IJDCST @ Aug-Sep-2024, Issue- V-8, I-2, SW-07
ISSN-2320-7884 (Online)
ISSN-2321-0257 (Print)

An Advanced Face Detection System Using Ensemble Learning Algorithms

Gosala Priyanka, Dr. S.Jhansi Rani,
Student, Computer Science and Systems Engineering, Andhra University, Visakhapatnam
Associate Professor, Computer Science and Systems Engineering, Andhra University, Visakhapatnam

Abstract: Face detection in extreme conditions, such as varying illumination, pose, occlusion, and background clutter, remains a significant challenge in computer vision. This paper presents a hybrid approach for robust face detection using the Local Binary Pattern (LBP) and Fuzzy C-Means (FCM) clustering algorithm. The Local Binary Pattern technique is employed to extract texture features from facial images, as it is robust to changes in illumination and can capture micro-patterns effectively. LBP converts an image into a binary matrix, which simplifies the subsequent feature extraction process. Fuzzy C-Means clustering is used to classify and segment the face region from non-face areas by grouping similar data points based on their proximity. FCM offers flexibility by assigning partial membership values to each data point, which is crucial in handling the uncertainties and variabilities introduced by extreme conditions such as occlusion or poor lighting. The proposed method was evaluated on several benchmark face detection datasets under extreme conditions. Experimental results demonstrate that the combination of LBP for feature extraction and FCM for classification provides superior accuracy and robustness compared to traditional face detection methods. The algorithm shows resilience to lighting variations, partial occlusions, and complex backgrounds, making it suitable for real-world applications, including surveillance and biometric systems.

Keywords: Face detection, Local Binary Pattern, Fuzzy C-Means, extreme conditions, texture feature extraction, clustering, illumination invariance.

1. Introduction

Computers, new sensing and analysis, and rendering technology, among other vital technologies, are becoming increasingly intelligent as processing power grows exponentially. Several academic projects and commercial products have demonstrated computers' inherent ability to engage with humans. This has been accomplished via observing and photographing people, as well as using microphones to listen to them, acknowledge inputs, and react in a nice manner with others. HCI, which stands for Human-Computer Interaction, is one of the essential approaches that enables such things to happen. Face detection is the first step in all facial analysis algorithms, which includes face alignment, face modelling, face relighting, facial recognition, face verification, head pose tracking, age and gender recognition, and so on. In this way, we might say that facial readings are simpler on computers. Psychological behaviour in terms of the face is a prospective research topic. Advances in the fields of face detection [1], recognition, and tracking are extremely beneficial for behaviour analysis. As a result, computers today have the ability to grasp human intentions and cognitive processes. The primary goal of the facial detection system is to determine the presence or absence of any faces in the image. This looks to be a straightforward operation to humans; nevertheless, for computers, it is extremely difficult, and it has been one of the most extensively researched problems in recent decades. Facial detection is a challenging task. This difficulty can be ascribed to variations in scale, location, and orientation, in-plane and out-of-plane rotation, human facial expressions, lighting circumstances, and numerous occlusions, among other factors. There are numerous ways in the literature for detecting faces. Face detection systems and this field have advanced significantly over the last decade. Many methods for face detection have been developed, and numerous literature articles have examined the benefits and drawbacks of previous techniques in face detection.

The cascade face detector introduced by Viola and Jones [5] employs haar-like features with AdaBoost to train cascaded classifiers, resulting in correct performance with real-time performance. However, some studies [6] imply that this detector may decline significantly in real-world applications with massive visual representations of human faces, despite additional improved features and classifiers. Recently, convolutional neural networks (CNN) made amazing success on a variety of computer vision tasks, including photo classification [7] and face popularity [8]. Li et al. [9] use cascaded CNNs for face detection; nevertheless, bounding box calibration from face detection incurs a higher computational cost and misses the natural link between facial landmark localisation and bounding field regression. Yang et al. [10] develop deep convolutional neural networks for facial characteristic reputation in order to generate excessive responses in face areas, resulting in candidate face windows. However, because of its intricate CNN form, this technique is time-consuming to implement. However, most accessible face identification methods ignore the underlying relationship between these duties. Even though there are multiple efforts that attempt to address them simultaneously, there are still barriers. However, the handcrafted elements limit its overall performance.

2. Literature Survey

Lal et al. [12] described the MSFP-ELM as an emerging technique that reduces some of the challenges presented by extra procedures, which has recently piqued the interest of a rising number of researchers. The prior is represented by single-hidden layer feedforward neural networks (SLFNs). It selects hidden channels and analytically determines their outcomes. So, in this article, we intend to provide you with a full overview of the MSFP-ELM developments for various tools. Finally, MSFP-ELM's weaknesses and strengths will be demonstrated, along with its potential viewpoints.

Zhou et al. [13] analyse the current state-of-the-art face detectors and their performance on the benchmark dataset FDDB, as well as comparing algorithm creation processes. Second, we look into their performance degradation when tested on low-quality photos with varying amounts of blur, noise, and contrast. Our findings show that both handcrafted and deep-learning-based face detectors are insufficiently robust for low-quality photos. It encourages academics to develop more robust designs for face detection in the field.

Louis et al. [14] look into the drawbacks of using a large number of features, a massive training dataset, and a lengthy training session to achieve a high-performance frontal face detector. The proposed face detector is based on a novel concept that suggests using a joint decision from two different feature-trained detectors, one trained with LBP features and the other with Haar features. Both detectors are trained using a small face/non-face dataset and in a reasonably short amount of time. Thus, both detectors agree on the face picture but rarely on the non-face image. Using a multi-detections merging process with a simple clustering method improves the results greatly. The detector's robustness is tested twice: once with a face/non-face dataset and compared to the Lienhart frontal face detector, and then again using a real-life sequence.

Karras et al. [15] suggested a robust feature extraction strategy based on a specific way of eigenanalysis of the unique classes identified in the challenge, using neural network-based classifiers. Such an eigenaiysis seeks to find the primary properties of the above-mentioned uniquely defined classes. During the testing phase, each unknown image is analysed using a sliding window raster scanning process to determine whether sliding windows correspond to one of the previously described distinct classes using a first stage neural classifier. After such a sliding window labelling technique, it is fair to apply a second stage neural classifier to the testing picture, which is considered as a succession of such labelled sliding windows, in order to make a final decision regarding whether or not a face exists within the test image. Although the proposed approach is a two-step procedure, it is clear that the first-stage classification process is the most important, because obtaining strong identification/labeling accuracy would greatly assist the final classification stage. As a result, the experimental component of this work focusses on analysing the labelling accuracy of face-specific classes at the initial classification stage.

Tsai et al. [16] demonstrated a CNN trained to enhance the accuracy of face identification while also capturing facial traits. The proposed solution addresses the condition in which the face is associated with occlusion. The face detection network calculates all of the face regions and facial landmarks from the inputted image. The face is then aligned using facial landmarks and fed into a face recognition network for identification. The experimental accuracy results were 96.15% and 88.46% with an occlusion ratio of 25% and 50%, respectively. The suggested technique significantly enhances facial recognition accuracy while the face is obscured.

Jha et al. [17]. Review of Face Recognition Technology: In recent decades, face recognition has become a major field in computer-based application development. This is due to its wide range of applications. Face recognition using database photographs, real data, recorded images, and sensor images is also difficult due to the large number of faces. Face identification relies heavily on image processing, pattern recognition, and computer vision. This report also discusses potential research problems and future work. In addition, this publication includes a prospective examination of facial recognition.

Arachchilage et al. [18] discuss some of the critical issues of face recognition in unfavourable settings. In this context, we present an end-to-end system for real-time video-based facial recognition. This system finds, monitors, and recognises people using a live video feed. The proposed method solves three major issues for video-based face recognition systems: end-to-end computational complexity, in-wild recognition, and multi-person recognition.

Wang et al. [19] investigated how to train face detectors without low-light annotations. Using current normal light data, we propose adapting face detectors from normal light to low light. This assignment is tough because the contrast between brightness and darkness is too wide and complex at the object and pixel levels. As a result, the present low-light enhancement or adaption solutions do not function well. To address this issue, we offer a Joint High-Low Adaptation (HLA) approach. We develop both bidirectional low-level adaptation and multitask high-level adaptation. For low-level, we boost dark photos while degrading normal-light ones, causing both domains to move closer together. Furthermore, our adaption technique can be applied to a variety of applications, including improved supervised learning and generic object detection. The project is publicly available at: https://daooshee.github.io/HLA-Face-v2-Website/.

3. Proposed Methodology

The methodology presented the hybrid approach for robust face detection using the Local Binary Pattern (LBP) and Fuzzy C-Means (FCM) clustering algorithm. The Local Binary Pattern technique is employed to extract texture features from facial images, as it is robust to changes in illumination and can capture micro-patterns effectively. LBP converts an image into a binary matrix, which simplifies the subsequent feature extraction process. Fuzzy C-Means clustering is used to classify and segment the face region from non-face areas by grouping similar data points based on their proximity. FCM offers flexibility by assigning partial membership values to each data point, which is crucial in handling the uncertainties and variability's introduced by extreme conditions such as occlusion or poor lighting.

3. 1 Local Binary Pattern (LBP)

Local Binary Pattern (LBP) is a simple and efficient texture operator initially developed for texture classification but has been widely adopted in face recognition systems due to its ability to represent both micro and macro patterns in images. The LBP operator labels each pixel of an image by thresholding the neighborhood of each pixel and considering the result as a binary number. The LBP operator captures local spatial patterns and intensity variations in the image, which are critical for face recognition tasks.

A. Key Concepts of LBP:

Texture Descriptor: LBP is a powerful texture descriptor that effectively captures the local features of an image. It works by comparing each pixel with its neighbors and generating a binary code. This binary code reflects the texture pattern at that pixel.

Binary Code Generation: The process involves taking a pixel and comparing it with its surrounding pixels (usually 8 neighbors). If the neighboring pixel's intensity is greater than or equal to the center pixel's intensity, it is encoded as 1; otherwise, it is encoded as 0. This forms a binary number, which is converted to a decimal value to form the LBP feature for that pixel.

Uniform Patterns: LBP also focuses on uniform patterns (binary sequences with at most two transitions between 0s and 1s). Uniform patterns are particularly useful because they represent most of the texture structures in images like edges, corners, and spots, making them highly relevant for face representation.

LBP Histograms: The LBP operator is applied to an entire image or divided into regions. A histogram of the LBP values is computed for each region, representing the distribution of local features. The combination of histograms from different regions forms the feature vector, which is used for face recognition.

Rotation Invariance: LBP can be extended to be rotation invariant by modifying the binary encoding process to account for changes in the rotation of the image.

B. Advantages for Face Recognition:

Simplicity and Efficiency: LBP is computationally efficient, making it suitable for real-time face recognition applications.

Illumination Invariance: LBP is relatively robust to changes in lighting conditions, as it relies on local intensity differences rather than absolute intensity values.

Robustness to Facial Expressions: LBP can handle variations in facial expressions and partial occlusions due to its focus on local texture patterns.

3.2 LBP-Based Face Recognition System:

In an LBP-based face recognition system, the steps typically involve:

Preprocessing: Normalize the input face images to ensure uniform size, lighting, and pose.

Feature Extraction: Apply the LBP operator to the face image, either globally or on smaller regions (e.g., dividing the face into blocks), and compute the LBP histograms for each region.

Classification: Use machine learning or template-matching algorithms to classify the face based on the extracted LBP features. Methods like Support Vector Machines (SVMs) or Nearest Neighbor classifiers are commonly used.

Matching and Recognition: Compare the LBP histograms of the query face image with those of the stored images in the database to find the closest match.

3.3 Fuzzy C-Means (FCM) Clustering Algorithm Face Recognition System

Face recognition is a critical technology used in various applications, including security, identity verification, and human-computer interaction. The need for reliable, robust, and adaptable algorithms in face recognition has led to the exploration of multiple techniques. One such algorithm is the Fuzzy C-Means (FCM) clustering algorithm, which has been widely studied for its efficiency in handling uncertainty and providing soft clustering solutions.

A. Overview of FCM Clustering Algorithm:

Fuzzy C-Means (FCM) is a type of clustering algorithm that groups data into clusters by minimizing an objective function. Unlike the traditional k-means algorithm, where each data point belongs strictly to one cluster, FCM allows each data point to belong to multiple clusters with a degree of membership, representing the "fuzziness" of the clustering. This is particularly useful in real-world problems like face recognition, where data may not clearly belong to one class or group.

B. Key Steps in FCM Algorithm:

Initialization: Randomly initialize the cluster centers or centroids.

Assign Membership: Calculate the degree of membership of each data point to every cluster based on the inverse distance between the point and the cluster center.

Update Cluster Centers: Calculate new cluster centers by minimizing the objective function, which depends on the distance between the points and the centers, weighted by their membership.

Iterative Optimization: Repeat the process until convergence (when the change in membership or centroids falls below a threshold).

4. Results and Discussions

In this project as per your instructions we have applied Viola Jones with Haar-cascade to extract faces from images and then used this faces features to train with ML algorithm to calculate face recognition accuracy.

In next part we have applied LBP algorithm on same faces to extract pattern histogram features and then applied Fuzzy-CMEANS algorithm whose cluster label will be assigned to LBP features which can help ML algorithm in accurately recognizing faces. Fuzzy-CMEANS will group similar faces into same cluster and this cluster label features will be input to LBP features to allow ML algorithm to accurately recognized faces.

Viola Jones features and LBP-Fuzzy-CMEANS features will be split into train and test where algorithm will use 80% features for training and 20% for testing. 80% training features will be input to Random Forest ML algorithm to train a model and this model will be applied on 20% test features to calculate Jones and LBP face recognition accuracy.

To train and test above algorithm performance we have used VIOLA JONES face cropped dataset which captures with different facial expression and in different lighting conditions. In below screen showing dataset images captured in different light

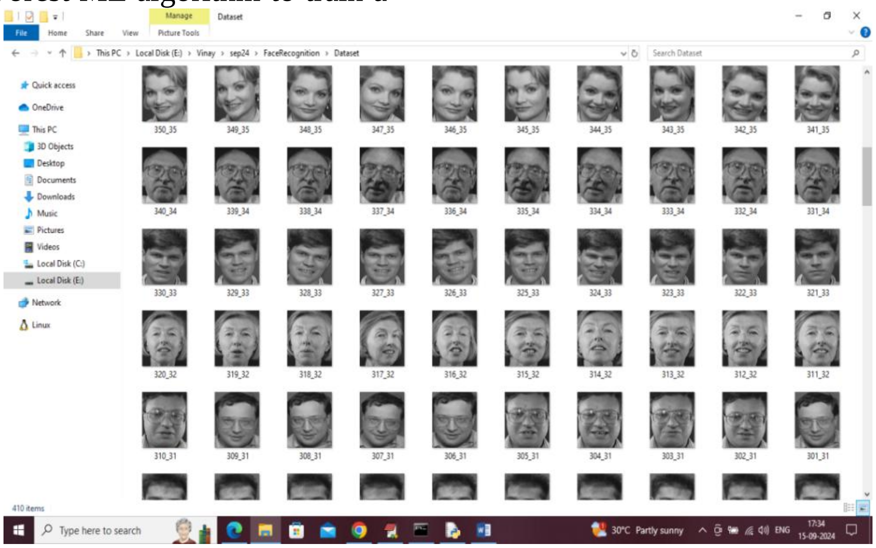

Figure 1 Dataset Images

ISSN-2320-7884 (Online)
ISSN-2321-0257 (Print)

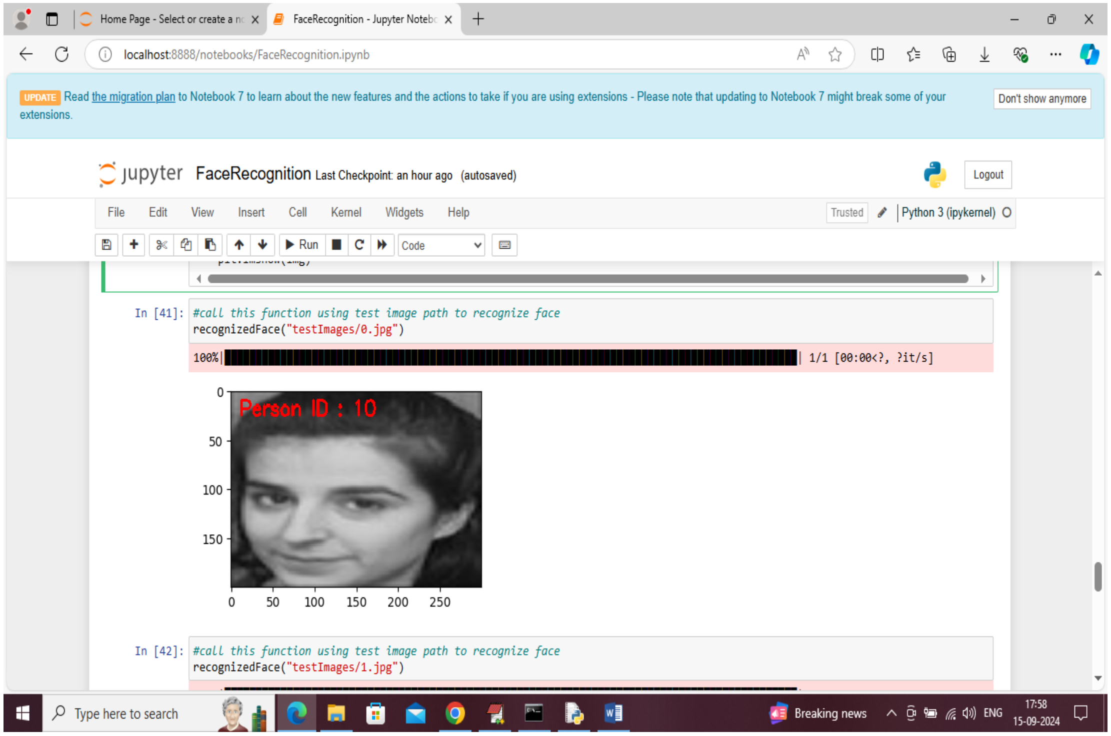

Figure 2 showing the face detection

In above screen calling ‘recognize’ function with image path and then algorithm recognize person id as 10

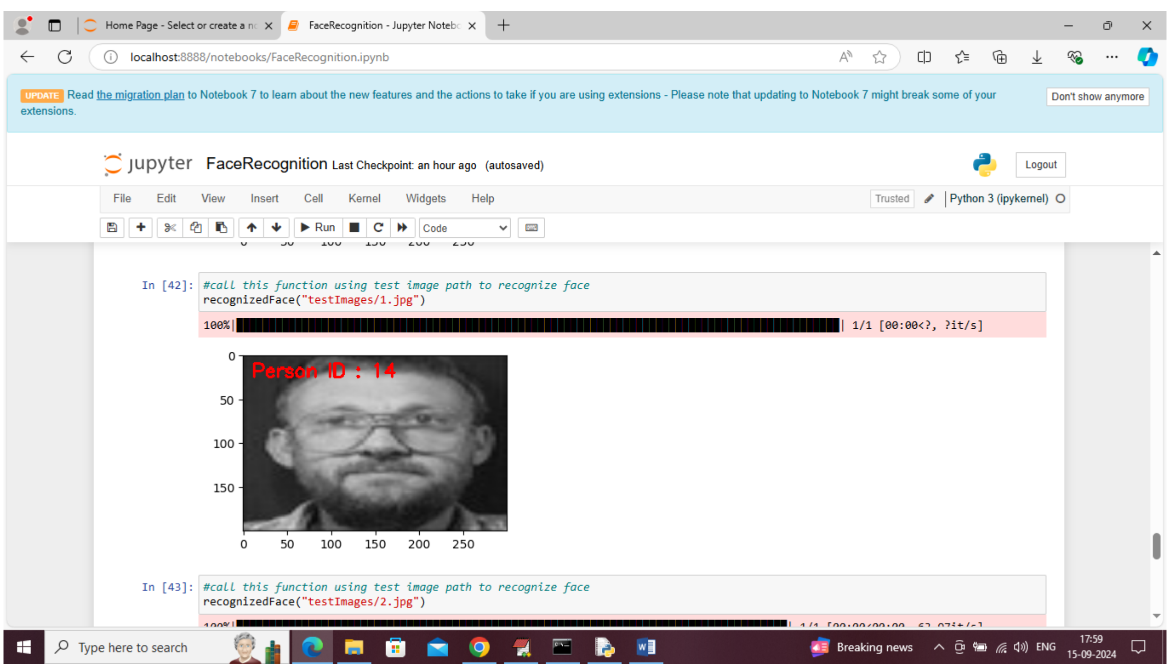

Figure 3 Recognizing the Face

In above screen another face recognize as person 14

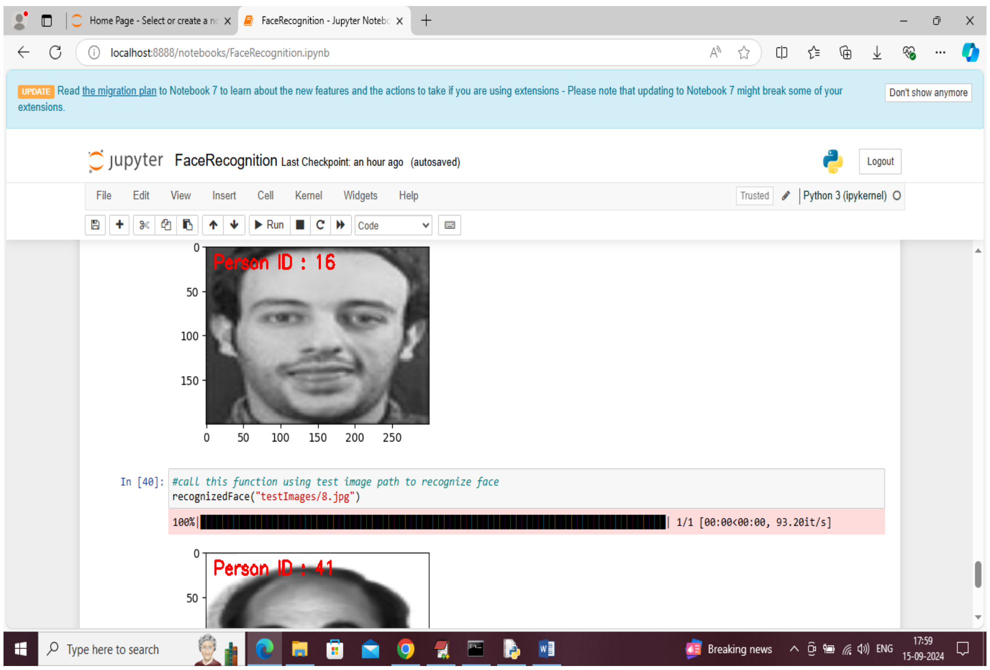

Figure 4 Final Output

In above screen another person recognized as person ID 16.

Similarly by using this function you can input test image path to recognize faces and all test images you can find inside 'testImages' folder

5. Conclusion

In the context of face detection in extreme conditions, the combination of Local Binary Pattern (LBP) and Fuzzy C-Means (FCM) algorithms has proven to be effective in addressing challenges such as illumination variations, occlusions, and pose variations. LBP is highly robust against illumination changes and works well in scenarios with varying light conditions. It captures texture information efficiently, providing essential features for face detection. LBP's computational simplicity ensures fast processing, making it suitable for real-time applications. However, its performance decreases when dealing with extreme occlusions or larger pose variations. FCM clustering enhances the segmentation of facial regions by grouping pixels based on similarity and dealing with uncertainty in data. FCM is adaptive to noise and variations in facial features. When used with LBP, FCM helps improve the precision of feature extraction and classification by better handling ambiguous areas, such as shadowed regions or occluded parts of the face.

Despite the effectiveness of LBP and FCM, extreme conditions like large-scale occlusions, severe angle shifts, and extreme low lighting still pose significant challenges. Another challenge lies in the need for higher computational power when integrating FCM in large-scale datasets, limiting its use in resource-constrained environments.

6. References

[1] Viola, P., & Jones, M. (2001). Robust real-time object detection. International journal of computer vision, 4(34-47).

[2] Zhonglong, Z., Yang, J., & Zhu, Y. (2006). Face detection and recognition using colour sequential images. Journal of Research and Practice in Information Technology, 38(2), 135-149.

[3] Lu, Y., Zhou, J., & Yu, S. (2003). A survey of face detection, extraction and recognition. Computing and informatics, 22(2), 163-195.

[4] J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia, and S. Zafeiriou, "Retinaface: Single-stage dense face localisation in the wild," ArXiv, vol. abs/1905.00641, 2019.

[5] P. Viola and M. J. Jones, "Robust real-time face detection. International journal of computer vision," vol. 57, no. 2, pp. 137-154, 2004.

[6] B. Yang, J. Yan, Z. Lei, and S. Z. Li, "Aggregate channel eatures for multi-view face detection," in IEEE International Joint Conference on Biometrics, 2014, pp. 1-8.

[7] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in Advances in neural information processing systems, 2012, pp. 1097-1105.

[8] Y. Sun, Y. Chen, X. Wang, and X. Tang, "Deep learning face representation by joint identification-verification," in Advances in Neural Information Processing Systems, 2014, pp. 1988-1996.

[9] H. Li, Z. Lin, X. Shen, J. Brandt, and G. Hua, "A convolutional neural network cascade for face detection," in IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 5325-5334.

[10] S. Yang, P. Luo, C. C. Loy, and X. Tang, "From facial parts responses to face detection: A deep learning approach," in IEEE International Conference on Computer Vision, 2015, pp. 3676-3684.

[11] C. Zhang, and Z. Zhang, "Improving multiview face detection with multi-task deep convolutional neural networks," IEEE Winter Conference on Applications of Computer Vision, 2014, pp. 1036-1041.

[12] P. V. Lal, U. Srilakshmi and D. Venkateswarlu, "Face Recognition system using Extreme Learning Machines," 2021 5th International Conference on Information Systems and Computer Networks (ISCON), Mathura, India, 2021, pp. 1-5, doi: 10.1109/ISCON52037.2021.9702467.

[13] Y. Zhou, D. Liu and T. Huang, "Survey of Face Detection on Low-Quality Images," 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), Xi'an, China, 2018, pp. 769-773, doi: 10.1109/FG.2018.00121.

[14] W. Louis and K. N. Plataniotis, "Weakly trained dual features extraction based detector for frontal face detection," 2010 IEEE International Conference on Acoustics, Speech and Signal Processing, Dallas, TX, USA, 2010, pp. 814-817, doi: 10.1109/ICASSP.2010.5494933.

[15] D. A. Karras, "A Robust Feature Extraction Methodology for Improved Face Detection Performance within a Complex Background," 2007 IEEE International Workshop on Imaging Systems and Techniques, Cracovia, Poland, 2007, pp. 1-5, doi: 10.1109/IST.2007.379596.

[16] A. -C. Tsai, Y. -Y. Ou, W. -C. Wu and J. -F. Wang, "Occlusion Resistant Face Detection

and Recognition System," 2020 8th International Conference on Orange Technology (ICOT), Daegu, Korea (South), 2020, pp. 1-4, doi: 10.1109/ICOT51877.2020.9468767.

[17] M. Jha, A. Tiwari, M. Himansh and V. M. Manikandan, "Face Recognition: Recent Advancements and Research Challenges," 2022 13th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2022, pp. 1-6, doi: 10.1109/ICCCNT54827.2022.9984308.

[18] S. W. Arachchilage and E. Izquierdo, "A Framework for Real-Time Face-Recognition," 2019 IEEE Visual Communications and Image Processing (VCIP), Sydney, NSW, Australia, 2019, pp. 1-4, doi: 10.1109/VCIP47243.2019.8965805.

[19] W. Wang, X. Wang, W. Yang and J. Liu, "Unsupervised Face Detection in the Dark," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 1, pp. 1250-1266, 1 Jan. 2023, doi: 10.1109/TPAMI.2022.3152562.

[20] V. Kumar, "A Multi-Face Recognition Framework for Real Time Monitoring," 2021 Sixth International Conference on Image Information Processing (ICIIP), Shimla, India, 2021, pp. 30-34, doi: 10.1109/ICIIP53038.2021.9702591.