatile and simple to apply, making them one of the most used ways of object localization. Their applicability is mostly restricted by available processing capacity, as identifying large and complicated templates can be time-consuming. Following Figure 1 shows that single apple is the template image and second is target image and third image shows the box around apple without green leaf. So this shows the basic limitation of template matching algorithm which is it can only detect the image which is provided as template.

Fig.1 template image, target image and detected image.

Template matching uses the concept of normalize cross-correlation. The main concept of template matching is to find small signal from the big signal. Normalized cross-correlation (NCC) is an improved version of the standard cross-correlation method that includes two new features: The results are insensitive to variations in global brightness, i.e. constant brightening or darkening of either image has no effect on the outcome (this is accomplished by subtracting the mean image brightness from each pixel value). The final correlation value is scaled to the [-1, 1] range, so that the NCC of two identical images is 1.0, whereas the NCC of an image and its negation is -1.0.

$$\gamma(u,v)=\frac{\sum_{x,y}\left[f(x,y)-\bar{f}_{u,v}\right]\left[t(x-u,y-v)-\bar{t}\right]}{\left\{\sum_{x,y}\left[f(x,y)-\bar{f}_{u,v}\right]^{2}\sum_{x,y}\left[t(x-u,y-v)-\bar{t}\right]^{2}\right\}^{0.5}}$$

Where f is the image, $\bar{t}$ is the mean of the template, $\bar{f}_{u,v}$ is the mean of f(x,y) in the region under the template.

Semantic segmentation takes image as an input and produces output as regions of interest or structure. Following image shows example of semantic segmentation where first image is given as an input, second image represents structures, curves, lines and regions and third image represents the output of semantic segmentation.

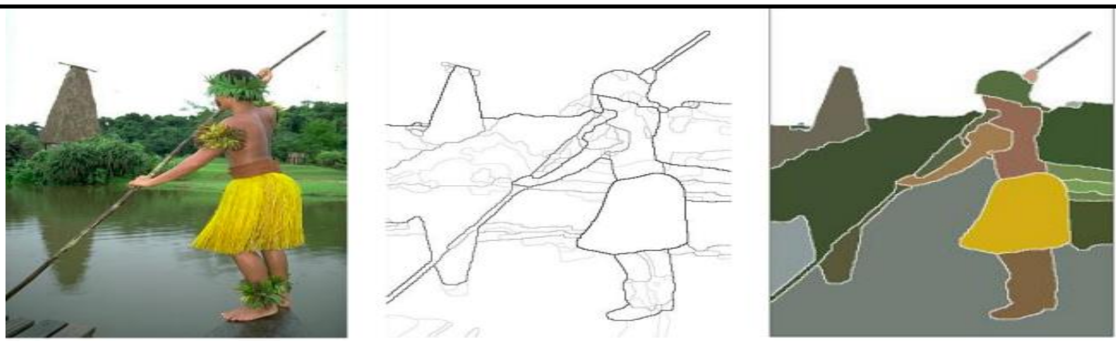

Fig.2 input image, region or structured image and semantic segmented image.

II. Literature Review

This section represent various facial mask detection techniques. The main novelty of paper lies in the combination of NCC with semantic segmentation. Combining template matching with semantic segmentation itself is very interesting area to explore. Paper [2-5] represents the template matching technique for various purpose. Paper [6-9] represents the use of semantic segmentation algorithm. Initially, researchers concentrated on the edge and grey value of the face image. Paper [10] was built on a pattern recognition model using prior knowledge of the face model. Adaboost [12] performed well as a training classifier. The famous Viola Jones Detector [13] was a breakthrough in face detection technology, dramatically improving real-time face detection. Viola Jones detector enhanced Haar [11] traits but failed to address real-world challenges and was influenced by aspects such as face brightness and face orientation. Viola Jones could only discern well-lit frontal faces. It didn't perform well in low-light situations or with non-frontal photographs. Because of these challenges, independent researchers have been working on developing new face identification models based on deep learning in order to achieve better outcomes for various facial circumstances. This paper has constructed our face identification model using the Multi Human Parsing Dataset [16], which is based on fully convolutional networks and can detect the face in any geometric condition, frontal or non-frontal. Convolutional Networks have long been used to classify images. AlexNet [14] and VGGNet [15] are examples of typical topologies that use stacked convolutional layers. AlexNet, with 5 convolutional layers and 3 fully connected layers, won the ImageNet LSVRC-2012 competition, but VGGNet outperforms AlexNet by replacing big kernels with 3x3 multiple kernels sequentially.

III. Proposed system

Proposed system is based on hybrid technique. First of all semantic segmentation is applied on the input image which segments the face and mask in two various layers. Mask layer shows the clear image of mask. For template matching various masks image can be cropped and stored as template image to use as database. The template of mask can be gathered in tilted position. So template dataset which can be prepared with around 500 various masks image in various position. Then the model can be trained to estimate that whether there is mask or not in the input image. Following proposed algorithm can be used.

Step 1: Input the image F(x, y)

Step 2: Perform Preprocessing for noise removal from the input image F(x, y)

Step 3: Perform segmentation

a. Choose the initial threshold value t;
b. Segment the image F(x,y) into two regions F1(x,y) and F2(x,y)
c. Find the mean of the pixel F1(x,y) and denote as m1
d. Find the mean of the pixel F2(x,y) and denote as m2
e. Now consider threshold t=(m1+m2)/2 and then go back to step b
f. Continue till $|t_i - t_{i+1}| <= t$

Step 4: After segmentation get Mask segmented image M(x, y)

Step 5: Apply template matching and apply normalized cross correlation

Step 6: Perform matching and detect and draw bounding box on input image by detecting mask

Step 7: Measure the accuracy

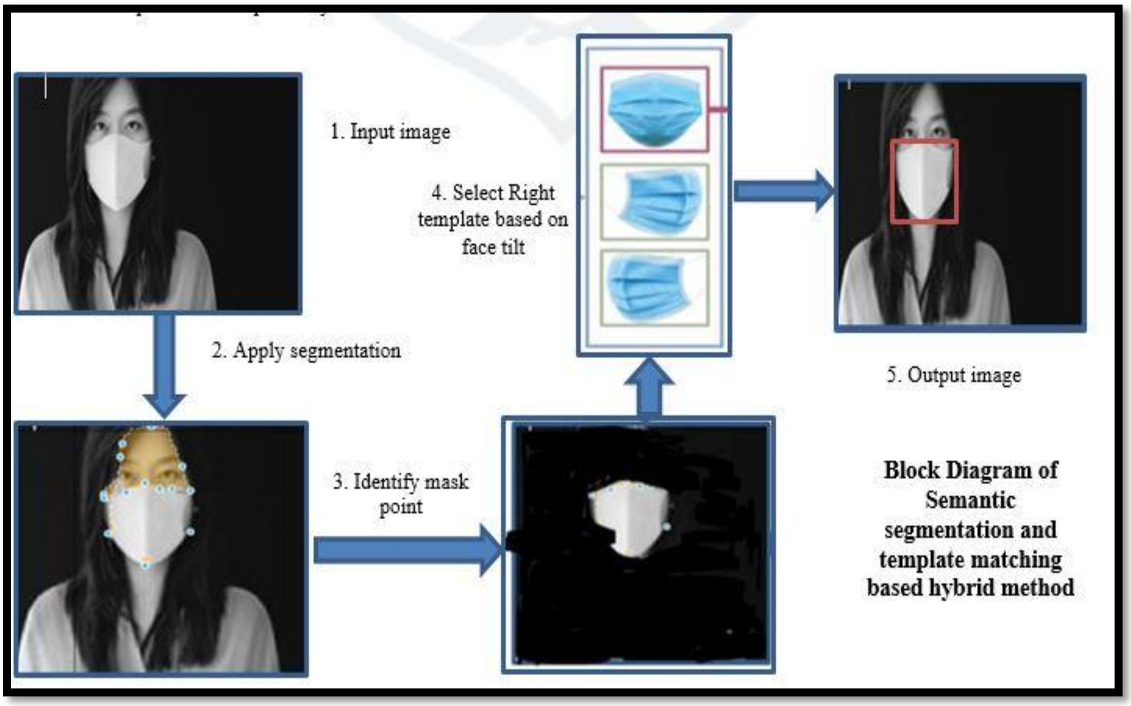

Figure 3: Block diagram of proposed system

IV. Analysis

Accuracy of the system can be checked with false positive and false negative rate. A true positive result is one in which the model accurately predicts the positive class. A real negative, on the other hand, is a result in which the model correctly predicts the negative class.

A false positive is an outcome in which the model forecasts the positive class inaccurately. A false negative is an outcome in which the model forecasts the negative class inaccurately.

There are two possibilities that if there is no facial mask in image but still a bounding box is shown and if there is actually facial mask in the image but system does not show bounding box.

V. Conclusion

This paper propose hybrid method for facial mask detection. This method uses segmentation and template matching algorithm both. Proposed method uses template matching algorithm to detect the facial mask. Many times when person are visiting hospital it is made compulsory to wear a mask to protect themselves from any contagious disease. In such scenario template matching based facial mask detection algorithm can be used. This paper propose the algorithm and also provides the review of various techniques. The main finding of the paper lies in the novel algorithm for facial mask detection which use segmentation and template matching algorithm.

References

1. Liang-Chieh Chen, Alexander Hermans, George Papandreou, Florian Schroff, Peng Wang, Hartwig Adam, MaskLab: Instance Segmentation by Refining Object Detection With Semantic and Direction Features; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2018, pp. 4013-4022
2. Dulari Bosamiya Mr. Akash K. Mehta, A Survey of the Farm Surveillance System for Animal Detection in Image Processing, International Journal Of Engineering Development And Research, 2014
3. Patel Milan, Dulari Bosamiya , Review Of Different Techniques For Ripe Fruit Detection, International Journal Of Engineering Development And Research, 2016
4. RACHANA PATEL, REEVA SONI, DULARI BHATT, Tumor Detection using Normalized Cross Co-Relation, IJRMEET, 2013
5. Bhatt, Dulari, Chirag Patel, and Priyanka Sharma. "Intelligent Farm Surveillance System for Bird Detection." (2011) in GRDET.
6. Masaya Kaneko, Kazuya Iwami, Toru Ogawa, Toshihiko Yamasaki, Kiyoharu Aizawa, Mask-SLAM: Robust Feature-Based Monocular SLAM by Masking Using Semantic Segmentation; Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2018, pp. 258-266
7. Daan de Geus, Panagiotis Meletis, Gijs Dubbelman, Panoptic Segmentation with a Joint Semantic and Instance segmentation Network, arXiv:1809.02110
8. Wang, Yufei and Lin, Zhe and Shen, Xiaohui and Zhang, Jianming and Cohen, Scott Concept Mask: Large-Scale Segmentation from Semantic Concepts Proceedings of the European Conference on Computer Vision (ECCV), September 2018.
9. T. Ojala, M. Pietikainen, and T. Maenpaa, “Multiresolution gray-scale and rotation invariant texture classification with local binary patterns,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 24, no. 7, pp. 971–987, July 2002.

10. T.-H. Kim, D.-C. Park, D.-M. Woo, T. Jeong, and S.-Y. Min, “Multi-class classifier-based adaboost algorithm,” in Proceedings of the Second Sinoforeign-interchange Conference on Intelligent Science and Intelligent Data Engineering, ser. IScIDE’11. Berlin, Heidelberg: Springer-Verlag, 2012, pp. 122–127.
11. P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, vol. 1, Dec 2001, pp. I–I.
12. A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in Advances in Neural Information Processing Systems 25, F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, Eds. Curran Associates, Inc., 2012, pp. 1097–1105.
13. K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” CoRR, vol. abs/1409.1556, 2014.
14. C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, “Going deeper with convolutions,” 2015
15. J. D. Fuletra and D. Bosamiya, “A survey on drivers drowsiness detection techniques,” International Journal on Recent and Innovation Trends in Computing and Communication, vol. 1, no. 11, pp. 816–819, 2013.
16. A. K. Katsaggelos, "Recent trends in image restoration and enhancement techniques," Proceedings of APCCAS'96 - Asia Pacific Conference on Circuits and Systems, 1996, pp. 458-459, doi: 10.1109/APCAS.1996.569312.
17. Bhagchandani, Ashish., Dulari Bhatt and Madhuri Chopade. Various Big Data Techniques to Process and Analyze Neuroscience Data. https://www.researchgate.net/publication/328529465_Various_Big_Data_Techniques_to_Process_and_Analyze_Neuroscience