Face Detection – white paper

Feb 2022

saiwa

simple artificial intelligence web application

FACE DETECTION

Face detection is a specific class of object detection and has always been an important problem for face, expression and gesture recognition. Detecting and localizing human faces in real-world images attract numerous attention between AI researchers from past to present. Recently, with continuous work on neural network technology, deep networks have achieved state-of-the-art results on standard benchmark face datasets. Here, at saiwa we provide two face detection methods:

Dlib [1] is a general purpose toolkit containing machine learning algorithms and tools to solve real world problems. Dlib provides two face detection functions, one for a HOG + Linear SVM face detector and another for the MMOD CNN face detector. Currently on saiwa we use HOG SVM face detector for its simplicity and efficiency. This detector is suitable for detecting frontal faces and may miss a face if it is not perfectly frontal (in this case please try saiwa MTCNN method).

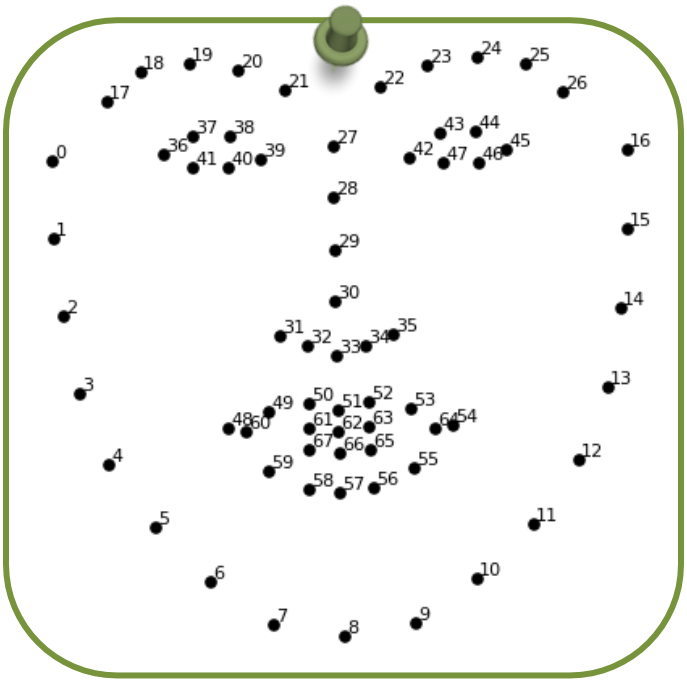

Figure 1. The 68 landmarks that locate on faces (printed from [2]

In saiwa we use a Python implementation of this method. Please refer to [2] for more details. Using this implementation, in addition to bounding boxes around faces, we extract 68 specific points (called landmarks), i.e. the top of the chin, the outside edge of each eye, the inner edge of each eyebrow, etc (Figure 1) [3]. Figure 2 shows a few images with detected or not-detected faces.

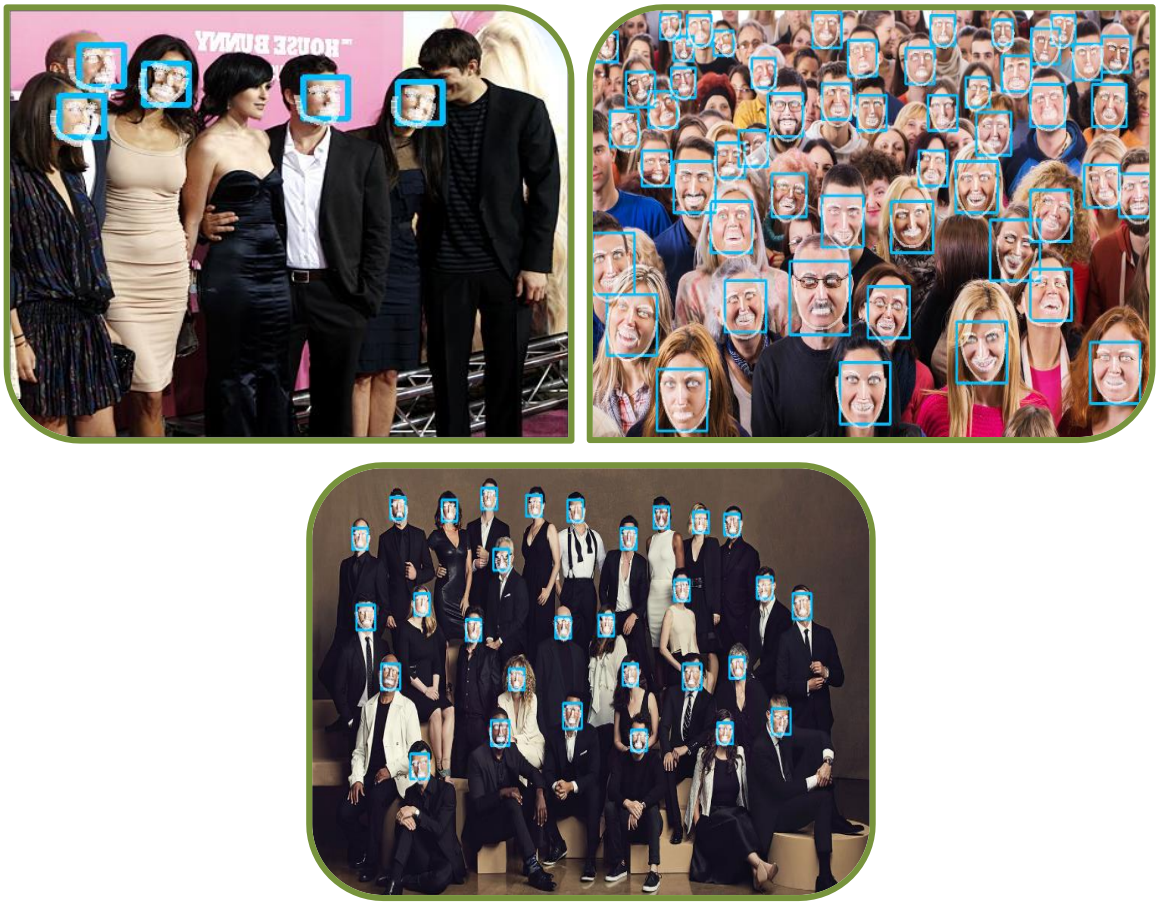

Figure 2. Face detection results using saiwa online interface and HOG + Linear SVM detector.

MTCNN (Multitask cascaded convolutional network) face detector is deep neural network that detects faces even with different sizes, lighting and strong rotations. MTCNN is leveraging a 3-stage neural network detector that predict faces and landmark locations (i.e. eyes, nose and corners of mouth) in a coarse-to-fine manner. In the first stage, it produces candidate windows quickly through a shallow convolutional neural network (CNN). Then, it refines the windows to reject a large number of non-faces windows through a more complex CNN. Finally, it uses a more powerful CNN to refine the result and output facial landmarks positions [4]. Figure 3 shows general architecture of MTCNN.

In saiwa face detection service, we use an open source Python implementation of MTCNN (please refer to [5] for more details). Figure 4 shows results of applying MTCNN on a few images using saiwa interface.

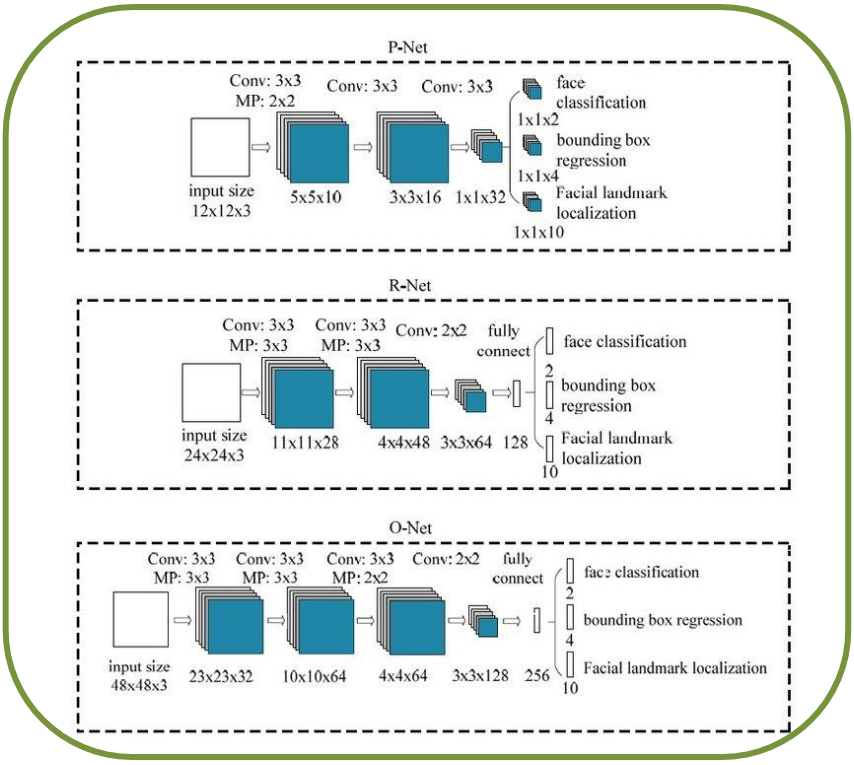

Figure 3. Three stage architecture of the MTCNN model used for face detection and landmark extraction (printed from [6]).

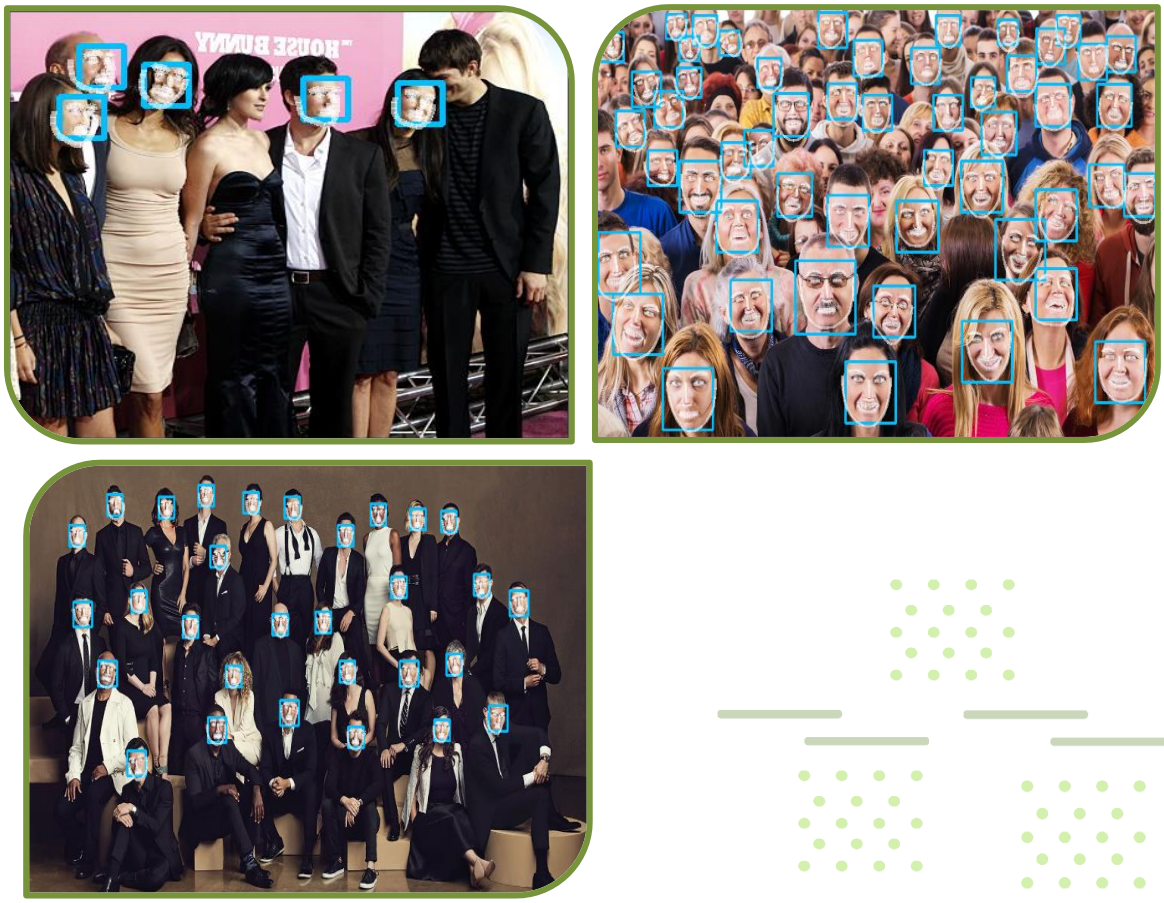

Figure 4. Face detection results using saiwa online interface and MTCNN detector.

References

[1] www.dlib.net

[2] https://github.com/ageitgey/face_recognition

[3] Kazemi, Vahid, and Josephine Sullivan. "One millisecond face alignment with an ensemble of regression trees." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

[4] Zhang, K., Zhang, Z., Li, Z., and Qiao, Y. Joint face detection and alignment using multitask cascaded convolutional networks. IEEE Signal Processing Letters, 23(10):1499–1503, 2016.

[5] https://github.com/ipazc/mtcnn

[6] Yuan, Meng, et al. "Minor privacy protection through real-time video processing at the edge." 29th International Conference on Computer Communications and Networks (ICCCN), 2020.

info@ saiwa.ai

720 Guelph Line Burlington, ON L7R 4E2 Building
+15148131809
www.linkedin.com/company/saiwa
www.instagram.com/saiwa.ai