Generating images of partial face using landmark based k-nearest neighbor

Israa Hadi, Alyaa Mahdi
College of Information Technology, Babylon University, Iraq

Article Info	ABSTRACT
	One of the most common approaches to address the partial face recognition challenge is to crop the full face image into segments. The problem is how the full face image must be cropped in a uniform way to generate informative segments. The un-blindly strategy was applied in this paper to generate informative segments, it depends on localizing the facial landmarks and selecting the more informative facial points as a key points, as more as the k-nearest neighbor concept was explored to select the k nearest landmark points to the key points. Two landmark localization techniques were experimented, the suitable technique resulted in segments which are overlapped due to the supervised clustering technique that explored in this paper to cover important biometric face regions, not repeated and covered most probabilities in which it is possible to distinguish the query face from the available part of it.

Corresponding Author:

Alyaa Mahdi,
College of Information Technology,
Babylon University, Iraq.
Email: alyaamahdi@itnet.uobabylon.edu.iq

1. INTRODUCTION

Partial face recognition, one of the face recognition challenges [1, 2], requires a specific strategy for the purpose of recognizing the face that is partially occluded. Occlusion is a real problem in face recognition and other tasks of facial analysis such as gender classification, age estimation, etc. Generally, there are two main methods to overcome occlusion which are a holistic-based method and a part-based method, each category has some ways to deal with occlusion problem [3].

The holistic-based approach does not crop the face image indeed, it handles the face as a single region used to predict the face identity. One of the holistic approach ways to handle the occluded face is to reconstruct it from the presented occluded one, but the main problem in this way is the variation in occlusion conditions must be handled by training the system on them [4]. Figure 1 shows examples of real-life occlusion.
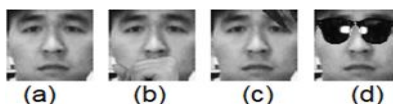

Figure 1. Face occluded with (a) no thing, (b) hand, (c) hair, (d) sunglasses [4]

Part-based method requires face cropping into specific segments which are overlapped or not from the face. There are some available methods that crop holistic face to generate face segments. One of the

cropping methods makes to crop the full face image into uniform parts [5], another one exploits facial landmarks to get face segments that are around the landmarks [6], another one makes to detect the occluded part [7], the detected occluded parts either reweighted or discarded. And also other ways available to overcome the face cropping of part-based partial face recognition challnge.

The majour problem faces who want to segments the human face in order to recognize it is the effectiveness of the segments and to some what they contain enformative facial region enable the system to recognize the face or partially occluded face from the generated segments. Accordingly, face image cropping into segments of facial region must be done accurately and carefully due to the importance of this preprocessing step as a building block that the next part-based face or partial face recognition steps depend.

This paper specifically aims to generate face segments which are served as a solution for the problem of generating informative segments that can be used in a partial face recognition system based part-based approach in order to recognize faces. There are a number of strategies explored in different research papers. Sun *et.al.* [8] proposed a randomly cropping strategy that relies on cropping the face image into 400 face patches from different positions and scales. Later, better 25 patches with the best features were selected by applying a specific algorithm. The cropping strategy proposed by *Hu et.al.* [9] relies on extracting five segments from the four face corners and center which are of specific squared equal sizes. Each segment has six different scales for it. The technique results in a total of thirteen segments with one size, different scales for the single face image. *Min et.al.* [10] made to address partial occlusion caused by facial accessories such as sunglasses, scarf, etc. They proposed a technique that makes to process only the non-occluded parts of the face. In this case the facial occlusion must be detected firstly, they detect it by using Gabor, PCA and support vector machine. The remaining non-occluded face parts are processed to be recognized, they use local binary pattern (LBP) for this purpose. In [11], *Zhang et.al.* offered a cropping strategy used it to recognize faces, this strategy depended on online cropping where the authors implemented a specific layer in their architecture called "crop-data" for this purpose. They generated six image segments are from the same size 136x136 pixels. All segments are from the same top-left positions with same width and height for all images. *Lopez-sanchez et.al.* [12] depended on Local Binary Pattern (LBP) to extract face features due to the nature of this method to separate occlusion from visible face regions features. In order to refine the effectiveness of extracted LBP features, they devided the full face image into arranged blocks as a grid and concatenated the histograms of each block finally in order to generate *"Local binary pattern histogram (LBPH)"* of the face image where the histogram of occluded regions differs from occluded-free regions.

In this paper the cropping strategy that followed makes to localize the facial landmarks on the detected face image as a first step, two types of landmark localization have experimented 15 and 68. Later a set of eight points from the most informative face regions based on trial and error are selected as key points from the overall localized points with optimal features [13]. The selected points are located at the same face positions in both cases with 15 and 68 points. Finally, the group of each selected key point from both of 15 and 68 detected landmark points with its k-nearest landmark points from the remaining detected landmark points is cropped to generate face segment, overlapping may occur in case of some points appeared in more than one group. The resulting segments are from different sizes, all of them are resized later. In this approach, we ensure that:
1. Get overlapped, not redundant facial segments.
2. All segments contain at least one informative key point to avoid being a useless segment.
3. Total segments are in same size and cover all of the informative face regions.

2. THE PROPOSED METHOD

Generaly, there are two groups of facial landmarks that detect the main points of human face both different in the number of poins, which are primary that detect the easily distinguished face point and secondary that obtained from primary which are surrounding the face components [14, 15]. There are many techniques available for the purpose of landmarks localizations, each one detects a specific number of points and each application chooses the reasonable points required for its purpose [16]. Some types of localizing landmarks shown bellow in Figure 2. Features may be extracted either by local or global method, local feature called hand crafted methods while the global feature using neural network to extract features [17, 18].

Assume the detected facial landmarks are the points that are to be classified or clustered in order to generate groups of landmarks points, these groups represents the facial segments. There exist numerous classifying and clustering techniques, k-nearest neighbor is the classification technique used in this paper not to decide a specific key-point belong to what class indeed, it used to generating groups of points as clustering dose.
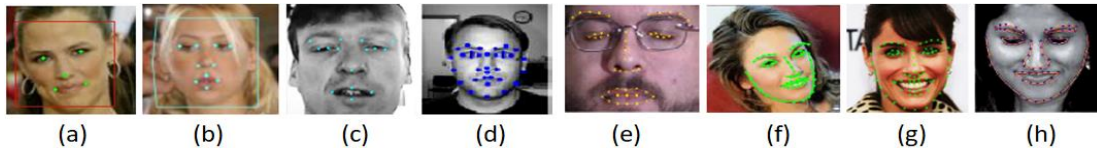
Figure 2. Some landmarks localization types with specific number of points detected.(a)5, (b)10, (c)15, (d)33, (e)46, (f)65, (g)68, (h)79 point [16]

The nearest neighbor (NN) technique [19] can be precisely identified as a supervised technique works to label the unknown query data by classifying them to the nearest neighbor class, Figure 3 (a) clarify NN concept. K Nearest Neighbor (KNN) [20, 21], developed from NN, calculate the distances between the query sample and each of the other related samples and finally take the K nearest one, Figure 7 (b) clarify this concept. Figure 3 shows the prediction of k-nearest neighbor. Star shape (*) represents the query to be classified, plus (+) and minus (-) represent two different classes. K is the number of nearest points required. Figure 3 (c) clarify the KNN classification technique [22, 23], in lowest square boundary the query sample (*) has 1 neighbor (+) so (*) classified into class1, in middle square boundary the query sample (*) have 7 neighbors, so (*) classified into class2 according to the largest number of its samples.
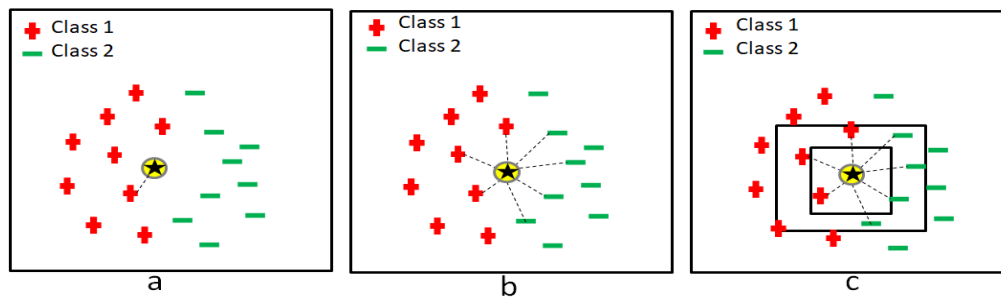

Figure 3. (a) Nearest Neighbor (1-NN), (b) K Nearest Neighbors (7-NN), (c) KNN classification

On the other hand, clustering strategy [24], such as k-mean and any other strategy, based on dividing the available data into N clusters without any data sharing, the clusters are isolated. Due to the importance of the overlapping between facial segments, and both classification and clustering strategies do not allow that, the supervised clustering technique that depends on distances using KNN concept was adopted in this paper, the adaptive knn algorithm of the proposed method shown in Algorithm1 as following:

Algorithm1: Adaptive KNN concept for overlapping group generation

for all the unknown samples UnSample(i)
   for all the known samples Sample(j)
     compute the distance between UnSamples(i) and Sample(j)
   end for
   find the k smallest distances
   group all k smallest distances in the new segment
   generate the segment (cropping from smallest-x to largest-x and from smallest-y to largest-y)
end fo

3. RESEARCH METHOD

The full face dataset used in a face recognition system must be preprocessed to be a database of segments used to be matched with the query image to recognize the identity. The block diagram in Figure 4 shows the proposed method to generate the required segments.
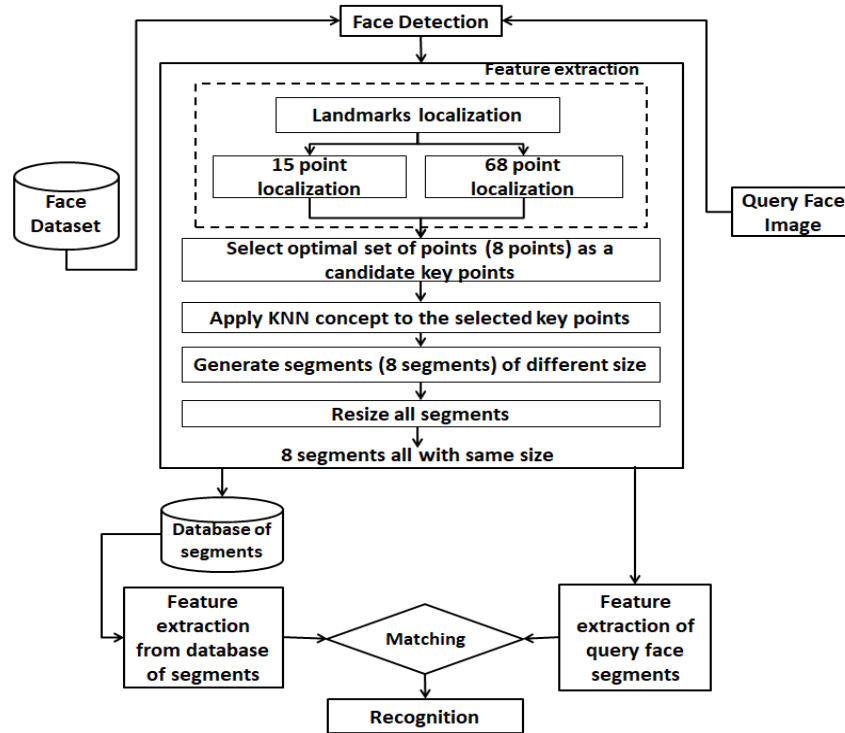
Figure 4. The proposed structure for segments generation approach

3.1. Face detection

Face detection can be defined as, finding the locations of the human face in the image, if present, regardless of its expression, partial occlusion, pose, illumination challenges. The location of the face window is detected and cropped to deal with it as the original image to be processed in the next steps presented in Figure 4. Figure 5 shows the original image and the detected face of the image.
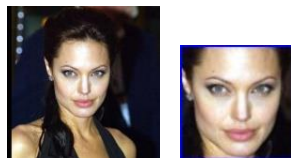

Figure 5. face detection for an example image from the LFW dataset using haar cascade

3.2. Feature extraction

Detecting facial key points is an important and very challenging problem due to a large variation, such as pose, size, position, viewing angle, and illumination conditions of facial features from one individual to another, and even for a single individual. It involves predicting the key-point's coordinates represented by (x, y) pair. There are many types of localization differ in the number and position of points to be localized, clarified in section 2. In this research, the cropping strategy was accomplished by experimented two localization strategies, one of them uses 15 points and another one using 68 points. The two landmark localization approach is summarized below.

3.2.1. The 15 point landmarks localization

The first landmark localization approach is done by using 15 points. The points are: Right eye center, left eye center, right eye inner corner, right eye outer corner, left eye inner corner, left eye outer corner, right eyebrow inner corner, right eyebrow outer corner, left eyebrow inner corner, left eyebrow outer corner, nose tip, mouth right corner, mouth left corner, mouth center upper lip, mouth center lower lip. The resulting image with the 15 facial landmark locations is shown in the Figure 6 (a).

3.2.2. The 68 point landmarks localization

The second proposed approach of landmark localization is by using 68 points. Different from the 15 points, the 68 points are surrounding the entire facial landmarks. The 68 points are distributed as following: Surrounding eyes with 12 points, 6 points each, eyebrows arc 10 points, 5 points each, nose 9 points, surrounding lips 20 points, face or cheekbone contour 17 points. The first step, face is aligned by identifying the 68 points to detect the facial landmarks. The 68 facial landmark points that applied to the face detected image are shown in Figure 6 (b).

3.3. Select the optimal set of facial points

In the previous step, two types of facial landmark localization were presented, 15 and 68 points. For both of mentioned landmark localization technique, key-points from the same locations on the face are selected. The only difference between the two techniques is the number and position of the remained points to be compared with the selected key-points according to the number of the neighbors and distances as will be addressed later in this paper. Eight key-points that cover all the informative regions of the face are selected. The Eight selected points are: two eyes corners, nose tip, tow mouth corners, and mouth center lower lip, all of them are projected to the both of previously mentioned landmarks localization techniques. The selected points can be noticed with a tiny square shape in Figure 6 (c).
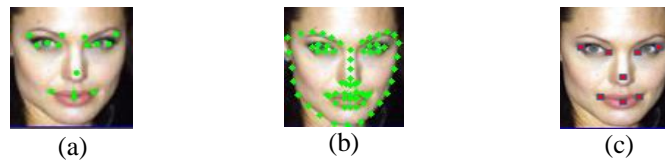

(a)   (b)   (c)

Figure 6. Sample from LFW dataset with two types of landmark localization (a) 15 point, (b) 68 point and (c) the selected key points

3.4. Apply the KNN concept

In this paper, the concept of nearest neighbor relies on distances that are mentioned in section 2 was explained, but not for supervised classification purpose indeed it is explored to the unsupervised image cropping purpose by obtaining the nearest points to the specific selected points and crop the image to get a segment with k nearest points to the specific facial key point previously selected. The aim of the cropping strategy proposed in this paper is to get 8 groups of points, segments, each with key point and K nearest points to it. The KNN classifier requires pre-specified groups to line data inside them, but in this case, we don't have pre-specified groups, new groups of points are already required that led us to the concept of clustering. On the other hand, clustering will not be allowed because the same data cannot be shared between more than single groups, and the required segments must be overlapped between them. So, a new strategy is explored that relies on selecting the nearest K landmark points to the selected key points and group them together as a new segment. If any selected points have the smallest distances with points that are previously grouped, they are re-grouped too in this new group and resulting in an overlapped segments where the overlapping is the objective of this approach. According to the 15 points landmark localization, K is determined 7 which prepared the best experimented K that result in informative segments, and in the 68 points landmark localization, K is determined 20 which prepared the best experimented K that result in informative segments. The presented adaptive KNN algorithm is shown in section2, Algorithm1 above.

3.5. Segments generation

The segments will be generated by grouping each key point with its k-nearest points together. With 15 facial landmark localization technique, the selected optimal 8 key points are projected on 15 points, the obtained result shown in the Figure 7. The tiny square shapes represent the selected point and the other points represent the nearest k neighbors to the specified square. As noted in the Figure 7, there are some segments which are exactly the same. This result is due to the small number of points compared among them that give the same smallest distances and result of the same nearest points that cause a redundant segment, the nose tip presented in the 5th segment from Figure 7, for example, have the nearest distance with left and right mouth corners, the same distances resulted in the left mouth corner, it also has the smallest distances with right mouth corner and nose tip points, so both of them resulting exactly the same segment. The resulting segments are redundant, furthermore, there is no, or small portion, of overlapping between segments. The second landmark localization technique uses 68 points. The same technique will be applied except a number of the

nearest neighbor, 20 in this case, is different due to a large number of neighbors compared to 15 landmarks and the small neighbors will result in a very small region of the face that cannot be recognized. Also, the same approach applied by projecting the same selected 8 points presented in Figure 6(c) on the 68 points. This approach compares each of the selected 8 points with the remaining 67 facial landmark points to select the nearest 20 ones. The result of this process gives overlapping segments without redundant segments as shown in Figure 8. More ever, the face contour, the area surrounding the face, is a good addition in this set of segments.

3.6. Resize all segments

As previously noticed, the resulting segments are of different sizes. The last segment in Figure 8, for example, the selected points located at the mouth that is indeed surrounded by 20 points, so its neighbors have smaller distances between them. On the other hand, the first segment in the same figure has large distances between the selected point and the nearest neighbor points to it. The new segments size depends on the previously largest segments. The segments resize will be done by checking the width (number of columns) and height (number of rows) of all segments, select the largest width and height among them, and finally adding the difference to all of the other segments to be of the same size. The rows and columns addition process involves taking the differences of width and height between the largest and current segment, dividing by two and finally adding the half of difference to the left and right of the current segment, in case of width increase, and adding the half of the difference to the upper and lower of the current segment, in case of height increase. Figure 9 shows the generated segment with 15 facial landmark point, showed in Figure 7, after resizing them. Figure 10 shows the generated segments with 68 facial landmark point, showed in Figure 8, after resizing them.

4. RESULTS AND ANALYSIS

This paper proposed an image cropping approach to give efficient features for partial face recognition part-based approach that makes to crop the full face image. The proposed cropping strategy passes throw numerous steps:

4.1. Localizing the facial landmarks

Two types of landmarks localization are experimented and compared, 15 and 68 localization technique. Figure 10 (a), (b) shown the 15 and 68 facial landmarks points respectively.

4.2. Selecting the optimal set of facial landmarks points as key points

Projecting the selected key points on both of the pre-localized landmarks points (15 and 68). Figure 6 (c) shown the selected points in a tiny square shape.

4.3. Apply the KNN concept and generate facial segments

Apply the KNN concept between each of the selected key points and the remaining points [25] to provide a new group of points, segment, as clarified in Algorithm1 and Figure 3 (c) previously in section 2. The segments result from applying Algorithm1 on 15 landmark points with 7 neighbors and 68 landmark points with 20 neighbors are shown in Figure 7 and Figure 8 respectively.

Figure 7. The segments with key points and 7 nearest neighbors to it from the 15 points
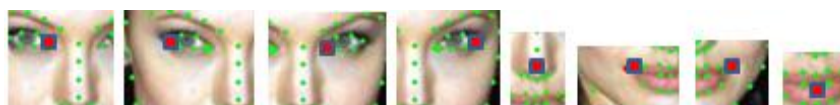

Figure 8. The segments with key points and 20 nearest neighbors to it from the 68 points

4.4. Resize all the augmented facial segments

All segments are resized to the largest image width and largest image height, the results after resizing segments presented in Figure 7 and Figure 8 are shown in Figure 9 and Figure 10 respectively.

Figure 9. The segments with key points and 7 nearest neighbors from the 15 points after resizing

Figure 10. The segments with key points and 20 nearest neighbors from the 68 points after resizing

The generated segments used as a training data that the system trained on them to be able to recognize a part of face when unseen (query) partially occluded face presented to the system, where the same cropping strategy is followed to generate a facial segments and discarding the occluded segment while keeping the only informative segments and extract their features to recognize the face according to them. Finally, a classifier such as neural network used to recognize only the occluded-free segments based on them features.

In order to evaluating with the other technique to generate facial segments for partially occluded face recognition, let us go back to the mentioned techniques presented previously in the introduction section. Opposite to other face cropping approaches such as approach presented by Sun *et.al.* [8] above in this paper that they offered a randomly cropping strategy where large number of facial segments were generated randomly, the cropping strategy presented in this paper is uniformly and not randomly. *Hu et.al.* [9] presented technique to generate thirteen facial segment for each face each of the same size and different scales (six scale for each of the five segments), the generated segments are large in number and redundant, opposite to the proposed technique where the generated segments are less redundant, less number of segments and at the same time they are informative segments surly have informative facial region. Another face cropping approach mentioned above which is presented by *Zhang et.al.* [11] where the generated segments are all of the same size and from the same top-left corners which is a constant for all dataset image, while the proposed cropping strategy presented in this paper didn't fix any thing where the top-left corner for any segment is different from image to another according to the nearest informative point position in that image and also the size of segments is also different from one image to another according to the size of the face inside the image, while the segments sizes of the same face finally resizd.

Krishnaveni and Sridhar [26] introduced a technique to recognize a partially occluded face using a dynamic approach. The authors assumed dividing the dataset and the query face images into k non-overlapping mxm patches to be combined later to generate a patch sequence, clarified in Figure 11. Euclidian distance used to measures the distances between the query image and dataset images patches and finally, dynamic time warping (DTW) retrieves the optimal alignment with less cost. As it is obvious, the patches don't have the guarantee of informative facial region as the proposed cropping strategy in this paper where the generating segments to be matched have a guaranty of informative facial regions consisting within.
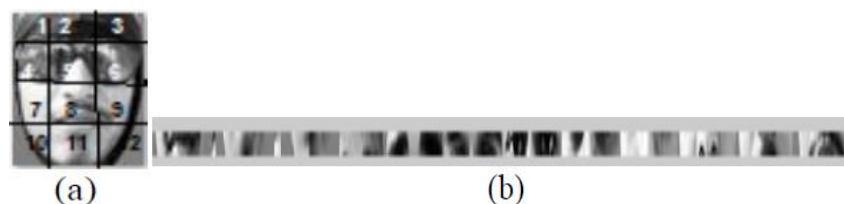

Figure 11. (a) Dividing face image into k (mxm) patches (b) Sequence of patches [26]

5.   CONCLUSION

As it is obvious from the presented two facial landmark points localization types with the same adapted KNN algorithm, the resulted segments from 68 landmark localization, Figure 10, have a more informative appearance, more segment's overlap, more partial face probabilities, less redundancy than the resulted segments from 15 landmark localization showed in Figure 9. Accordingly, the resulting segments from 68 facial landmark points with adapted KNN algorithm are more informative for the recognition system, convolution neural network in this paper, to recognize the face from other segments if one of them was occluded, because the system was trained on more partial face probabilities due to less redundance. As more as, the system has more probabilities in this case with more overlap, if some of them are occluded, the probability of more informative segments existing is heigher.

Compared to other informative facial segments augmentation technique, the proposed technique is easy and accurate at the same time, the technique is doesn't work randomly and at the same time it does not require a great effort to ensure that the segment contains informative facial regions, the adopted KNN algorithm does it automatically because the nature of its work that select the K nearest points to the specific key point. Since all of the prelocalized points are informative because they are facial points resulted from some facial landmarks localization technique, all segments ensure containing informative facial region.

REFERENCES

[1]   Sonia Ohlyan, Sunita Sangwan, Tarun Ahuja. "A Survey On Various Problems & Challenges In Face Recognition." *IJERT*. June – 2013; 2(6): 2533-2538.
[2]   G. Ramkumar, E. Logashanmugam. Multimodal Verge for Scale and Pose Variant Real Time Face Tracking and Recognition. *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*. 2019; 13(2): 665-670.
[3]   Yong Li, Jiabei Zeng, Shiguang Shan and Xilin Chen. *Patch-Gated CNN for Occlusion-aware Facial Expression Recognition*. 24th International Conference on Pattern Recognition (ICPR). Beijing, China. 2018: 2209-2214.
[4]   X. Mao,Y. Xue, Z. Li,K. Huang,and S. Lv. *Robust Facial Expression Recognition Based on RPCA and Adaboost*. 10th Workshop on Image Analysis for Multimedia Interactive Services.London. 2009:113-116.
[5]   Lin Zhongy, Qingshan Liuz, Peng Yangy, Bo Liuy, Junzhou Huangx, Dimitris N. Metaxasy. *Learning Active Facial Patches for Expression Analysis*. Department of Computer Science, Rutgers University, Piscataway, NJ, 08854. Nanjing University of Information Science and Technology, Nanjing, 210044, China. Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX, 76019. 2012.
[6]   Wei Li, Farnaz Abtahi, Zhigang Zhu. "*Action Unit Detection with Region Adaptation, Multi-labeling Learning and Optimal Temporal Fusing*". IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA. 2017: 6766-6775.
[7]   Xiaohua Huang, Guoying Zhao, Wenming Zheng, Matti Pietikainen. "Towards a Dynamic Expression Recognition System under Facial Occlusion". *Pattern Recognition Letters.*2012; 33(16): 2181-2191.
[8]   Yi Sun, Yuheng Chen, XiaogangWang, Xiaoou. "Deep Learning Face Representation by Joint Identification-Verification." *Advances in neural information processing systems.* 2014: 1988-1996.
[9]   Guosheng Hu, Yongxin Yang, Dong Yi, Josef Kittler, William Christmas, Stan Z. Li, Timothy Hospedales. "*When Face Recognition Meets with Deep Learning: an Evaluation of Convolutional Neural Networks for Face Recognition*" IEEE International Conference on Computer Vision Workshop (ICCVW). Santiago, Chile. 2015: 384-392.
[10]   RuiMin, Abdenour Hadid, Jean-Luc Dugelay. "*Improving the Recognition of Faces Occluded by Facial Accessories*". Face and Gesture 2011. Santa Barbara, CA, USA. 2011:442-447.
[11]   Yanhong Zhang, Kun Shang, Jun Wang, Nan Li, Monica M.Y. Zhang. "Patch Strategy for Deep Face Recognition". *IET Image Processing.* 2018; 12(5): 819 – 825.
[12]   Daniel L´opez-S´anchez1, Juan M. Corchado1;2, Ang´elica Gonz´alez Arrieta. "*Inhibition of Occluded Facial Regions for Distance-Based Face Recognition*". Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI-18). Stockholm, Sweden. 2018; 5324-5328.
[13]   Ruaa Mohammed Hamza, Dr. Tawfiq A. Al-Assadi. Genetic Algorithm to Find Optimal GLCM Features. Department of Computer Science College of Information Technology.2012.
[14]   Yue Wu, Qiang Ji. "Facial Landmark Detection: a Literature Survey". *International Journal on Computer Vision*. 2018:1-28.
[15]   Oya Celiktutan, Sezer Ulukaya and Bulent Sankur. "A Comparative Study of Face Landmarking Techniques". *EURASIP Journal on Image and Video Processing*. 2013; 2013(13),1-27.
[16]   Dhananjay Rathod, Vinay A, Shylaja SS and S Natarajan."Facial Landmark Localization – A Literature Survey". *International Journal of Current Engineering and Technology*. 2014; 4(3), 1901-1907.
[17]   Yongjin Lee, Kyunghee Lee, Sungbum Pan. *Local and Global Feature Extraction for Face Recognition*. International Conf. on Audio- and Video-Based Biometric Person Authentication. Rye Brook, NY, USA. 2005; 3546: 219-228.
[18]   Nik Noor Akmal Abdul Hamid, Rabiatul Adawiya Razali, Zaidah Ibrahim. Comparing Bags of Features, Conventional Convolutional Neural Network and Alexnet for Fruit Recognition. *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*. 2019; 14(1): 333-339.

[19] Nitin Bhatia, Vandana. Survey of Nearest Neighbor Techniques. *(IJCSIS)*. 2010; 8(2), 302-305.
[20] Gongde Guo, Hui Wang, David Bell, Yaxin Bi, and Kieran Greer. *KNN Model-Based Approach in Classification*. On the Move to Meaningful Internet Systems.2003; 2888: 986-996.
[21] A. Ibrahim, M.K.Osman, N.A.M. Yusof, K. A. Ahmad, N.H. Harun, R.A.A Raof. Characterization of Cracking in Pavement Distress Using Image Processing Techniques and K-Nearest Neighbor. *Indonesian Journal of Electrical Engineering and Computer Science (IJEECS)*. 2018; 14(2), 810-818.
[22] Aman Kataria, M. D. Singh. A Review of Data Classification Using K-Nearest Neighbour Algorithm. *International Journal of Emerging Technology and Advanced Engineering*. June 2013; 3(6), 354-360.
[23] Sadegh Bafandeh Imandoust and Mohammad Bolandraftar. Application of K-Nearest Neighbor (KNN) Approach for Predicting Economic Events: Theoretical Background. *Journal of Engineering Research and Applications*. Sep-Oct 2013; 3(5), 605-610.
[24] Prof. Dr. Tawfiq A. Abbass, Dr. Mahdi Jasim. *New Geometrical Similarity-based Clustering Algorithm for GIS Vector Data*. Proceedings of International Conf. on Engineering and Information Technology. Toronto, Canada. 2012; 1-4.
[25] Aher Pravin D., Galande S.G. Ecludian Distance based Partial Face Recognition. *International Journal of Engineering Research & Technology (IJERT)*.2017; 6(7): 291-295.
[26] B Krishnaveni, Dr. S Sridhar. Partially Occluded Face Recognition Using Dynamic Approach. *(IJSETR)*. 2016; 5(12):3341-3347.