Smart Intelligent Computing and Communication Technology
V.D. Ambeth Kumar et al. (Eds.)
© 2021 The authors and IOS Press.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/APC210033

Facial Expression Recognition Under Occlusion Conditions Using ACNN

BalasundaramA[a,1], AshokkumarS[b], Saravanan A[c], Sathish Kumar P J [d], BalajiS[d]

[a,1]*Assistant Professor, School of Computer Science and Engineering, Vellore Institute of Technology (VIT), Chennai, India*
[b]*Assistant Professor, Department of CSE, Saveetha School of Engineering (SIMATS), Chennai, India*
[c]*Professor, Dept of ECE, SMK Fomra Institute of Technology, Chennai, India*
[d]*Associate Professor, Department of CSE, Panimalar Engineering College, India*

Abstract. This work proposes a Convolution Neural Network with attention mechanism (ACNN), resembling that human understanding the emotion. Naturally, human perceives the facial emotions depending on particular parts of the face. Several clues are gathered from different parts of face like mouth, lips, cheeks etc, and these are in turn analyzed by the learning model to decide on the facial expression. This work uses ACNN to handle the scenario of occlusion as well.

Keywords. Facial Expression Recognition, FER, CNN

1. Introduction

Recognizing facial expressions (FER) has got huge interest from researchers and analysts over the years, as it caters to a number of uses, for example, human-computer interaction, influence examination, and psychological well-being evaluation. The FER frameworks that acts impeccably in the gathered dataset images, are plausible to act inadequately while perceiving people expressions under normal and abnormal scenarios. To enclose the void between the FER precision in the controlled appearances and uncontrolled portraits, scientists put forth attempts for gathering huge scope facial expressions datasets in nature. In spite of the utilization of information from the wild, outward appearance acknowledgment is as yet testing because of the presence of incompletely impeded countenances. Occlusion might brought about by hair, sun glasses, muffler, mask, beard, arms, moustache, and different items that might be set before the faces in our routine life. These items may hide some portion of cheeks, eye, eyebrows, mouth, and some more facial features

ACNN naturally sees the impeded facial fixes and focuses principally to the unblocked and enlightening patches. Each and every Gate Unit in ACNN learns a versatile load by un-obstructiveness or significance. At that point, the Occluded portrayals are linked and utilized in the classification part. In this manner ACNN can be

[1]Balasundaram A,Assistant Professor, School of CSE, Vellore Institute of Technology, Chennai, India; E-mail: balasundaram2682@gmail.com[1]

accurate in on particular just as unhampered areas in facial picture. Examination on emotion detection is a difficult field that target strategies to make compelling human computer interaction. Face Image signal contains huge amount of important information with mental disorder and emotion of people. Later enhancements here have urged the specialists to expand the pertinence of facial emotion acknowledgment to zones like chat room symbols and video conferencing avatars. Thus identification of emotions from occlusion based images is the current research area. The major problem that starts with this system is that dataset images could be of low quality and our model does not supports that kind of images. Though increased number of epochs increases accuracy, our system gets slower due to load of the model. Also choosing dataset images according to our system and face has to fit in our frame, it will be another difficulty while training the model.

2. Related Works

As indicated by Nithisha Raut Face discovery has been around for a long time. Stepping forward, human feeling showed by face and felt by mind, caught in either video, electric sign (EEG) or picture structure can be approximated. Human emotion recognition is the need of great importance so present day by insightful frameworks can copy and check responses from face. This can be useful to settle on choices be it with respect to recognizable proof of aim, advancement of offers or security related dangers. Perceiving feelings from pictures or video is an insignificant assignment for natural eye, however ends up being trying for machines and requires many picture preparing procedures for include extraction. Several works using computer vision [12-22] have been done in FER analysis. A few AI calculations are appropriate for this work. Any identification or acknowledgment by AI requires preparing calculation and afterward testing them on an appropriate dataset. This paper investigates a few AI calculations just as highlight extraction methods which would help us in precise recognizable proof of the human feeling [1]. In the work of Naveen Kumar H N, Jagadeesha S [2], subject independent FER in the case of from semi-occluded images has been discussed.. As indicated by, Jiabei Zeng, Shiguang Shan and Yong Li [3], a trainable neural network that uses patch-gated convolution (PG-CNN) has been used to study FER in occluded scenario. In the work done in [4], K-nearest neighbor (K-NN), facial expression recognition and Support Vector Machine (SVM) classifiers are applied and tested. According to Susmita Moitra and Souvik Choudry [5], several CNN based approaches to classify the basic human emotions in occluded scenarios has been analyzed.

3. Proposed System

In real time, occlusion is a serious issue and may have direct bearing on FER performance. The obstruction may be caused by beard, sun glasses, muffler, hairs, food and other objects that could be placed in front of the faces in daily life. These objects may block the eyebrows, lips, cheeks, and any other part of the face, thus making the emotion recognition difficult. Thus the occlusion problem statement is not yet addressed in the previous research papers, just emotion recognition on controlled environment is been discussed.

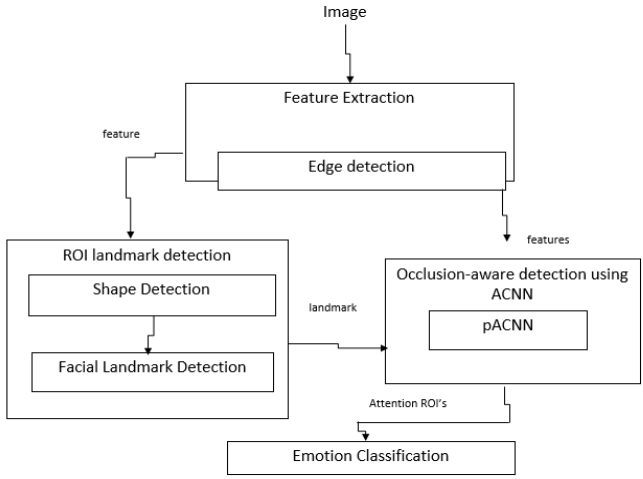

Figure 1. Proposed System Architecture diagram.

The proposed framework tends to the complexity of impediment, we introduce a Convolutional Neural Network mindfully instrument (ACNN) which can see the impediment districts of the face and work in the most discriminative un-blocked areas. It joins the numerous portrayals from facial areas of intrigue (ROIs). Thinking about various RoIs, we present fix based ACNN (pACNN). pACNN just focuses on neighbourhood facial patches. The proposed ACNNs are assessed on standard self-gathered outward appearance dataset with certifiable impediments, the two biggest in-the-nature outward appearance datasets (RAF-DB and AffectNet) and their changes with blended facial impediments. Outward appearance is recognized in specific facial areas, in light of the fact that the articulations are facial exercises summoned by group of muscle movements. Figure 1 shows the architecture of the proposed system. The image is given as input into the convolutional model and get some feature inlets. Initially, image size will be 224*224 and gets encoded into 512*28*28 when it enters into nine convolutional layer.ACNN breaks feature maps into several sub feature maps. The region decomposition is completed through aligning the facial image by fixing the 68 facial landmark round the face and dividing the facial landmark into 24 patches which covers all the informative region.The cropped patches enter pg-unit and divided into two branches. the primary branch represent the vector-shaped feature maps where it's un-weighted image and therefore the second branch represent the eye net where each patches are weighted. pACNN is meant to specialize in local discriminative and representative patches. In addition, partitioning the face into various local parts assists to find the placements of obstructions. To identify the standard parts of face that are combined with expression, We first find certain landmark of faces pointed by the proposed method ,Then supported the identified points, we chose or recompute few details that hides the essential parts of the face, such as nose, mouth, cheeks, eyes. Then we take the patches consistent with the angles of every subject's landmarks of face. ACNN sent the feature maps into gACNN. The global local-based ACNN has local information in the image and represent the global clues. The weight ranges from [0,1] where the 0 represent the occluded part and 1 represent the un-occluded part. This module indicates the benefaction of universal people face presentation [22]. Universal

representation is then calculated by the computed weight. Unit considers the feature maps ˜ pi as input, then learns localized facial features ψi:

ψi =ψ(˜ pi) (1)

and a corresponding weight αi:

αi =Ii(˜ pi)

ψi is a variable that represents the non-weighed features. αi is a constant that represent the patch i's importance. The fully connected layers integrates the weight of pg-unit and gg-unit. Based on their weight the emotion are classified. It is obvious that gACNN results in promising classification validity on surprised, scared sections respectively. The potential candidate expressions for confused expressions are surprise and fear, sad and neutral, fear and neutral are little concern.

4. Experimental Results

For the experimental result analysis, we have used pycharm tool with python script for image processing and machine learning techniques.

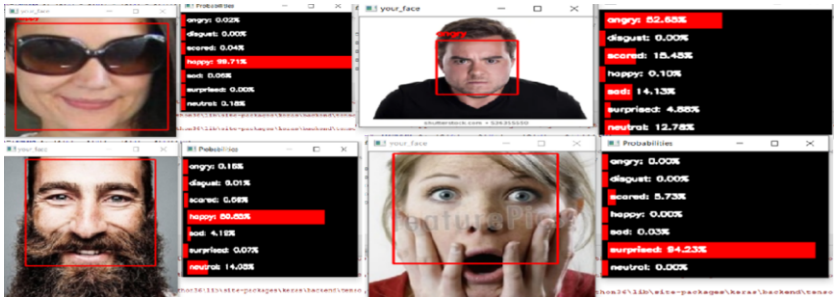

Figure 2. Sample output of the experiment.

Occlusion dataset images are collected from kaggle and given as the input file to the application. Finally based on the input images provided, our application recognizes the respective emotions. Figure 2 shows the sample outputs of the proposed solution. The results of the model looks promising. It shows the percentage of the key emotions from the occluded face. Also using this model, we can be able to view mixed emotions in occluded face of a specific person.

5. Conclusion and Future Works

As part of this work, a CNN based approach namely - ACNN for emotion recognition in occlusion scenarios has been presented. The Gate Unit present in the ACNN facilitates the model to move attention from the occluded parts to other unobstructed also as unique facial areas. Taking that countenance is distinguished in marked facial areas, a patch based pACNN was designed that comes up with region decomposition to identify facial parts that determine facial expressions. Also, an efficient gACNN to feed global facial details for FER in occlusion scenarios was developed. The proposed system analyses the potential of ACNNs to move the attention from obstructed areas to other similar areas. As part of future work, the way to produce monitoring parts in

faces without regions of interest, as ACNNs believe identical face detection and facial regions of interest placement modules can be explored.

References

[1] Y. Li, J. Zeng, S. Shan, and X. Chen, Patch-gated CNN for occlusionaware facial expression recognition, in Proc. Int. Conf. Pattern Recognit. (ICPR), Aug. 2018, pp. 2209–2214. [7] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, Robust face recognition via sparse representation, IEEE Trans. Pattern Anal. Mach. Intell., vol. 31, no. 2, pp. 210–227, Feb. 2009.

[2] S. Li, W. Deng, and J. Du, Reliable crowdsourcing and deep locality preserving learning for expression recognition in the wild, in Proc. CVPR, Jul. 2017, pp. 2584–2593.

[3] Naveen Kumar H N, JagadeeshaSSubject Independent Automatic Facial Expression Recognition from Partially Occluded Still Images by Employing Appearance Based Feature Extraction Schemes, 2019.

[4] Zhao.G, X. Huang, M. Taini, and S. Z. Li, Facial expression recognition from near-infrared videos, Image Vis. Comput., vol. 29, no. 9, pp. 607–619, 2011.

[5] Mollahosseini.A, B. Hasani, and M. H. Mahoor. (2017). AffectNet: A database for facial expression, valence, and arousal computing in the wild. [Online]. Available: https://arxiv.org/abs/1708.03985

[6] Vikram.K and S. Padmavathi, Facial parts detection using Viola Jones Algorithm, International Conference on Advanced Computing and Communication Systems (ICACCS-2015) on, 2017.

[7] P. Lucey, J. F.Cohn, T. Kanade, J. Saragih, Z. Ambadar, and I. Matthews, The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression, in Proc. CVPRW, Jun. 2010, pp. 94–101.

[8] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, Web-based database for facial expression analysis, in Proc. ICME, Jul. 2005, p. 5.

[9] L. Wiskott, J. M. Fellous, N. Kuiger and C. V. Malsburg, Face Recognition by Elastic Bunch Graph Matching, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 19, no. 7, pp. 775-779, 1997.

[10] M. S. Bartlett, J. R. Movellan and S. Sejnowski, Face Recognition by Independent Component Analysis, IEEE Transactions on Neural Networks, vol. 13, no. 6, pp. 1450-1464, 2002.

[11] Balasundaram, A., Chellappan, C. An intelligent video analytics model for abnormal event detection in online surveillance video. J Real-Time Image Proc 17, 915–930 (2020).

[12] A. Balasundaram and C. Chellappan, Vision Based Motion Tracking in Real Time Videos, 2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), Coimbatore, 2017, pp. 1-4, doi: 10.1109/ICCIC.2017.8524504.

[13] A. Balasundaram, C. Chellappan, Vision Based Gesture Recognition: A Comprehensive Study, The IIOAB Journal, Vol.8, Issue.4, pp.20-28, 2017.

[14] Balasundaram, A., Ashok Kumar, S., &Magesh Kumar, S. (2019). Optical flow based object movement tracking. International Journal of Engineering and Advanced Technology, 9(1), 3913–3916.

[15] Balasundaram, A., &Chellappan, C. (2019). Computer vision based system to detect abandoned objects. International Journal of Engineering and Advanced Technology, 9(1), 4000–4010.

[16] Balasundaram, A., Ashokkumar, S., Sasikumar, A. N., Kajendran, K., &Sathish Kumar, P. J. (2020). Computer vision and machine learning based facial expression analysis. International Journal of Advanced Trends in Computer Science and Engineering, 9(5), 7421–7426.

[17] S. Magesh Kumar; Balasundaram A; Sasikumar A N; Sathish Kumar P J. Human Emotion Prediction – A Detailed Study. European Journal of Molecular & Clinical Medicine, 7, 4, 2020, 1991-2001.

[18] A Balasundaram et al Computer vision based fatigue detection using facial parameters"2020 IOP Conf. Ser.: Mater. Sci. Eng. 981 022005

[19] D Kothandaraman, A Balasundaram, SeenaNaikKorra, E Sudarshan and B Vijaykumar Enhancing dull images using discrete wavelet families and fuzzy"2020 IOP Conf. Ser.: Mater. Sci. Eng. 981 022020

[20] Balasundaram, A. Computer Vision based Detection of Partially Occluded Faces. International Journal of Engineering and Advanced Technology, 9(3), 2188-2200.

[21] Balasundaram, A, Ashokkumar, S. Study of Facial Expression Recognition using Machine Learning Techniques. JCR. 2020; 7(8): 2429-2437.

[22] Ambeth Kumar.V.D et.al..Facial Recognition System For Suspect Identification Using A Surveillance Camera. Pattern Recognition And Image Analysis (Springer), Volume 28, Issue 3, Pp 410–420, 2018. (Doi: 10.1134/S1054661818030136)