



RESEARCH ARTICLE – COMPUTER SCIENCE

CNN-Mesh: Enhancing Face Recognition with Convolutional Neural Networks and Mesh Algorithm

Nibras Abdel Mohsen^{1*}, Enas Mohammed Hussien Saeed²

¹ N Master's Student at Mustansiriyah University

² E Doctor at the University of Mustansiriyah

* Corresponding author E-mail:

Article Info.	Abstract
<i>Article history:</i> Received 30 July 2024 Accepted 19 September 2024 Publishing 30 September 2025	Several areas of AI have found uses for face recognition (FR) technologies, such as biometrics, authentication, security, surveillance, and law enforcement. Results from FR using deep learning (DL) models, especially CNNs, have been promising. By utilizing learning methods such as Convolutional Neural Networks (CNN) and mesh algorithm to improve face recognition accuracy. our goal is to ensure privacy protection while tackling issues related to varying lighting conditions and facial expressions. The proposed model (CNN-Mesh) is combined CNN algorithm with mesh technique for prediction the fully landmark of face. the CNN-Mesh approach performs in comparison to that of other facial recognition methods that are currently in use. The purpose of this study is to determine whether or not the integrated model offers improved resilience against variables such as varied facial expressions, age, or cosmetics, in addition to speed and computing efficiency. The purpose of this research is to determine the extent to which merging CNNs with Mesh algorithms results in improvements in both accuracy and processing speed. The combination has two stage: first stage is that the network collects feature extraction to solve the classification problem. Second stage apply mesh network after collects whole features for used it to describe the edge of faces. This study investigates how convolutional neural networks (CNNs), which are particularly effective at extracting hierarchical information from images, with used to provide Mesh networks, which are primarily concerned with predicting certain face landmarks, with a more accurate input. The purpose of this study is to determine which elements retrieved by CNNs (such as edges, textures, and deep semantic representations) are the most helpful for the succeeding Mesh network to reliably predict face structures. Moreover, it able the network to construct predictions based on those traits. During the process of learning, the model is instructed to recognize patterns and traits in the images that are associated with certain individuals. Our strategy involves the use of 68 Landmark markers to optimize facial recognition precision resulting in advancements in this field. The results of our CNN-Mesh algorithm achieved a 97% accuracy rate for classification and detecting faces, while the accuracy of the model is increased after mesh network to 99%. Moreover, we train ResNet model for comparative with our model which achieve accuracy 31.73%. while CNN 88.46%.

This is an open-access article under the CC BY 4.0 license (<http://creativecommons.org/licenses/by/4.0/>)

The official journal published by the College of Education at Mustansiriya University

Keywords: Deep learning, Face Recognition, CNN, Mesh Network, Face landmarks, Deep learning, ResNet.

1. Introduction

In recent years, artificial intelligence has entered many areas of life [S. Kadhem, Z. Ali, and A. Suhad Malallah Zuhair Hussein Ali, 2017] ,[A. A. Maryoosh and S. Pashazadeh , 2024]. In this changing environment the uses of recognition technology go beyond security purposes to areas, like healthcare, education and retail garnering interest, in recent years like healthcare, education and retail garnering interest, in recent years [I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed,2020]. In the past face recognition technology involved steps, like taking image of faces identifying faces preparing images

extracting characteristics and making matches using those features. Advancements, in networking and the widespread adoption of learning have made deep learning based facial recognition techniques the practice. With all these progressions recognizing faces can still be tricky because of things, like varying conditions, age variations, obstacles, makeup changes and shifts, in posture [K. Zhang, Y. L. Chang, and W. Hsu, 2018]. Furthermore, there are issues regarding recognition that require deeper investigation in this area. Other studies have also explored methods, for combining images to tackle challenges associated with changes in perspectives. The advancements in technology and research has integrated facial recognition technology into our routines. Convolutional neural networks (CNNs) are commonly acknowledged as tools, for image recognition known for their proficiency in classifying images. [C. R. Vishwanatha, V. Asha, B. Saju, N. Suma, T. R. M. Reddy, and K. M. Sumanth, 2023] The need, for solutions in identification and verification is increasing, with features offering valuable insights into a person's identity and age. Facial recognition stands out as a method that can be used without consent making it particularly useful for surveillance purposes. Additionally, face recognition plays a role in biometrics by capturing both the aspects of an individual's face. However, recognizing faces in environments presents challenges due to factors like head position variations, lighting conditions, age related changes and emotional expressions [R. Alejandra Di Mattia, 2023.]. External elements such as hair, makeup or accessories like scarves or glasses can further complicate the process. An additional hurdle arises when identifying individuals who share similarities—such as twins or relatives—adding complexity to facial recognition systems. Unlike humans who effortlessly distinguish faces computers struggle with processing information due, to the pathways involved in the human visual system that analyze both static features and subtle movements to differentiate faces within their surroundings. The suggested model performs better than current baseline face recognition algorithms due to its unique equivariance features. The problem statement that we tackled in our study is that working with and managing sets of data can be quite demanding in terms of computational resources and time. Consequently, the disproportion of data in The problem of data imbalance is a typical problem that arises in face recognition datasets. This problem occurs when certain persons have samples relative to others. Because of this imbalance, you may end up with biased models that have difficulty identifying people. Taking action to rectify this data imbalance is very necessary in order to guarantee accurate face recognition systems. The use of face recognition technology, on the other hand, raises worries about privacy and ethics in terms of monitoring and the possible abuse of personal information. These concerns are a result of the fact that the technology is being used. Overall, the contribution of our experiment can be describe as below:

- 1- the convolved model have achieved superior accuracy to solve the classification problem.
- 2- preprocessing images and fine tune the model apply to train the network and increase the enhancement prediction.
- 3- time consuming for training stage is very crucial problem, in our experiment the time of training is very faster to finish all epochs in less than ten second.
- 4- The proposed model has undergone testing in four distinct scenarios: analysis of face image resolution, analysis of face position, analysis of face occlusion and non-uniform lighting, and analysis of face rotation with non-uniform illumination and blurred face images.
- 5- combination the Convolutional network with mesh network to develop and enhance the prediction of face-landmark.

2. Literature Review

Facial recognition technology has made significant progress in terms of accuracy, as shown by its performance on different datasets. Several studies have explored the use of Convolutional Neural Networks (CNN) in learning for recognizing faces. Zhou and Yu demonstrated face recognition accuracy, up to 99.8% by combining a binary pattern with a customized LeNet 5 CNN model. [L. Zhou and W. Yu,2022] utilized models, like CNN, ResNet50, VGG16 and InceptionV3 to achieve accuracy levels. Notably Inception V3 achieved an accuracy rate of 97.93% on the FER 2013 dataset. The CNN model was employed as a feature extractor for emotion detection. However deep learning models such as CNNs have their limitations, including training times susceptibility to inputs and difficulty, in learning multiple tasks consecutively.

[N. Bukhari, S. Hussain, M. Ayoub, Y. Yu, and A. Khan,2022] The authors employed a convolutional neural network to detect and recognize masks in real time, within crowded public areas and communal settings particularly where face recognition poses challenges due to increased target density. Meanwhile Sapakova and Yilibule utilized a YOLOv5 model for real time detection and recognition of masks achieving a average precision (mAP) value of 92.9%, for recognizing face masks.

The method introduced a Hybrid Ensemble CNN model, for recognition achieving accuracies of 99.35% on LFW 91.58% on cross pose LFW and 95% on a custom criminal dataset. Although the study discusses the application of training to improve recognition performance it falls short in providing a breakdown or evaluation of the computational resources and time investment needed for this training regimen, which holds significant importance, for real world application. [S. Anwarul, T. Choudhury, and S. Dahiya,2023] investigated how various machine learning algorithms affect emotion recognition by utilizing trained CNN features. They found that SVM and Ensemble classifiers performed better on Alex Net and VGG16 architectures leading to enhanced accuracy.

The study analyzed the impact of classifiers, on the FC6, FC7 and FC8 layers of CNNs; however, it did not provide an explanation for why SVM and Ensemble classifiers surpassed the SoftMax classifier on these layers. This gap in information suggests the need, for a examination of the underlying factors. [H. M. Shahzad, S. M. Bhatti, A. Jaffar, S. Akram, M. Alhajlah, and A. Mahmood,2023].

2.1 Face recognition

Facial recognition involves using technology to compare an individuals captured image with a collection of photos, for identification purposes. There are three types of facial recognition methods; 2D, 2D 3D and 3D. Studies have indicated that advanced learning algorithms, like networks (CNNs) have been instrumental, in this field [12 D. S. Ramesh, S. Heijnen, O. Hekster, L. Spreewers, and F. de Wit,2022] Facial expressions can be identified with precision by paying attention to features, like the nose and mouth. Additionally, Research has been studied confirming its importance in recognition individuals involved in criminal incidents 許琇娟,2023]. Using recognition models, in scenarios like identifying Roman emperors from their portraits has yielded positive outcomes using methods such, as learning through transportation [A. J. Russ, M. Sauerland, C. E. Lee, and M. Bindemann,2018].

[M. He, J. Zhang, S. Shan, M. Kan, and X. Chen, 2020] For the purpose of facial identification, a Deformable Facial Net (DFN) is used to manage the various poses. With the help of the deformable convolution module, an effort is made to concurrently learn face recognition directed alignment and identity-preserving feature extraction. It is recommended that the displacement consistency loss (DCL) be used to ensure that the learned displacement field is locally consistent in both the orientation and the amplitude. This would result in improved alignment for face recognition. It has been shown via extensive testing that the DFN that has been developed is capable of achieving pretty promising performance with a relatively light network topology, particularly for those big postures.

[Y. Taigman, M. Yang, M. A. Ranzato, and L. Wolf, 2014] In terms of accuracy, the Deep Convolutional Neural Network (DCNN)-based FR algorithm that you suggest is superior to a few well-known approaches. Even when there is a significant amount of posture fluctuation and lighting circumstances, the authors claim to have achieved substantially greater accuracy. Furthermore, the method was subjected to training and testing on the LFW dataset, the SFC dataset, and the YouTube Face dataset (YTF). system (DeepFace) that has closed the majority of the remaining gap in the most prominent benchmark in unconstrained facial recognition, DeepFace has managed to bridge the gap. [M. J. Er, S. Wu, J. Lu, and H. L. Toh, 2002] The authors describe a method that is both effective and efficient for face recognition. This method is based on the Radial Basis Function (RBF), which helps to decrease the amount of computing work required and prevents over-fitting. Within the framework of the suggested method, the Principal Component Analysis (PCA) technique is first used to extract low-level facial characteristics. After some time has passed, the features that were produced are subjected to Fisher's Linear Discriminant (FLD) approach in order to gain discriminant patterns with reduced dimensions. [L. Lu, Y. Mao, L. Wenxin, and H. J. Zhang, 2003] The Multiple Classifier System (MCS) has been created by an intelligent fusion of Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Independent Component Analysis (ICA). The method was evaluated on the ORL, Yale, AR, and a dataset generated by the authors. The published data demonstrate an average accuracy over 90%.

3. Methodology

Preprocessing for face recognition is a series of techniques aimed at preparing input images to achieve accurate and efficient face identification. The pixel resolution of images differs across different databases. For this experiment, the size of these photos was decreased to 250×250 pixels. The first step of our suggested method involves locating and aligning the face in order to ensure that the eyes, nose, and mouth are positioned identically in every image. The suggested technique used Face Landmark Localization (FLL) because to its ability to align the face by identifying 68 particular landmarks in an image. Therefore, certain attributes of these properties are recognized in order to determine the position of the human face. The face displays significant features such as the top portion of the chin, lips, eyes, nose, and eyebrows. Therefore, to identify the human face, certain attributes of these qualities are established. The FLL efficiently handles variations in posture. The process of face alignment involves creating a triangle by linking the locations of the eyes on facial landmarks to the normal vector of the image. Subsequently, the angles between the eye points and the normal vector are calculated. Subsequently, the computed angles are used to rotate the face images. Consequently, certain characteristics of these qualities are identified in order to locate the human face. The face exhibits prominent characteristics like the upper part of the chin, lips, eyes, nose, and eyebrows. Thus, in order to locate the human face, distinct features of these characteristics are recognized. The FLL efficiently manages pose variation. Face alignment is constructing a triangle by connecting the eye positions of facial landmarks to the image's normal, and then determining the angles between the eye points and the normal. Afterwards, the calculated angles are used to rotate the facial photographs.

3.1 Dataset description

Dataset Preparation; the code acquires the LFW dataset [G. B. Huang et al., 2008] and readies it for training a facerecognition model illustrating how to manage datasets for machine learning purposes. A collection of information and pictures of faces pasted in the wild (LFW) were used for facial detection and identification investigations. A database of face photos called Wild Classified Faces (LFW) was created to research the issue of facial identification in an unconstrained setting. The collection includes 5,000 image of 80 well-known individuals that were gathered from the Internet and recognized by the sequential CNN detector. Overall It consists of 13,233 photos of 5749 different persons, each with a wide range of image conditions. The inherent diversity and complexity of this collection enables the application of learnt models to novel, unseen image, such as those obtained from the internet. This

database is also compatible with the DAR process. Fig. 1. Illustrate some samples of the datasets that used in our experiment.



Fig. 1. Some samples of LFW dataset images.

3.2 Convolutional neural network for classification

Create a predictive model using the training data to determine the names linked to images.

The training set should include many photos of certain persons to provide the model with the ability to handle the same face in various contexts.

Evaluate the model's performance on novel photos and make predictions about the identification.

The test set should include both familiar persons from the training set (albeit not the exact same photographs) and unfamiliar individuals to evaluate the algorithm's accuracy.

We used a fine tune for Convolutional Neural Network (CNN) layers on the LFW dataset. Prior to constructing any models, it is crucial to comprehend the composition and objective of a Convolutional Neural Network (CNN). There is an abundance of material accessible, however, the primary concepts are succinctly summarized.

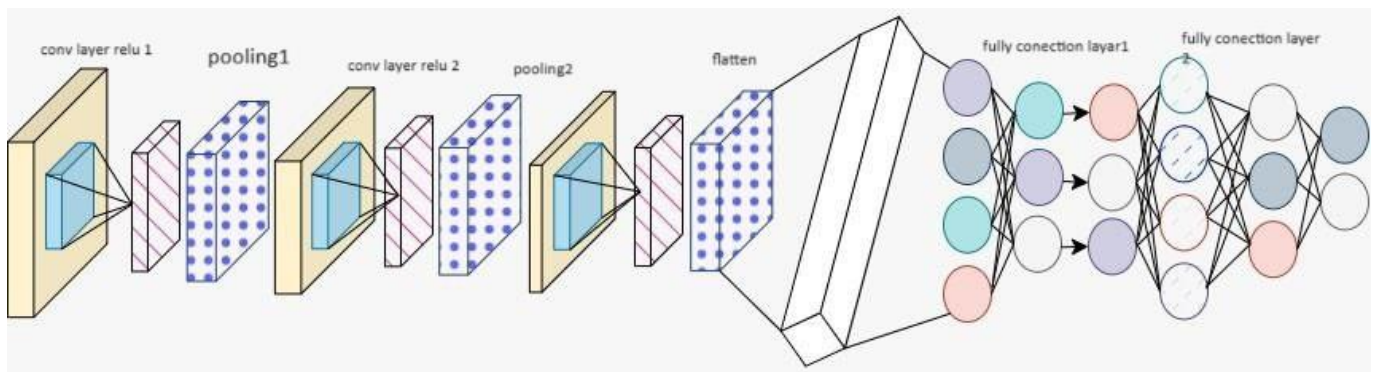


Fig.2. Convolutional neural network layers of CNN-Mesh model

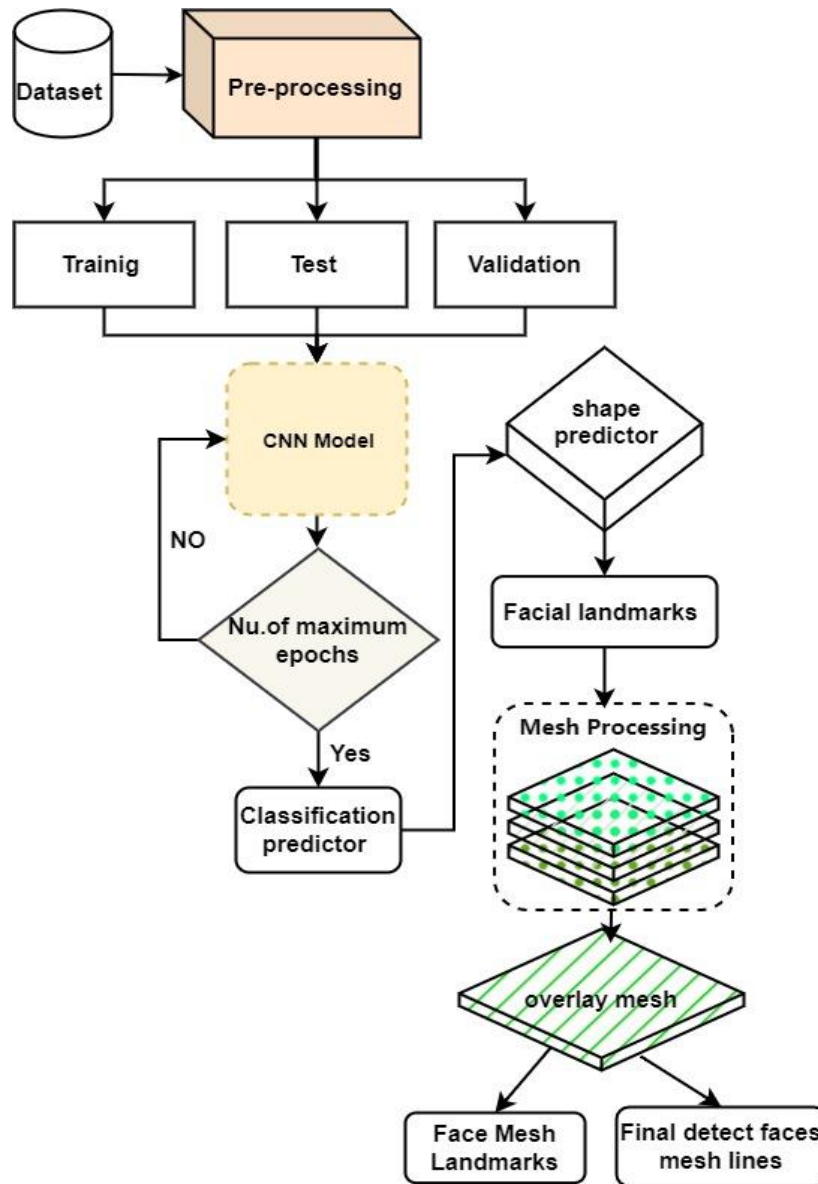


Fig. 3. flowchart of CNN-Mesh model

Fig. 2. Illustrate our proposed convolutional neural network which is a neural network typically used for the purpose of image processing. The main idea is that by scanning images regions, one may get convolved representations that have a smaller dimensionality and yet maintain the links between pixel areas. The convolved model consists of four essential step, first, Convolution used to apply a convolution operation to the images using a feature detector in order to generate a feature map, such as edge detection.

Secondly, Non-Linearity that used to introduce non-linarites in the data by using a non-linear activation function. Third, Pooling is a technique used to decrease the dimension of feature maps by calculating the average of convolved values inside a certain area. Finally, establish interconnections between the layers for arrange all the previous components of the model into a network that is completely interconnected, and enabling it to provide predictions. Fig. 3 shows the flow chart and steps of our proposed CNN model that are discussed below:

1. Model Creation and Training; It. Trains a Convolutional Neural Network (CNN) model using Keras and Tensor Flow for face recognition. The model design incorporates layers pooling layersbatch normalization and dense layers showcasing recommended practices for designing CNNs.

2. Performance Assessment; The code assesses the trained models performance through metrics like accuracy, classification report, confusion matrix and ROC curve to offer insights into the models efficacy.
3. Facial Landmark Detection; It showcases how to identify landmarks using Dlib and MediaPipe demonstrating techniques for extracting and analyzing features.
4. Visualization; The code presents visualizations such as training and validation loss and accuracy plots, pixel intensity distribution as a collection of images with predicted labels, alongside true labels to deepen understanding of model performance and data characteristics.

3.3 Face Occlusion and Non-Uniform Illumination Analysis

The LFW dataset is a well-recognized and respected benchmark dataset in the area of computer vision. The images in the LFW collection were acquired from several sources, including the internet, leading to disparities in quality, lighting, and location. Therefore, the LFW dataset poses a challenging test for face recognition algorithms. The LFW dataset has been extensively used in research, including areas such as face detection, recognition, verification, and identification. The LFW dataset has been used for the advancement and assessment of several algorithms, and it is extensively adopted as a standardized benchmark for evaluating the efficacy of various strategies.

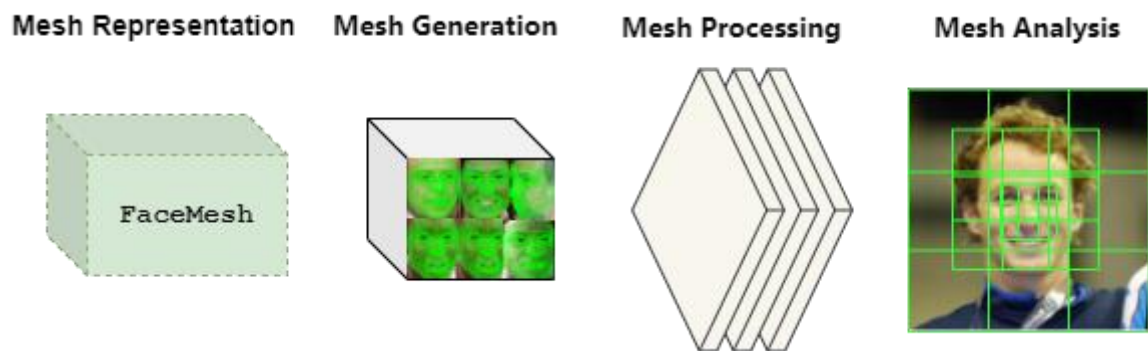


Fig 4. Second stage of proposed model of mesh network

The second objective is to examine different Convolutional Neural Network (CNN) structures for the purpose of identifying and verifying faces. This includes studying face alignment and metric learning, while using the unique dataset for training. Fig. 4. shows the steps of the second stage of our proposed algorithm. Accordingly, Mesh representation shows at FaceMesh block that feed the features to mesh generation block. Last two blocks of the mesh network are mesh processing and mesh analysis for generate and predict the final mask covered each face that show the final output results of prediction.

As follows is the procedure that the CNN-Mesh system follows in order to function: In the first place, the tracking thread will extract feature points after reading the RGB picture and the depth image. Meanwhile, the target thread will follow suit. The detection thread is responsible for identifying information about items in the environment and selecting things that have the ability to move. These objects are referred to as prospective moving objects. In the next step, the dynamic object judgment module receives the bounding boxes of these objects as well as the feature points that were retrieved by the tracking thread.

The CNN-Mesh module is used to update the probability of feature points and evaluate the motion state of objects. This is accomplished via the utilization of a dynamic probability propagation model that is founded on target detection and the coupling of numerous geometric constraints. After that, we send the bounding boxes that we have determined to be dynamic objects to the outlier rejection module. This module makes use of an object foreground image segmentation algorithm that is based on depth image histogram statistics in order to extract information about the contours of dynamic objects and reject feature points that are located within the contour.

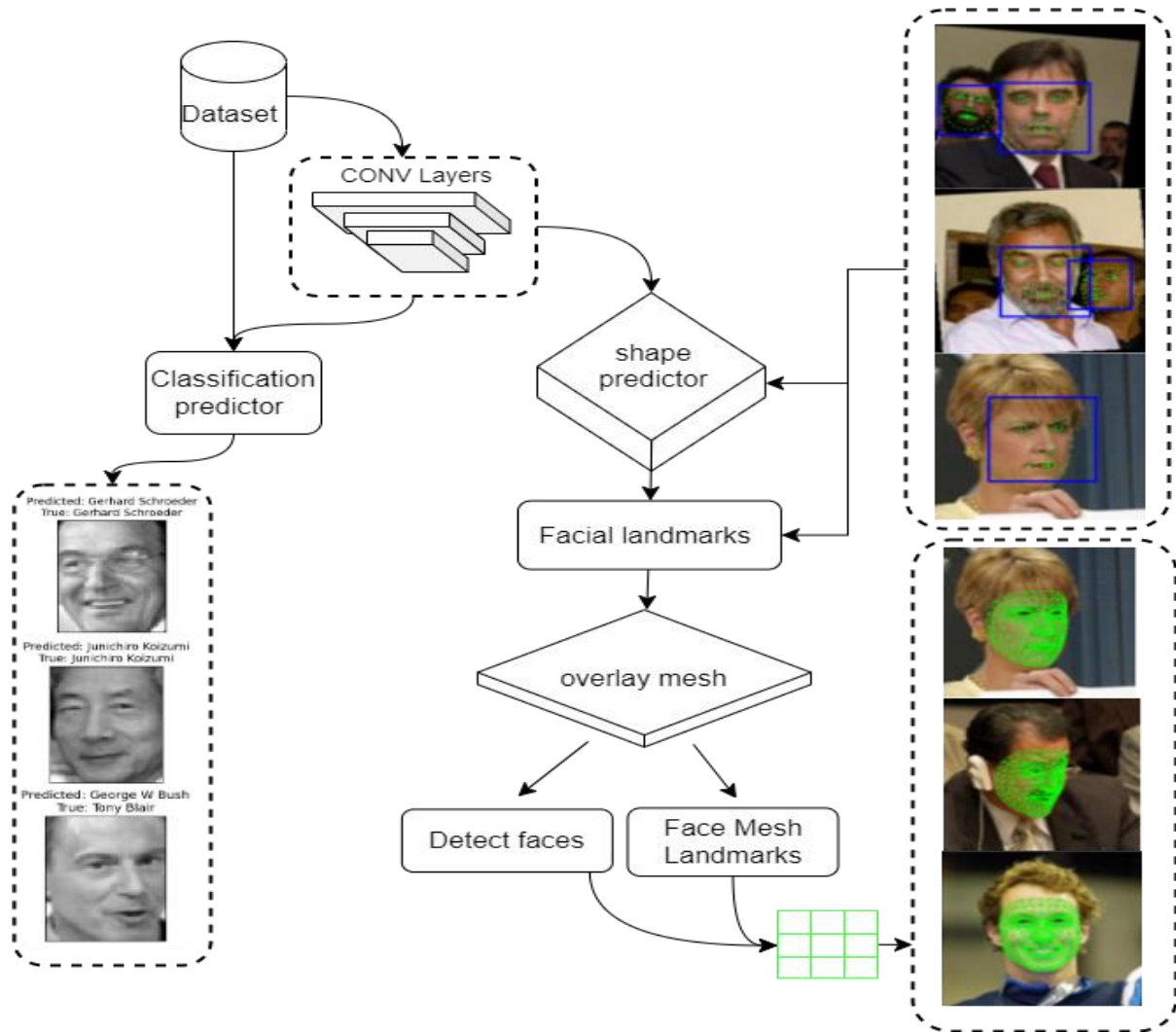


Fig. 5. Second stage of proposed model of mesh network CNN-Mesh

The architecture of our model is proposed in Fig.5. the figure is shows the datasets collection feed to the convs layers for showing the classification predictor depend on each class of persons. On the other hand, the output of CNN model feeds to shape predictor to process features for facial landmarks prediction with blue box.

3.4 Face landmarks

In the computer vision task of face landmark detection, the important features of the human face are identified, as illustrated below. In deep learning, facial landmark detection is the act of recognizing particular points on a face, like the mouth, nose, and eye corners. Numerous applications, such as facial recognition, facial expression analysis, and artwork analysis as shown in figure 6. Although deep learning-based techniques have demonstrated impressive results in the identification of face landmarks, they sometimes call for intricate models with a lot of parameters, which increases computational complexity and execution time.

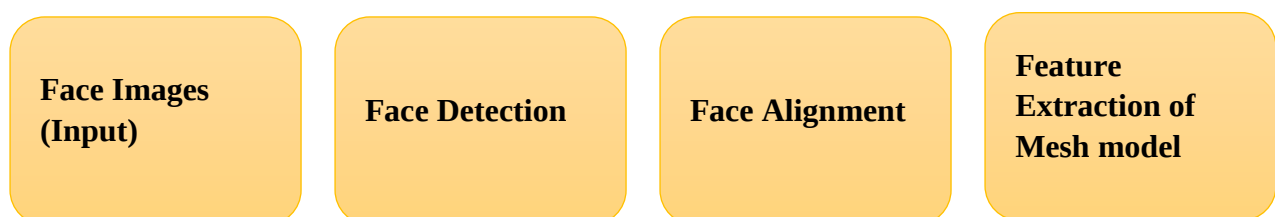


Fig. 6. the preprocessing of face textures

4. Experiment Results

the suggested model performed effectively while recognizing faces images having several difficulties. Fig. 7 displays the validation prediction of the gathered images dataset for 250×250 pixels.

The classification part of our model enhanced by using fine tune parameters with batch size 32. the faster prediction is used gray scale images for increase the accuracy. Fig. 7. Precisely depicts the true person related matching with the prediction name at each class. This robust prediction is depending on feature extracted that collects automatically from the convolved model.

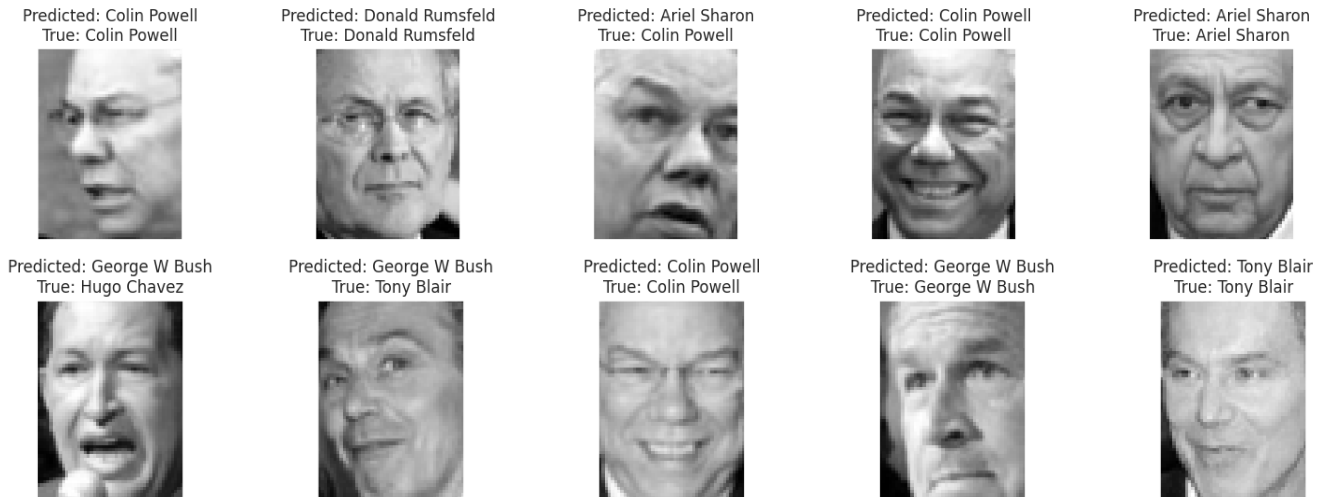


Fig. 7. The prediction results of CNN model

Prediction stage when a trained model is used to figure out some one's identity from an image its referred to as prediction, in the field of face recognition. The classifier is employed to guess the identities of faces, in data having been trained on different classes. This involves passing the input image through the trained model, which then provides a distribution of identities or classify. The highest probability value is used to identify the category or identity. In terms the model uses features and patterns to determine which specific individual most likely matches the face in the image.



Fig. 8. landmark box prediction for edge face recognition

Fig. 8. Depicts the results of landmark box prediction to describe the edge of face by using bounding box. In the realm of analysis, in the learning domain a tool is employed to measure how well authentication models work. This tool assesses the tradeoff between identifying instances (positive rate) and mistakenly labeling impostors (false positive rate) at various decision points. the results of detection show the power and the best performance of the proposed model that precisely draw the green point to represent the face edge including eyes mouth eyebrow and noise. Plotting these rates on a ROC curve provides insights into the models sensitivity and specificity. The area under this curve (AUC) acts as a benchmark for evaluating the models performance with an AUC indicating discriminatory abilities. This assessment technique is especially beneficial in biometric authentication systems that heavily utilize deep learning algorithms. It enables an evaluation of system performance, by taking into account both impostor scores into the model's effectiveness.

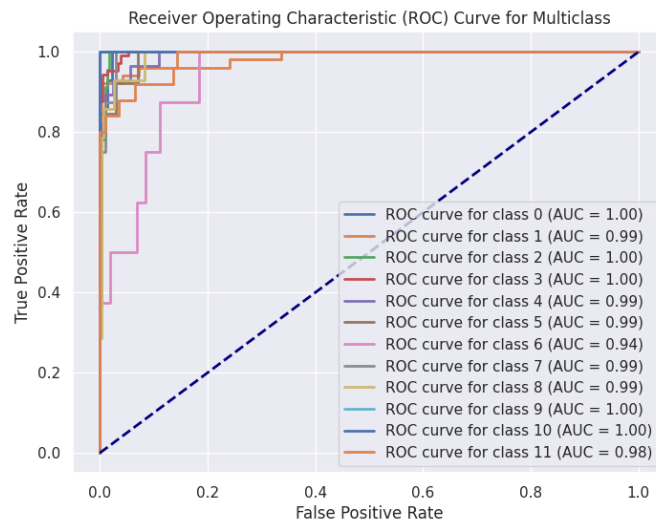


Fig. 9. Receiver Operating Characteristic ROC Curve for multiclass.

The model includes visualizations such as training/validation loss and accuracy plots, pixel intensity distribution, and confusion matrix with predicted and true labels. Also enhancing the predicted of model performance and data characteristics. to evaluate the effectiveness of binary classifiers, we plot the sensitivity (true positive rate) versus the 1-specificity (false positive rate) at different threshold values. Fig. 9. Depicts Receiver Operating Characteristic ROC Curve for multiclass. statistical measure that calculates the average performance across all classes or categories in a dataset. Calculates the average receiver operating characteristic (ROC) curve by combining the contributions of all classes, giving equal weight to each class. Training, validation accuracy and loss in recognition are metrics for assessing the performance of facial recognition models. The training loss indicates the difference between expected and actual outputs during training showing how well the model learns from the data. On the hand, validation loss measures identifications in a separate dataset to estimate how the model generalizes to unseen data. Training accuracy reflects the percentage of classified samples during training while verification accuracy shows the percentage of identified samples in the verification set at 97% in our proposed model. These metrics play a role, in evaluating how effectively the model can learn from training data and accurately recognize faces. Fig. 10 describe the training and validation accuracy and loss for the whole implementation the model with number of epochs 20. Time of training our model is very fast as we used fine tuning the hyper parameters of the model.

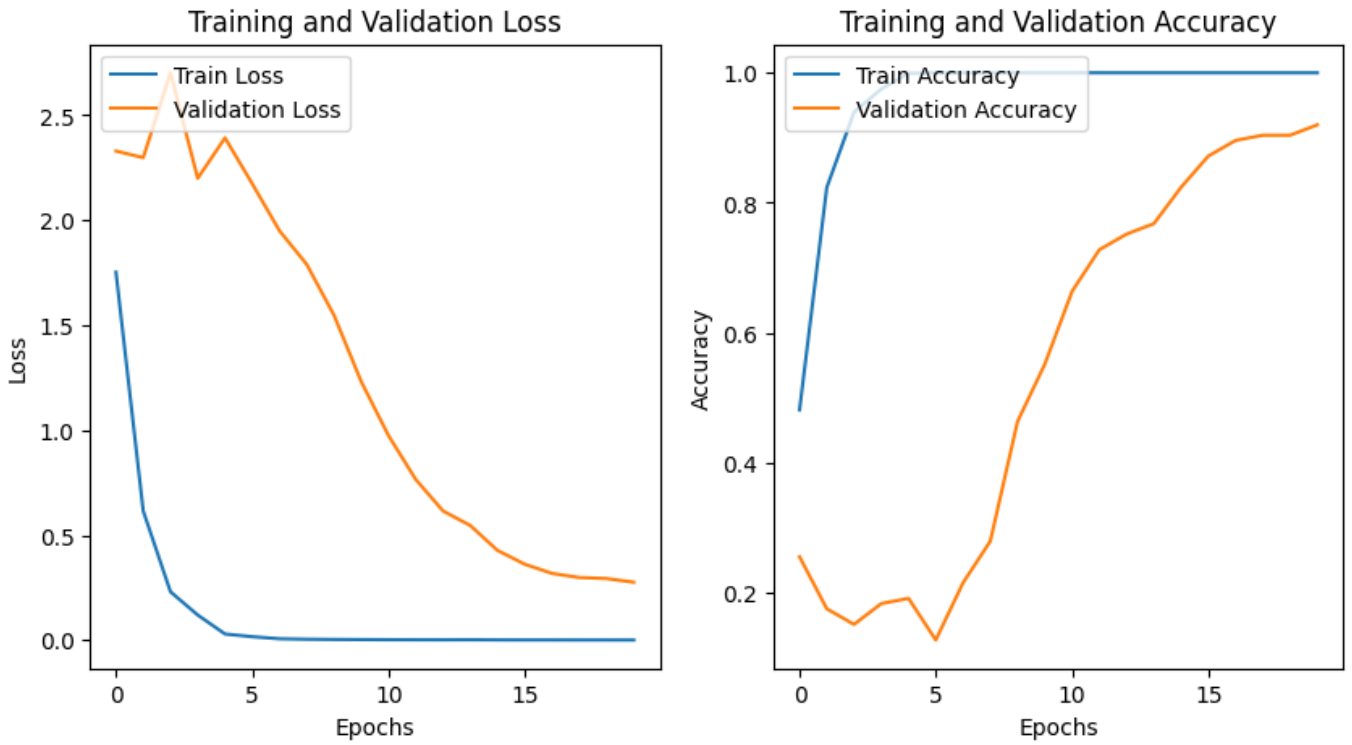


Fig. 10. Train and Validation loss curve over whole epochs (Left). Accuracy for train and valid of CNN-Mesh Model (Right).

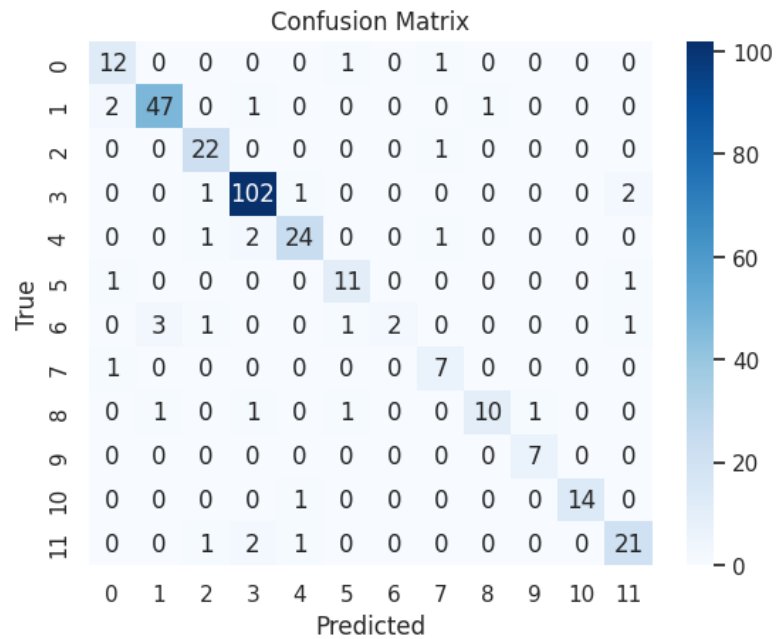


Fig. 11. Confusion Matrix of multiclass true over predicted classification.

Fig.11. shows the prediction multi class with true class to describe the best performance of our proposed CNN-Mesh model.

To enter mesh technique used Mesh Algorithm utilizes the Media Pipe library's Face Mesh module to detect facial landmarks and draw a face mesh on the given image. The specific mathematical equations used in the Media Pipe Face Mesh algorithm are not explicitly provided in the model snippet. Media Pipe Face Mesh employs a deep neural network for facial landmark detection, but the detailed

equations are part of the internal model architecture are not implemented in the code. The pipeline is implemented as a Media Pipe graph that uses a face landmark sub graph from the face landmark module, and renders using a dedicated face renderer sub graph. Facial Landmark Detection It demonstrates how to detect facial landmarks using dlib and Media pipe, showcasing techniques for facial feature extraction and analysis.

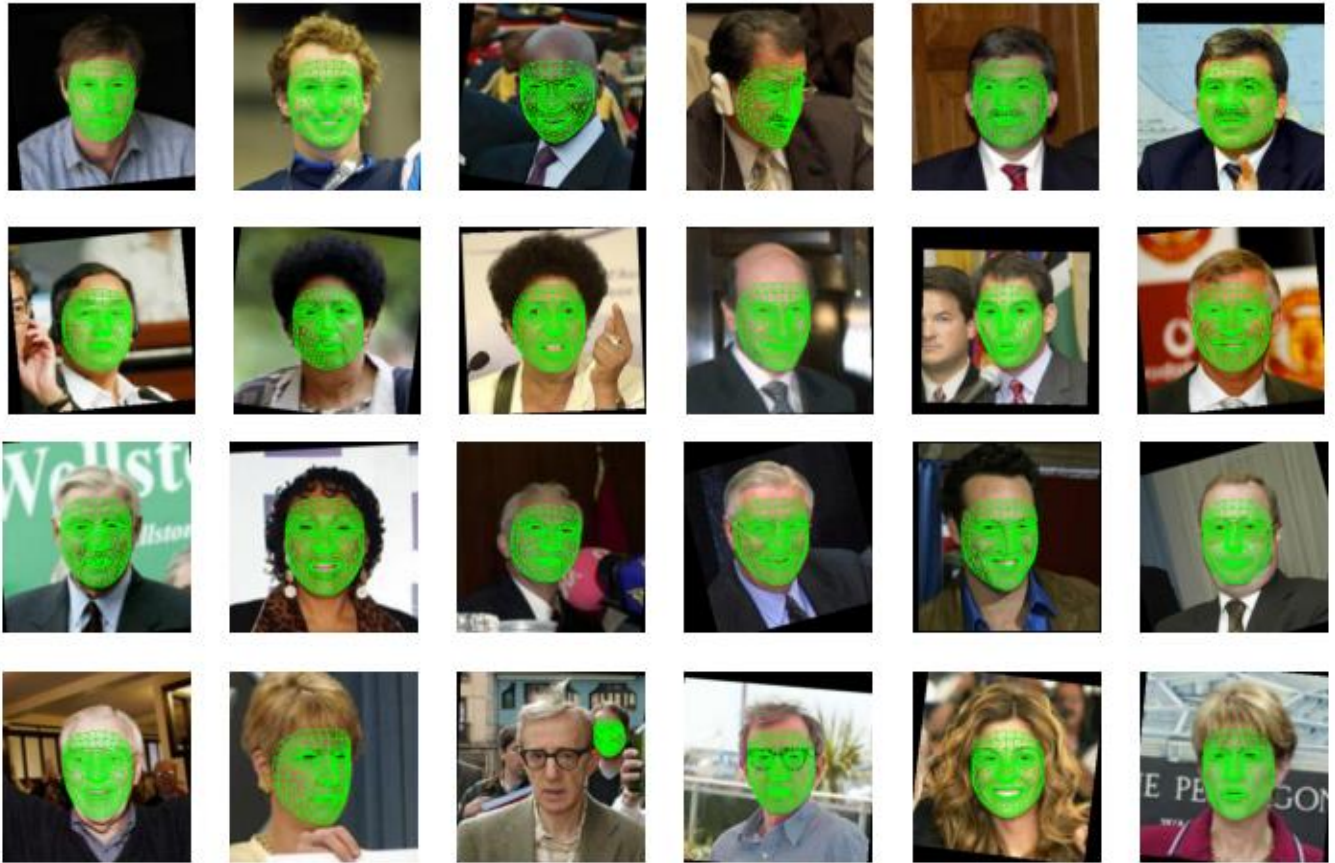


Fig. 12. Mesh network results of prediction the whole face edges.

Fig.12. shows the best prediction of the CNN-Mesh model. the results show the power detection for the whole face and describe the corner edge precisely.

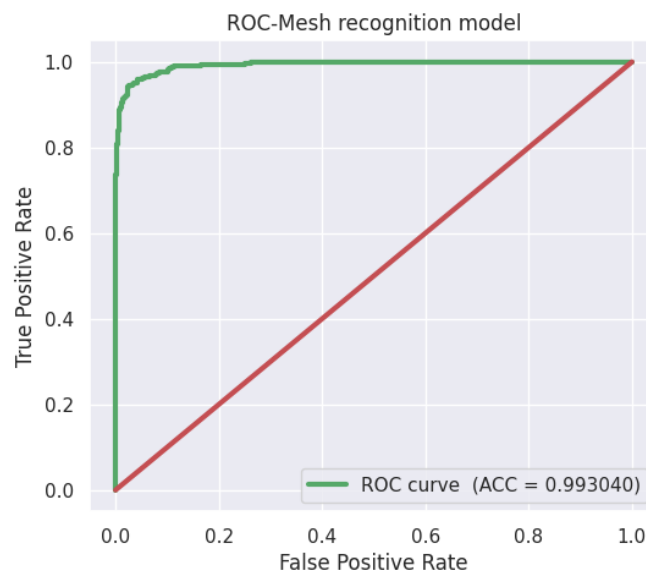


Fig. 13. ROC Curve of mesh model

Fig.13. shows the performance of our CNN-Mesh model with accurate classification and prediction. ROC curve of model accuracy with 0.9930 % .



Figure. 14. depicts the combative results of ResNet model with our proposed network.

Figure 14 displays two line charts that compare the training and validation performance of two distinct neural network models, CNN-Mesh and ResNet, throughout several epochs. The loss value, which reflects how well the model's predictions match the actual labels, is shown along the Y-axis, which is referred to as the loss value. Over a period of epochs, both models demonstrate advances in terms of lowering loss and raising accuracy and precision. Convergence seems to occur more quickly using the CNN-Mesh model, which also exhibits significant gains in training accuracy and loss. Over the course of time, the ResNet model demonstrates more progressive increases, with validation accuracy improving more steadily. Both models eventually perform well, with ResNet displaying a more constant and continuous increase in validation accuracy than the other model

Table 1. the comparative results for face recognition

Authors	Algorithms	Accuracy (%)
DFN-CNN [M. He, J. Zhang, S. Shan, M. Kan, and X. Chen,2020]	Deformable Face Net	93.7%
EIGEN FACES [. Turk and A. Pentland,1991]	near-real-time computer system	53.46%
ADABOOST-LDA [J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, and S. Z. Li,2006]	boost performance of traditional Linear Discriminant Analysis	76.33%
PAL [H. Joo, T. Simon, and Y. Sheikh,2018]	unified deformation model	86.91%
MMDFR [C. Ding and D. Tao,2015]	Multimodal Deep Face Representation	76.31%
Caps Net-FR [M. Ul Haq, M. A. J. Sethi, N. Ben Aoun, A. S. Alluhaidan, and S. Ahmad,2024]	Capsule Networks based FR	97.3%

ResNet Model (ours)	Residual blocks	31.73%
CNN Model (our)	Convolutional network	88.46%
Proposed (CNN-Mesh)	Convolutional and Mesh network	99.30%

Table 1. describe the comparative results with state of art models with our proposed model CNN-Mesh model for face recognition. The performance of proposed model has achieved 99.30 % accuracy.

4.1 Discussion

This study presents a CNN-Mesh model that exhibits substantial improvements in the accuracy of facial recognition (FR), with a classification and detection accuracy rate of 97%. After using the mesh network, this performance further improves to 99%. A multitude of elements contribute to the efficacy of the CNN-Mesh model.

Conventional convolutional neural network (CNN) models suffer with fluctuating illumination conditions, which may greatly affect the process of extracting features. To tackle this issue, the CNN-Mesh model integrates a CNN algorithm with a mesh network, therefore augmenting the model's capacity to effectively capture intricate face characteristics. The mesh network facilitates the construction of precise face markers that maintain consistency even under varying illumination conditions, hence enhancing the resilience of the model.

Facial expressions provide ubiquitous obstacles in the field of face identification. The CNN-Mesh model utilizes a mesh network to accurately delineate the boundaries and shapes of faces, enabling it to accurately map and forecast facial landmarks, even in situations when some areas of the face are hidden or when facial emotions exhibit significant variations. The capacity to accurately and reliably record comprehensive face landmarks guarantees the model's accuracy in such circumstances.

On the other hand, applied supervised learning enables the model to acquire knowledge from annotated data, hence improving its capacity to identify patterns and characteristics linked to particular persons. The inclusion of 68 Landmark markers enhances the model's accuracy by guaranteeing faithful capture of small variations among individual faces. This exceptional degree of accuracy is especially beneficial in situations that need utmost confidentiality.

5. Conclusion

In this study, we suggested a unique CNN-Mesh model for face recognition (FR) that combines Convolutional Neural Networks (CNNs) with a mesh network approach to increase accuracy in demanding settings such as shifting illumination, facial expressions, and partial occlusions. This model was developed in order to improve the accuracy of face identification in challenging conditions. The model that has been suggested functions in two stages: the first stage makes use of convolutional neural networks (CNNs) for the purpose of feature extraction in order to solve the classification issue. The second stage makes use of a mesh network that makes use of these features in order to forecast and identify face landmarks. This dual-stage strategy allows the model to obtain a high classification accuracy of 97%, which further climbs to 99% with the integration of the mesh network, representing a major progress in the area of face recognition. Moreover, the implementation of the mesh network further boosts the accuracy of the classification. The CNN-Mesh model is extremely resistant against fluctuations in illumination and face emotions, which are major obstacles in real-world FR applications. This is because the CNN-Mesh model is able to record specific facial landmarks. Because of its resilience, the model is a competitive alternative for applications that need high levels of security and dependability. Some examples of these applications include biometrics, authentication systems, surveillance, and law enforcement authorities. The results of our CNN-Mesh

algorithm achieved a 97% accuracy rate for classification and detecting faces, while the accuracy of the model is increased after mesh network to 99%. On the other hand, the research also outlines a number of drawbacks that are linked with the model that was developed. Its computational complexity, which may make its deployment in real-time or resource-constrained situations more difficult, and its sensitivity to training data quality and variety, which may have an impact on its capacity to generalize to other datasets, are two of the factors that contribute to these limitations. Due to the fact that the model relies on 68 landmark markers, its efficiency may be hindered in situations in which these landmarks are not well defined or when faces are significantly obscured. Furthermore, it is necessary to carefully evaluate the risk for overfitting to certain face traits, as well as ethical problems relating to privacy and fairness.

It is recommended that future study concentrate on refining the CNN-Mesh model in order to improve its computing efficiency and lessen the dependency on certain landmark markers while doing so. Exploring hybrid architectures, introducing unsupervised learning techniques, and increasing the variety of training datasets are all potential ways to further increase the resilience and applicability of the model. the CNN-Mesh model is a huge step forward in terms of enhancing the accuracy and robustness of face recognition under difficult settings. Despite the fact that there are certain areas that might be improved, the technique that has been provided successfully establishes a solid basis for future developments in facial recognition technology. It strikes a balance between accuracy, efficiency, and privacy concerns.

Reference

- [1] S. Kadhem, Z. Ali, and A. Suhad Malallah Zuhair Hussein Ali, "Multi-Document Summarization using Fuzzy Logic and Firefly Algorithm Multi-Document Summarization using Fuzzy Logic and Firefly Algorithm . [Online]. Available: 2018 <https://www.researchgate.net/publication/330006097>
- [2] A. A. Maryoosh and S. Pashazadeh, "Leukemia Detection Using Machine Learning Algorithms: Current Trends and Future Directions (Literature Survey)," 2024. [Online]. Available: <https://mjpas.uomustansiriyah.edu.iq/index.php/mjpas>
- [3] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," *Electronics (Switzerland)*, vol. 9, no. 8. MDPI AG, pp. 1–53, Aug. 01, 2020. doi: 10.3390/electronics9081188.
- [4] K. Zhang, Y.-L. Chang, and W. Hsu, "Deep Disguised Faces Recognition."
- [5] C. R. Vishwanatha, V. Asha, B. Saju, N. Suma, T. R. Mrudhula Reddy, and K. M. Sumanth, "Face Recognition and Identification Using Deep Learning," in *2023 3rd International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies, ICAECT 2023*, Institute of Electrical and Electronics Engineers Inc., 2023. doi: 10.1109/ICAECT57570.2023.10118154.
- [6] S. Shakya, G. Papakostas, and K. A. Kamel Editors, "Lecture Notes on Data Engineering and Communications Technologies 166 Mobile Computing and Sustainable Informatics Proceedings of ICMCSI 2023."
- [7] L. Zhou and W. Yu, "Improved Convolutional Neural Image Recognition Algorithm based on LeNet-5," *Journal of Computer Networks and Communications*, vol. 2022, 2022, doi: 10.1155/2022/1636203.
- [8] N. Bukhari, S. Hussain, M. Ayoub, Y. Yu, and A. Khan, "Pakistan Journal of Engineering and Technology, PakJET A Deep Learning-based Framework for Emotion Recognition using Facial Expression".
- [9] S. Anwarul, T. Choudhury, and S. Dahiya, "A novel hybrid ensemble convolutional neural network for face recognition by optimizing hyperparameters," *Nonlinear Engineering*, vol. 12, no. 1, Jan. 2023, doi: [10].1515/nleng-2022-0290.

- [11] H. M. Shahzad, S. M. Bhatti, A. Jaffar, S. Akram, M. Alhajlah, and A. Mahmood, "Hybrid Facial Emotion Recognition Using CNN-Based Features," *Applied Sciences (Switzerland)*, vol. 13, no. 9, May 2023, doi: 10.3390/app13095572.
- [12] D. S. Ramesh, S. Heijnen, O. Hekster, L. Spreeuwiers, and F. de Wit, "Facial recognition as a tool to identify Roman emperors: towards a new methodology," *Humanities and Social Sciences Communications*, vol. 9, no. 1, Dec. 2022, doi: 10.1057/s41599-022-01090-y.
- [13] Z. Y. Huang *et al.*, "A study on computer vision for facial emotion recognition," *Scientific Reports*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-35446-4.
- [14] A. J. Russ, M. Sauerland, C. E. Lee, and M. Bindemann, "Individual differences in eyewitness accuracy across multiple lineups of faces," *Cognitive Research: Principles and Implications*, vol. 3, no. 1, Dec. 2018, doi: 10.1186/s41235-018-0121-8.
- [15] M. He, J. Zhang, S. Shan, M. Kan, and X. Chen, "Deformable face net for pose invariant face recognition," *Pattern Recognition*, vol. 100, Apr. 2020, doi: 10.1016/j.patcog.2019.107113.
- [16] Y. Taigman, M. Y. Marc', A. Ranzato, and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification."
- [17] M. Joo Er, S. Wu, J. Lu, S. Member, and H. Lye Toh, "Face Recognition With Radial Basis Function (RBF) Neural Networks," 2002.
- [18] L. Lu, Y. Mao, L. Wenyin, and H.-J. Zhang, "AUDIO RESTORATION BY CONSTRAINED AUDIO TEXTURE SYNTHESIS."
- [19] G. B. Huang, M. Mattar, T. Berg, E. Learned-Miller, and E. Learned-Miller, "Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments." [Online]. Available: <http://vis-www.cs.umass.edu/lfw/>.
- [20] M. He, J. Zhang, S. Shan, M. Kan, and X. Chen, "Deformable face net for pose invariant face recognition," *Pattern Recognition*, vol. 100, Apr. 2020, doi: 10.1016/j.patcog.2019.107113.
- [21] M. Turk and A. Pentland, "Eigenfaces for Recognition." [Online]. Available: <http://direct.mit.edu/jocn/article-pdf/3/1/71/1932018/jocn.1991.3.1.71.pdf>
- [22] J. Lu, K. N. Plataniotis, A. N. Venetsanopoulos, and S. Z. Li, "Ensemble-based discriminant learning with boosting for face recognition," *IEEE Transactions on Neural Networks*, vol. 17, no. 1, pp. 166–178, 2006, doi: 10.1109/TNN.2005.860853.
- [23] H. Joo, T. Simon, and Y. Sheikh, "Total Capture: A 3D Deformation Model for Tracking Faces, Hands, and Bodies." [Online]. Available: <http://www.cs.cmu.edu/>
- [24] C. Ding and D. Tao, "Robust Face Recognition via Multimodal Deep Face Representation," Sep. 2015, doi: 10.1109/TMM.2015.2477042.
- [25] M. U. Haq, M. A. J. Sethi, N. ben Aoun, A. S. Alluhaidan, S. Ahmad, and Z. Farid, "CapsNet-FR: Capsule Networks for Improved Recognition of Facial Features," *Computers, Materials and Continua*, vol. 79, no. 2, pp. 2169–2186, 2024, doi: 10.32604/cmc.2024.049645.