International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 12 Dec. 2024, pp: 291-298 www.ijaem.net ISSN: 2395-5252

An Improved System for Fiducial Point Detection under Various Occlusions

[1]Habib Shehu Jibrin, [2]Nyako Alhaji Baba, [3]Ali Abubakar, [4]Usman Adamu Muhammad

[1,2,3,4]Lecturers, Department of Computer Science, Federal Polytechnic Kaltungo, Gombe, Nigeria.
Corresponding Author: Habib Shehu Jibrin

Date of Submission: 10-12-2024 Date of Acceptance: 20-12-2024

ABSTRACT: Fiducial point detection is essential in applications like facial recognition, medical imaging, and human-computer interaction. Despite advancements in deep learning, occlusions present a significant challenge in face recognition systems. This research proposes an enhanced facial landmark detection system that utilizes Principal Component Analysis (PCA) and Active Shape Models (ASM) to improve detection accuracy under various occlusions, lighting conditions, and expressions. By employing color segmentation for face detection and landmark detection to compensate for occlusions, the system provides a robust solution for accurate recognition in complex environments. The model was evaluated using datasets such as PICS, UMB, and HELEN, achieving a detection accuracy of over 97% across various conditions. Experimental results demonstrate the system's superior performance in face detection and occlusion handling compared to existing methods, making it a valuable tool for real-world applications in security, biometrics, and human-computer interaction.

KEYWORDS: Deep learning, Detection, Facial recognition, Fiducial point, Occlusion.

I. INTRODUCTION

Fiducial point detection is a fundamental component in various applications, including facial recognition, medical imaging, and human-computer interaction. With the advancements in deep learning, face recognition methods have achieved remarkable performance, often surpassing human capabilities on benchmark datasets [1]; [2]; [3].

Face recognition has been a subject of extensive research for decades [4]. Compared to other biometric modalities like fingerprints, hand geometry, iris, signature, and voice (Figure 1), facial features offer a less intrusive and more convenient means of identification [1]. This has led to the widespread adoption of face recognition in applications such as surveillance, forensics, and border control [2]. The development of deep learning techniques and the availability of large-scale face datasets have significantly improved face recognition performance [3].

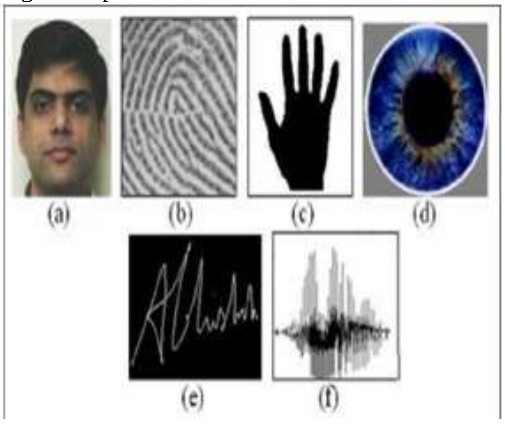

Figure 1: Typical Examples of biometric characteristics: (a) face, (b) fingerprint, (c) hand geometry, (d) iris, (e) signature, and (f) voice[4]

Occlusions, whether natural or synthetic, continue to pose significant challenges for face recognition systems. Gallery images are typically of high quality, while probe images often suffer from missing data due to occlusions [1]. This mismatch between available features in probe and gallery faces can hinder recognition accuracy.

Facial occlusion is a particularly difficult problem because the occluded part can be unpredictable in terms of location, size, and shape [5]. Collecting a large-scale training dataset with all possible occlusions is impractical, making it challenging to apply deep learning techniques effectively. Occlusions can occur due to accessories like scarves, masks, glasses, hats, or

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 12 Dec. 2024, pp: 291-298 www.ijaem.net ISSN: 2395-5252

random objects obstructing the face. These occlusions can increase intra-class variations and reduce inter-class similarity, compromising recognition accuracy [1]. Figure 2 illustrates how facial appearance can be significantly altered by occlusions.


Figure 2: Example of Facial Images with Occlusions[1].

Table 1 categorizes the various occlusion challenges across different scenarios, along with typical examples of occlusions [1].

Occlusion Scenario	**Examples**
Facial accessories	Eyeglasses, face masks, scarves, hat, hair
External occlusions	Occluded by hands and other random objects
Partially captured faces	Partially captured due to Limited field of view
Self-occlusions	Non-frontal pose
Extreme illumination	Part of face highlighted
Artificial occlusions	Occluded with random white or black rectangles, random salt & pepper noise

To address these drawbacks, this work proposes an improved system for fiducial point detection that can accurately detect faces and facial features under various occlusions, illuminations, and expressions.The system employs a model that uses Principal Component Analysis (PCA) to extract features from face images and reduce image dimensions, allowing for faster computation and memory conservation. Additionally, it compensates for occlusions, illumination, and expression changes in the test images using landmark detection, ensuring robust face recognition across different environments and use cases. This proposed solution overcomes the limitations of existing methods, providing a more reliable approach to fiducial point detection even in complex, occluded settings.

II. LITERATURE REVIEW

Occluded face detection is a crucial area of research in computer vision with significant applications in security, biometrics, and human-computer interaction. This literature review provides an overview of face detection techniques under occluded conditions.

A. Face detection under occlusions

Detecting occluded objects presents significant challenges in unconstrained environments, particularly when large areas are obscured due to intra-class similarity and variations. Various approaches have been explored to address this issue, including convolutional correlational filters [6].

Existing face detection models often experience reduced accuracy when dealing with occluded faces. To mitigate this, specialized algorithms for occluded face detection have been developed [6]. These algorithms aim to enhance the robustness and accuracy of face detection systems in the presence of occlusions.\

B. General Face Detection

Face detection algorithms generally perform well in unconstrained environments and can be categorized into three main approaches: rigid templates, deformable part models, and convolutional neural networks (ConvNets).

The Viola-Jones face detection model, which uses Haar-like features and AdaBoost, falls under the rigid template category but may exhibit decreased performance in real-time applications [7]. Despite its efficacy, its performance can drop significantly in real-world scenarios [8]. In contrast, deformable part models (DPMs) are more suitable for real-time applications but come with high computational complexity [9].

The most promising recent development is the use of deep convolutional neural networks (DCNNs) for face detection [10]. Some methods integrate face detection with face alignment to leverage their inherent correlation, thereby enhancing performance [11]. Recent advancements in DCNNs have yielded impressive results, particularly with the introduction of the Widerface benchmark, which encompasses wide pose variations, significant scale differences, expression variations, makeup, severe illumination, and occlusion [12]. The RetinaFace model, a single-stage pixel-wise face localization method, employs extra-supervised and self-supervised multi-task learning, achieving 92.5% accuracy on the Widerface hard test set [13].

Handling occlusions in face detection is

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 12 Dec. 2024, pp: 291-298 www.ijaem.net ISSN: 2395-5252

challenging due to the variability and unknown locations of occlusions [6]. Efforts to detect occluded faces can be categorized into three main strategies: locating visible facial segments, discarding features from occluded sub-regions, and utilizing occlusion information. Attribute-aware CNNs categorize facial features based on properties like prominent lips or large eyes to create a facial response map, often achieving accuracies around 99% [1]. Approaches such as FAN, LLE-CNN, and AdaBoost cascade classifiers are typically trained as segment-based face detectors that focus on discarding features from occluded sub-regions [1].

C. Detecting Occluded Faces

Detecting partially occluded faces involves locating the face region in an image where occlusions are present [1]. The challenge in handling occlusions lies in the unknown location and type of occlusions [14]. Recently, there has been increased interest in occluded face detection, and several publications have explored this area. Techniques from pedestrian detection, a well-studied field, have been adapted to improve occluded face detection by addressing occlusion as the primary challenge [24]; [25]; [26]. This method emphasizes the need for training models to recognize and segment occluded areas, which is crucial for accurate face detection [15].

Another innovative approach involves using fuzzy clustering methods to detect and recover occluded facial regions. Studies have demonstrated that applying fuzzy C-means clustering can effectively identify and reconstruct occluded areas, thereby improving detection accuracy [16]. This method underscores the potential of advanced clustering techniques in enhancing the recovery of occluded facial features.

Additionally, the Sparse Representation Classification (SRC) method has gained attention in the field of occluded face recognition. SRC utilizes a linear combination of training images to reconstruct unoccluded faces, which can enhance detection accuracy in scenarios where traditional methods may struggle [17]. Despite its promise, SRC faces challenges with scalability, highlighting the need for further research into more efficient algorithms.

Furthermore, the integration of deep learning techniques, such as Long Short-Term Memory (LSTM) Autoencoders, has been explored for face de-occlusion [18]. A proposed method uses LSTM Autoencoders to automatically remove facial occlusions before recognition, demonstrating the effectiveness of deep learning in addressing occlusion challenges [18]. This approach marks a shift towards leveraging advanced neural networks to improve face detection capabilities in the presence of occlusions.

III. RELATED WORK

The problem of 3D face recognition under occlusion has become increasingly relevant due to the widespread use of facial recognition systems in security applications. Researchers have explored various strategies to detect and restore occluded facial regions. [23] provided a comprehensive review of existing approaches, highlighting their limitations and identifying areas for future research. These methods, based on image processing, machine learning, and deep learning, aim to detect occlusions and reconstruct missing facial information to enhance 3D face recognition performance. Similarly, [24]addressed occlusion robustness in 2D face recognition, proposing the Occlusion-aware face REcOgnition (OREO) approach. OREO uses an attention mechanism to extract local identity-related regions while combining them with global representations to improve generalization under occlusions, achieving significant performance improvements in both single-image and image-set scenarios.

categorized methods of face recognition under occlusion into three primary approaches: occlusion-robust feature extraction, occlusion-aware face recognition, and occlusion recovery. They analyzed the strengths and weaknesses of these approaches, providing valuable insights for future research. [25] addressed the limitations of face recognition datasets by developing Webface-OCC, a simulated occlusion dataset. This dataset, which includes various occlusion types such as masks and glasses, improved the performance of models like ArcFace, showing enhanced recognition in occluded scenarios.

examined the challenges introduced by face masks during the COVID-19 pandemic, which hinder facial recognition systems. Despite advancements in deep learning, they emphasized that masks and other large-area occlusions remain one of the most formidable challenges for face recognition algorithms. Lastly, [9] explored the growing role of face recognition in various applications, noting that occlusion remains a major challenge that impacts system accuracy, especially in real-world conditions. They called for further research to improve the robustness of face recognition under occluded scenarios.

IV. MATERIAL AND METHOD

The framework for facial landmark detection in occluded scenarios comprises three key modules as shown in Figure 3: face detection, landmark detection, and occlusion detection. The face detection module preprocesses images and uses color segmentation to identify faces. The second module applies an Active Shape Model (ASM) for iterative fitting to facial landmarks, involving shape labeling, alignment, variation modeling, and fitting. The final module detects and estimates occlusion regions by assessing the amount of detail covering facial features. This structured approach aims to improve performance and applicability in real-world settings where occlusions pose challenges.

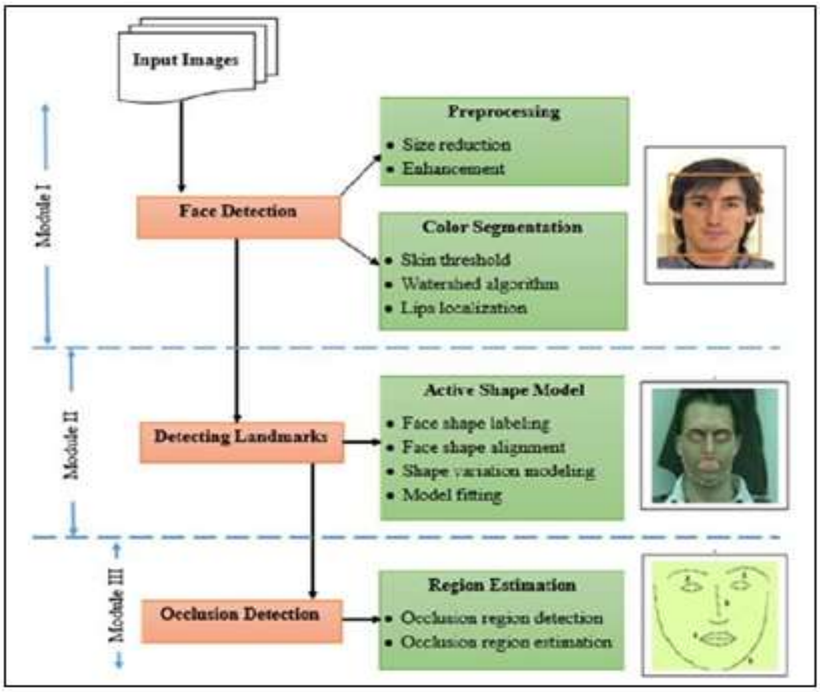

Figure 3: Framework of the Proposed Model for Facial Landmark Detection under Various Occlusions

A. FACE DETECTION MODULE

In building a face detection tool, preprocessing is key to handling issues like lighting, background complexity, and noise. It simplifies the image, making face detection easier and faster. The tool uses color segmentation to efficiently locate faces without overloading the computer. Adjusting contrast helps with poor lighting, and resizing images reduces computational load. By exploring different color models like RGB, CMYK, YCbCr, HSV, and Lab*, the process ensures effective face detection even under varying lighting conditions.

The proposed technique finds the maximum energy of the histogram signal for skin, limited to the narrowest ranges for each component of the evaluated color spaces (c1, c2, ..., ck). The maximum value is calculated by taking the cube root of the average component energy across the selected color space ranges.Figure 4 shows the proposed framework for the face detection using skin color information.

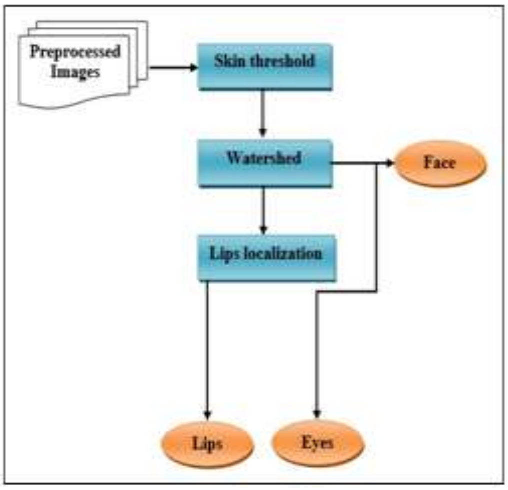

Figure 4: Proposed Framework for Face Detection using Skin Color Information

The face detection process identifies skin regions using a **histogram signal** that evaluates the energy distribution across different components of selected color spaces. The segmentation process isolates facial features by utilizing a **watershed algorithm** to remove noise and then applies **lips localization** using the color differences in the lips region. The expected result of the face detection module is given in Figure 5.

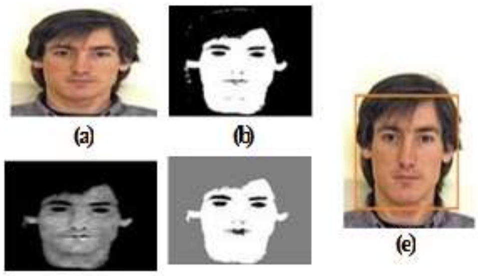

Figure 5:Expected Result of the Face Detection Module

From the above figure, image (a) represents the preprocessed image, (b) represents the segmented skin region, (c) the most likely face area, (d) the localized eye and nose and (e) the face region of interest in that order.

B. DETECTING LANDMARK MODULE

The **Active Shape Model (ASM)** is employed for facial landmark detection, which involves several steps:

1. **Labeling Training Images**: Training data consists of manually annotated images, where facial landmarks are labeled as coordinate points representing key facial features (e.g., eyes, nose, mouth).
2. **Face Shape Alignment**: The ASM aligns shapes using **Procrustes Analysis** to remove variations caused by translation, scaling, and rotation. This standardizes the shapes for statistical analysis.
3. **Shape Variation Modeling**: **Principal Component Analysis (PCA)** is applied to analyze shape variations, reducing the dimensionality while capturing essential features of the face.

Model Fitting: The ASM iteratively refines the position of each landmark point to fit the face, adjusting parameters until the model converges to the best possible configuration.

C. OCCLUSION DETECTION MODULE

The landmark detection approach presented here utilizes 2D data to identify potential points of interest, which are then classified and labeled as anatomical landmarks. In developing the facial landmark model, an anatomical landmark system is employed, dividing the face into five distinct regions as illustrated in Figure 5. This regional division methodology aligns with the approaches previously outlined in the work of [30], building upon established frameworks in the field of facial analysis and landmark detection.

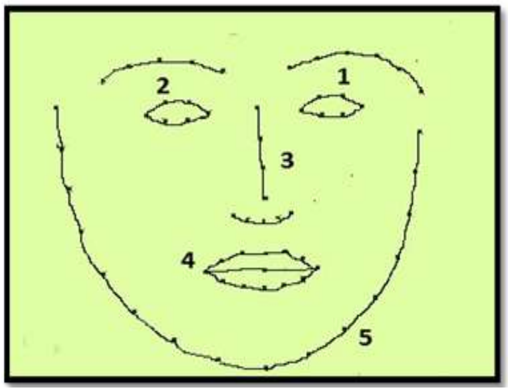

Figure 6: Face Shape Model Divided into Five Regions

The facial regions used in the study include five key areas: the Right Eye and Right Eyebrows (REREB), the Left Eye and Left Eyebrows (LELEB), the Nose Tips (NT), the Mouth Tips (MT), and the Edges of the Face (FE). These regions were critical for facial landmark detection and analysis in the experiments.

The landmark points from the study's complete set of 66 landmarks are categorized into five regions, which are typically visible on frontal and nearly frontal faces, except when occlusions or semi-profile views are present. The positioning of these landmark points according to each region is detailed in Table 2.

Table 2: Regions and their Corresponding Positions of the Landmark Points

Region name	Position of the landmark points	No. of landmark points
REREB	1 – 11	11
LELEB	12 – 22	11
NT	23 – 31	9
MT	32 – 48	16
FE	59 – 66	19

V. RESULTS AND DISCUSSION

A. EXPERIMENTAL DATA

The proposed facial landmark detection model was evaluated using three datasets: PICS, UMB, and HELEN. The PICS dataset was used to

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 12 Dec. 2024, pp: 291-298 www.ijaem.net ISSN: 2395-5252

select the best color model for face detection, while the UMB dataset was utilized for evaluating facial landmark detection and occlusion detection. The HELEN dataset enabled comparisons with other studies. These datasets provided a comprehensive assessment of the model's performance across different facial analysis tasks, demonstrating robust results in face recognition research.

DATASET

The HELEN dataset consists of 2,000 high-quality, frontal face images sourced from Flickr, selected through a rigorous process to ensure image quality and exclude profile views and false positives. The images were precisely annotated by Amazon Mechanical Turk workers, marking key facial features such as the nose, eyes, eyebrows, jaw, and mouth. The dataset, in JPEG format with a minimum resolution of 500 x 500 pixels, provides a robust resource for facial analysis research and algorithm development.

EXPERIMENT SETUP

In this research, images from three selected databases were used as training sets for face detection, incorporating various conditions that could affect the detection process. The PICS database provided equal numbers of male and female images for training, with male images labeled from Pm1 to Pm18 and female images from Pf1 to Pf18. Figure 7 illustrates samples of these training images.

Figure 7: Sample of Images Obtained from UMB Database

B. EXPERIMENTAL PROCEDURES AND RESULTS

In Procedure I, face images were evaluated using five color models: RGB to YUV, HSV, CMYK, and YCbCr. The YCbCr model proved most effective in determining skin color for the PICS and UMB datasets. The proposed method showed high detection accuracy, with PICS achieving a 97.51% detection rate and UMB 96.00%, outperforming previous studies using standard datasets.

In Procedure II, facial landmarks were detected under various occlusions. Landmarks for eyes, eyebrows, nose, lips, and face edges were annotated. The model maintained a 99.21% detection rate across different image sizes. It handled occlusions like eyeglasses, caps, and multiple objects, with eyeglasses achieving the highest detection rate (99.04%) and multiple objects the lowest (94.65%). The model showed faster detection times and higher accuracy compared to other methods but struggled with extreme head movements and strong occlusions.

VI. CONCLUSION

This research demonstrates that addressing the issue of sparsity at a deeper level can eliminate the need for traditional feature extraction methods. While traditional Principal Component Analysis (PCA) has been effective for over a decade under ideal conditions, it often fails in inconsistent environments, such as those involving partial occlusion or significant changes in illumination. The integration of Active Shape Models with PCA for compressive sampling presents a novel approach in computer vision. This combined method allows for improved representation of test images and effective modeling of occlusions or noise. The algorithm has been rigorously tested across various scenarios, including different lighting conditions, facial expressions, poses, and occlusions. The experimental results indicate that the proposed algorithm consistently outperforms other methods in all test cases. The incorporation of Active Shape Models with PCA marks a significant advancement in the field, particularly in image processing and face recognition, due to its enhanced robustness and potential. This makes it a highly attractive solution for automating facial feature extraction.

ACKNOWLEDGEMENT

The authors express their profound gratitude to the Tertiary Education Trust Fund (TETFund) for their invaluable support. Their financial contribution has been essential in making this research possible.

REFERENCES

[1]. D. Zeng, R. Veldhuis and L. Spreeuwers, "A survey of face recognition techniques under occlusion," IET biometrics, vol. 10, no. 6, pp. 581-606, 2021.

[2]. H. Wang, Y. Wang, Z. Zhou, X. Ji, D. Gong, J. Zhou, ... and W. Liu, "Cosface: Large margin cosine loss for deep face recognition," in In Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake, USA, 2018.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 6, Issue 12 Dec. 2024, pp: 291-298 www.ijaem.net ISSN: 2395-5252

[3]. D. S. Trigueros, L. Meng and M. Hartnett, Face recognition: From traditional to deep learning methods, arXiv preprint, 2018.
[4]. F. O. Aranuwa, " Efficient Multimodal Biometric Authentication System Based on Multilevel Decision Threshold," Technology, vol. 8, no. 2, pp. 49-58, 2020.
[5]. M. O. Oloyede, G. P. Hancke and N. Kapileswar, "Evaluating the effect of occlusion in face recognition systems," in In 2017 IEEE AFRICON, Cape Town, South Africa, 2017.
[6]. S. Hemathilaka and A. Aponso, "An Analysis of Face Recognition under Face Mask Occlusions," in In CS & IT Conference Proceedings (Vol. 11, No. 18), 2021.
[7]. P. Viola and M. J. Jones, "Robust real-time face detection," International journal of computer vision, vol. 57, no. 2, pp. 137-154, 2004.
[8]. B. Yang, J. Yan, Z. Lei and S. Z. Li, "Aggregate channel features for multi-view face detection," in In IEEE international joint conference on biometrics, Clearwater, FL, USA, 2014.
[9]. J. Yan, Z. Lei, L. Wen and S. Z. Li, "The fastest deformable part model for object detection," in In Proceedings of the IEEE conference on computer vision and pattern recognition, Ohio, USA, 2014.
[10]. Y. Bai, Y. Zhang, M. Ding and B. Ghanem, "Finding tiny faces in the wild with generative adversarial network," in In Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake City, USA, 2018.
[11]. K. Zhang, Z. Zhang, Z. Li and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," IEEE signal processing letters, vol. 23, no. 10, pp. 1499-1503, 2016.
[12]. S. Yang, P. Luo, C. C. Loy and X. Tang, "Wider face: A face detection benchmark," in In Proceedings of the IEEE conference on computer vision and pattern recognition , Las Vegas, USA, 2016.
[13]. J. Deng, J. Guo, Y. Zhou, J. Yu, I. Kotsia and S. Zafeiriou, Retinaface: Single-stage dense face localisation in the wild, arXiv preprint, 2019.
[14]. J. Chen, S. Shan, S. Yang, X. Chen and W. Gao, "Modification of the adaboost-based detector for partially occluded faces," in In 18th International Conference on Pattern Recognition (ICPR'06), Hong Kong, China, 2016.
[15]. Y. Chen, L. Song, Y. Hu and R. He, "Adversarial occlusion-aware face detection," in In 2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS), 2018.
[16]. W. Hadi, "Face occlusion detection and recovery using fuzzy c-means," Engineering and Technology Journal, vol. 28, no. 18, pp. 5744-5756, 2010.
[17]. Y. Li, "Face detection algorithm based on double-channel cnn with occlusion perceptron," Computational Intelligence and Neuroscience, vol. 2022, pp. 1-10, 2022.
[18]. F. Zhao, J. Feng, J. Zhao, W. Yang and S. Yan, "Robust lstm-autoencoders for face de-occlusion in the wild," IEEE Transactions on Image Processing, vol. 27, no. 2, pp. 778-790, 2018.
[19]. J. Deng, J. Guo, N. Xue and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition," in In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, California, USA, 2018.
[20]. W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj and L. Song, "Sphereface: Deep hypersphere embedding for face recognition," in In Proceedings of the IEEE conference on computer vision and pattern recognition, Honolulu County, USA, 2017.
[21]. B. Liu, W. Deng, Y. Zhong, M. Wang, J. Hu, X. Tao and Y. Huang, "Fair loss: Margin-aware reinforcement learning for deep face recognition," in In Proceedings of the IEEE/CVF international conference on computer vision, Seoul, Korea, 2019.
[22]. L. Best-Rowden and A. K. Jain, "Longitudinal study of automatic face recognition," IEEE transactions on pattern analysis and machine intelligence, vol. 40, no. 1, pp. 148-162, 2018.
[23]. R. Ranjan, S. Sankaranarayanan, A. Bansal, N. Bodla, J. C. Chen, V. M. Patel, ... and R. Chellappa, "Deep learning for understanding faces: Machines may be just as good, or better, than humans," IEEE Signal Processing Magazine, vol. 35, no. 1, pp. 66-83, 2018.
[24]. Z. Guo, W. Liao, P. Veelaert and W. Philips, "Occlusion-robust detector trained with occluded pedestrians," in In 7th International Conference on Pattern

Recognition Applications and Methods ICPRAM 2018, Madeira, Portugal, 2018.
[25]. S. Zhang, L. Wen, X. Bian, Z. Lei and S. Z. Li, "Occlusion-aware R-CNN: Detecting pedestrians in a crowd," in In Proceedings of the European conference on computer vision (ECCV), Munich, Germany, 2018.
[26]. C. Zhou and J. Yuan, "Bi-box regression for pedestrian detection and occlusion estimation," in In Proceedings of the European Conference on Computer Vision (ECCV) , Munich, Germany, 2018.
[27]. X. Xu, N. Sarafianos and I. A. Kakadiaris, "On improving the generalization of face recognition in the presence of occlusions," in n Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, VIRTUAL, 2020.
[28]. N. Dagnes, E. Vezzetti, F. Marcolin and S. Tornincasa, "Occlusion detection and restoration techniques for 3D face recognition: a literature review," Machine Vision and Applications, vol. 29, pp. 789-813, 2018.
[29]. B. Huang, Z. Wang, G. Wang, K. Jiang, K. Zeng, Z. .. Han and Y. Yang, "When face recognition meets occlusion: A new benchmark," in In ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Toronto, Canada, 2021.