International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8322
***Article Received:** 12 July 2023 **Revised:** 31 August 2023 **Accepted:** 20 September 2023*

RESDEN: A Novel Deep Unified Model for Face Recognition System

Kavita[1], Rajender Singh Chhillar[2]
[1]Research Scholar, Department of Computer Science and Applications
Maharshi Dayanand University, Rohtak-124001, Haryana, India
kavitavce@gmail.com
[2]Professor, Department of Computer Science and Applications
Maharshi Dayanand University, Rohtak-124001, Haryana, India
Chhillar02@gmail.com

Abstract—The Face Recognition technology plays a significant role in the field of Computer Vision in contemporary times. The research article is centered on a Facial attendance system that utilizes a deep learning technique to recognize face photos. To execute face identification and classification via the use of deep learning processes, many Convolutional Neural Network (CNN) models are taken into account. Previous studies have mostly focused on either the ResNet or DenseNet-based convolutional neural network model. The present research utilizes the merging of ResNet and DenseNet to propose a hybrid model. The proposed work is expected to provide enhanced efficiency and accuracy. In the training and testing stages of the simulation, considerations are made for both binary and category classifications. The current research focuses on the use of the LFW dataset. The pictures undergo an initial step of the noise reduction process. The evaluation of picture quality is conducted by taking into account metrics such as Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM). After the proposed model has undergone training, it generates photographs of superior quality. Finally, the proposed system incorporates the RESDEN framework, which integrates DenseNet with a noise reduction technique, a segmentation mechanism, and a CNN based on ResNet. A comparative analysis has been conducted to evaluate the accuracy of several filtered picture sets across different convolutional neural network (CNN) models. The simulation results indicate that the suggested model exhibited a good level of performance and accuracy.

Keywords- Face Recognition, Segmentation, Deep Learning, LFW dataset, CNN.

I. INTRODUCTION

Computer algorithms have been used by facial recognition (FR) systems to identify unique characteristics of an individual's face [1]. The feature presented on faces in an FR database is compared to the mathematical representation of such traits, such as the distance between the eyes or the shape of the chin [2]. Facial recognition is commonly used with different biometric technologies while analyzing identification documents to avoid ID fraud and identity theft [3, 4].

In the current era, Face recognition is also applicable to biometric applications that have been used to take attendance in educational and professional institutions. Face recognition system development is a difficult and important endeavor for security issues. Multiple face recognition systems use various deep learning algorithms. Existing research focused on the CNN model for face detection and classification. One of the key issues with traditional mechanisms is that little work has been done on the quality of facial photographs. Before the training operation, the picture quality must be improved [5].

There have been several limitations in work related to facial recognition. Those systems ignored the performance and accuracy factors. Considering the conventional CNN model, it is observed that there is a need to do more work on overfitting and underfitting problems. Moreover, there is an issue of image quality during preprocessing. To improve the quality there is a need to integrate noise removal techniques.

A. CNN Model for Face Recognition

By using a smart CNN-based model with an enhanced facial image dataset, facial traits can be extracted with some degree of success, leading to improved identification accuracy [6]. One useful feature of CNNs is that they can create a representation of a two-dimensional image. The model can learn the location and size of faces in a picture as a result. After receiving training, CNN can identify a person's face in a photograph. One of CNN's greatest strengths is its ability to detect important features without human intervention. CNNs are a kind of ANN that was designed to manage pixel input in image identification and processing. In the case of CNN, the input layer collects image data and the hidden layer performs internal work. The output layer produces output [7]. There have been different CNN models with different features.

The types of CNN models are:

1) VGGNET:

A 16-layer CNN (VGG-16) may be imported and has been pre-trained on the ImageNet database's astronomical number of photos [8]. The acronym VGG refers to the many layers of the common deep Convolutional Neural Network (CNN) architecture used by the Visual Geometry Group. The "deep" in VGG-16 and VGG-19 refers to the number of convolutional layers, respectively 16 and 19.

2) AlexNet:

AlexNet is a state-of-the-art framework for any object identification task and has the potential for widespread use in the AI computer vision field. This is due to AlexNet's origins as a tool for addressing issues in computer vision. It's possible that soon, AlexNet will replace CNNs as the go-to solution for picture jobs [9].

3) Inception:

Inception V3 is a CNN model used for image processing and object identification. It was originally designed as a module. A convolutional neural network (CNN) Inception Module is a piece of its picture model that seeks to approximate the best local sparse structure [10]. Simply said, it enables the usage of several filter sizes inside a single picture block, which may then be concatenated and sent to the subsequent layer.

4) GoogleNet:

Google's Inception Convolutional Neural Network has undergone two significant revisions [11]. GoogleNet is a 22-layer convolutional neural network developed by Google. Pre-trained models of the network created with either the ImageNet [12] or Places365 datasets may be loaded. The network can divide photos into 1000 different categories, such as "keyboard," "mouse," "pencil," "animal," and so on.

The lack of sophistication in this form is reflected in the fact that even the 55 convolutional layers are computationally fairly costly, meaning it takes a long time and a lot of processing power to complete. An operation like maxpool greatly slows down a Convolutional neural network [13].

5) ResNet:

ResNet is a kind of traditional neural network that serves as the foundation for several computer vision applications. ResNet50 has 50 such filters used in layers. ResNet's most famous contribution was enabling the efficient training of very deep neural networks, including the use of skip connections and more than 150 layers in this case to enhance accuracy [8, 14].

6) DenseNet's levels:

The first layer of DenseNet is composed of a convolution and pooling operation [15, 16]. Next comes a dense block, then a transition layer, then another dense block, another transition layer, and lastly a dense block, then a classification layer. It's common knowledge that deeper networks perform better and are simpler to train than their shallower and broader counterparts [17].

B. Noise Removal Filters

The conventional system did not consider a noise removal mechanism. Therefore, the accuracy during identification by machine learning gets reduced [18]. Moreover, need to integrate noise removal filters that are capable of eliminating noise from images in the preprocessing phase. Filtering has been used to minimize visual noise. Filters have various uses in the area of Image Processing [19], including noise reduction, interpolation, and resampling [20]. This is a crucial component of any image processing system. Image processing present research considers median, Gaussian, blur, and bilateral filters and analyzes that the Gaussian filter provides better results [21]. Due to its features such as the Gaussian filter Gaussian kernels offer less importance to pixels that are farther out from the center of the window. Its primary function is to reduce harshness or noise in images. The Gaussian filter may be used independently to reduce contrast and smooth out edges [22]. The Gaussian function, for instance, is very sensitive to changes in its standard deviation. Gaussian filtering may eliminate noise and extract fine features from an image [23].

C. Segmentation using HAAR Cascade

To calculate and match features, Haar features move across the image in a window-sized fashion. The Haar cascade is an effective classifier. Data points are sorted into two categories: those that include a detected face (positive) and those that do not (negative) [24]. Real-time performance of Haar cascades is possible because of their speed. It's an algorithm that can identify human faces in pictures, no matter how little or far away they may be. Trainable items for a haar-cascade detector include human faces, vehicles, buildings, fruits, and more. A face is immediately recognized by it [25]. Concurrently, CNN takes advantage of the convolutional process by propagating a convolution (filter) kernel of a predetermined size from the product of the image's multiplication with the filter to the next picture [26].

D. Paper Organization

Section 1 has introduced conventional models that are based on CNN and used for facial detection and classification.

Section 2 considers existing research related to facial identification with their methodology and results.

Section 3 presents the proposed research methodology along with the proposed system architecture.

Section 4 illustrates the simulation of noise removal, techniques, and measures of accuracy for a hybrid model.

Section 5 concludes the work with future aspects of research.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8322
***Article Received:** 12 July 2023 **Revised:** 31 August 2023 **Accepted:** 20 September 2023*

II. LITERATURE REVIEW

There have been several kinds of research in the area of facial recognition where face identification and classification mechanisms are applied.

K. Kavita et al. (2022) introduced machine learning recognized for human faces. Tools like Matlab and Python were examined with machine learning and deep learning methods. They reviewed trending techniques for face detection, face identification, facial expression, and age estimation. The article also focused on the present and future direction of research in facial recognition systems [1].

P. T. Waghmare et al. (2021) analyzed the literature on deep learning-based, general-purposed face recognition systems. This article provided a comprehensive bibliometric analysis of previous research on deep learning-based face recognition systems. Data analysis was performed using the Scopus database, and the retrieved information was visualized using some programs [2].

F. Kong et al. (2019) refined LBP combined with the deep CNN-learned abstract aspects of facial expression. In a head-to-head test, the model outperformed the competition by 91.28% [3].

S. M. Bah et al. (2020) introduced a novel method to enhance the performance of face recognition systems by combining an LBP algorithm with cut-edge image-processed methods such as contrast adjustment, bilateral filtering, histogram equalization, and picture blending [4].

Fahad P (2017) proposed research on face recognition. They presented a real-time system that has been used for intelligent decision-making during surveillance [5].

Kirtiraj Kadam (2017) did research work on the attendance Monitoring System. Their research focused on Image Processing. The classification process has been made using a Machine Learning mechanism [6].

R. Zhang (2018) worked on automatic segmentation. Their research focused on acute ischemic stroke from DWI. The author made use of 3-D fully convolutional DenseNets [7].

The unique CNN-based technique suggested by Lu and Jiang (2018) is a deep-coupled ResNet model for low-resolution face recognition. The suggested DCR model consistently outperformed the state-of-the-art on both the LFW and SCface datasets [8].

M. Z. Khan, et al. (2019) focused on a deep unified model for face recognition based on CNN & edge computing. Performance evaluated on a publicly accessible dataset. The algorithm could distinguish 30 students' faces from 40 from a single photograph and found that the method was 97.9% accurate [9].

Zhu, et al. (2019) considered scalable sample learning for ID vs. feature-based face recognition. For IvS face recognition, presented a deep learning-based LBL approach. Moreover, softmax was added to deep learning to enable its use with massive datasets. LBL was performed on an IvS face dataset that included over 2 million unique individuals [10].

Almabdy, et al. (2019) presented methods for face recognition using deep CNN. The architecture of AlexNet and ResNet-50 achieved the top results in ILSVRC during the previous several years. Improved classification rates were found after rigorous tests on the ORL, GTAV face, and FEI faces datasets. The model's 94%-100% accuracy is better than that of most state-of-the-art models [11].

To extract features from defective facial data, A. Elmahmudi et al. (2019) discovered the capability of CNNs using the VGGF model. Cosine similarity and linear support vector machine classifiers were utilized to evaluate the model's recognition rates. Using the FEI and LFW information, forecasts were made [12].

S. Sharma et al. (2019) presented a framework based on the sequential deep learning approach. The model generated deep features to extract and improve the reconstructed voxels. The characteristics were evaluated using the SVM [13].

The Multipath-DenseNet was used to be deeper and broader, reducing trained time and using fewer parameters by B. Lodhi et al. (2020). Four object recognition datasets, including CIFAR-10, CIFAR-100, SVHN, and ImageNet, were used to test their suggested architecture. The experimental findings demonstrated that Multipath-DenseNet significantly outperformed its predecessor, DenseNet [14].

X. Han conducted a sample image recognition study in 2020. In this study, they used a technique based on ResNet and transfer learning [15].

L. Ke, in their next work (2020), focused on creating a self-constrained 3D DenseNet model. This study focuses on nasopharyngeal cancer and the automated identification and segmentation of this disease. A magnetic resonance image was employed for this study [16].

In 2021, Khawla Alhanaee et al. proposed a deep learning CNN-based face recognition attendance system. The validation accuracy for the three networks utilized ranged from 98.33% to 93.33% to 100%. Many businesses might benefit by implementing the suggested strategy into their attendance and security door access systems [17].

Y. Said, et al. (2021) focused on intuitive facial expression analysis using deep learning and high-resolution photos to decipher human emotions. The FS-CNNs proposed two phases; a patch cropper and a CNN. Faces were detected in the first phase, and the cropped face was then employed in the second phase [18].

Chowdary, et al. (2021) provided human-computer interaction software based on deep learning for recognizing facial expressions of emotion. Resnet50, vgg19, Inception V3, and Mobile Net frameworks were employed for the pre-trained networks. The fully connected layers of the pre-trained ConvNets were removed and replaced with their own, custom-built layers that were sized to match the task's instruction set. Finally, adjusting the weight of the extra layers was all that was required during training [19].

A completely automated, highly effective convolutional architecture for disguise-invariant face detection was created by Junaid Khan et al. (2021) using noise-based data augmentation and deep transfer learning. The proposed method employs the Viola-Jones face detector to locate human faces in a picture, followed by a CNN that had been specifically trained and optimized for DIFR [20] to assign labels to the detected faces.

One of the CNN designs, ResNet-50, was employed for facial recognition in a study by Pratama et al. (2021). Different configurations of ResNet's main parameters were tested to establish their efficacy. The best model, achieved after tested with 22 different setups, had an accuracy of 99% [21].

Sunitha, et al. (2022) reviewed facial picture ethnicity identification and categorization powered by intelligent deep learning. Since PCA can successfully overcome the "curse of dimensionality," which happens when the recovered features are highly dimensional, it was utilized in the feature reduction process [22].

Rehmat Ullah et al. (2022) used a variety of techniques to get excellent CNN accuracy. They compared the performance with KNN, decision trees, random forest, and CNN. 40K real-time image set was used with a different environment. Finally, the system recognized faces with an accuracy of more than 90% in minimum computing time [23].

Zhigang Yu et al. (2022) reviewed the classification research for recognized human faces using an enhanced GoogleNet. The network achieved the highest performance, with a recall rate of 0.97 and an accuracy of 0.98 [24].

S. Anwarul et al. (2023) presented a hyper parameter-optimized hybrid ensemble CNN for facial recognition. The model used progressive training dramatically improved recognition accuracy. The model achieved Best-in-class results. Accuracy of 99.35, 91.58, and 95% was achieved by the proposed HE-CNN model on the LFW, cross-pose LFW, and custom-built datasets, respectively [25].

Md. Fazlay Rabbi et al. (2023) introduced autism spectrum disorder. They chose to tackle the issue of autism by developing an image classification system for early diagnosis, using transfer learning with VGG 19, inception V3, and DenseNet 201. They used the picture dataset in conjunction with deep learning methods. They attempted to develop a system that uses facial expressions as a proxy for autism [26].

III. PROPOSED WORK

The fundamental goal of this research is to create a facial attendance system that can analyze digital photos of people's faces using a deep learning mechanism. To execute face recognition and classification using a method based on deep learning, numerous standard CNN models are being examined. ResNet or DenseNet-based CNN models have been studied in conventional research. The proposed hybrid model makes use of an integration of ResNet and DenseNet to provide better accuracy and performance during classification. Throughout the training and testing phases of the model, simulation work considers both binary and categorical classes. The present work considers resizing and image filtering before applying a hybrid CNN model that is based on ResNet and DenseNet. Challenges with traditional processes, such as ignoring the quality of facial photos, are being taken into consideration by research efforts.

A. Proposed Research Methodology

The proposed research effort has mostly concentrated on enhancing the picture quality before the training operation. The current research effort takes into consideration the LFW dataset, in which photos are first processed by a noise reduction mechanism employing a variety of noise removal techniques including median, blur, bilateral, and Gaussian filters, and observe that Gaussian filters gave good results as compared to others. Therefore, a Gaussian filter is used to improve the quality of the image in the preprocessing phase before segmentation. Then, an analysis of the image quality is carried out, taking into account MSE, PSNR, and SSIM. After that, high-quality images are sent to a proposed Hybrid model that is based on the ResNet and DenseNet model i.e. RESDEN model. The primary goal of the work being done using simulations is to evaluate the degree of success achieved by various CNN models using a variety of filtered picture sets.

B. Algorithm for Proposed Work

The novel algorithm of the proposed work describes step-by-step quality evaluation of images. For the process get Image I as input. The preprocessing phase is applied before the simulation is performed with the proposed RESDEN model. The steps followed are:

1. Input image I
2. Resize image I and get RI
3. Apply noise removal mechanisms on NRI
4. Get the Mean squared error ratio by Eq. 1

$$MSE = \frac{1}{mn}\sum_{0}^{m-1}\sum_{0}^{n-1}\|f(i,j)-g(i,j)\|^2 \quad (1)$$

5. Get peak-sensitive noise ratios

The PSNR metric calculates how much noise detracts from a signal and the formula given in Eq. 2

$$PSNR = 20\log_{10}\left(\frac{MAX_f}{\sqrt{MSE}}\right) \quad (2)$$

6. Find the SSIM from the image

The SSIM index is calculated using many different window sizes from a single picture. When determining the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8322
***Article Received:** 12 July 2023 **Revised:** 31 August 2023 **Accepted:** 20 September 2023*

separation between two windows of the same size, we use the formula that can be represented in Eq. 3 as

$$SSIM(x,y) = \frac{(2\mu_x\mu_y+C_1)(2\mu_{xy}+C_2)}{(\mu_x^2+\mu_y^2+C_1)(\sigma_x^2+\sigma_y^2+C_2)} \quad (3)$$

7. Perform segmentation
8. Initialize the proposed RESDEN model for training
9. Split the training and testing dataset for the RESDEN model
10. Perform RESDEN simulations and get accuracy.
11. Make a comparative analysis of accuracy in the case of the RESDEN hybrid model that considers noise-removed images.

Abbreviations:

m: total number of rows of picture pixels,
n: image's column count
f: original image's matrix data
g: matrix information of our deteriorating picture.
MAX_f: highest possible signal value in our reference "excellent" picture.
μ: Pixel sample mean
σ: Variance
C: Variables to stabilize the division with a weak denominator

C. Proposed System Architecture

In the proposed work after getting facial images, preprocessing is made on face sets to get a common dimension of the images. The noise removal section eliminates the noise using Gaussian filters. Then MSE, PSNR, and SSIM are calculated. Classifiers, iterations, and layers are initialized in the RESDEN model to predict the accuracy of a model. The suggested model's person is assessed by a comparison of accuracy measurements.

The Suggested system design is represented in Fig.1. The proposed work focuses on facial attendance systems that employ a deep learning method to analyze captured facial photos. Using a deep learning method for face identification and classification requires taking into account some standard CNN models. Typically, conventional research used a ResNet or DenseNet-based CNN model. The proposed approach has used a State-of-the-Art hybrid model by Integration mechanism on ResNet and DenseNet. Improved efficiency and precision are promised by the proposed work. Model training and testing in simulations consider both binary and categorical classes.

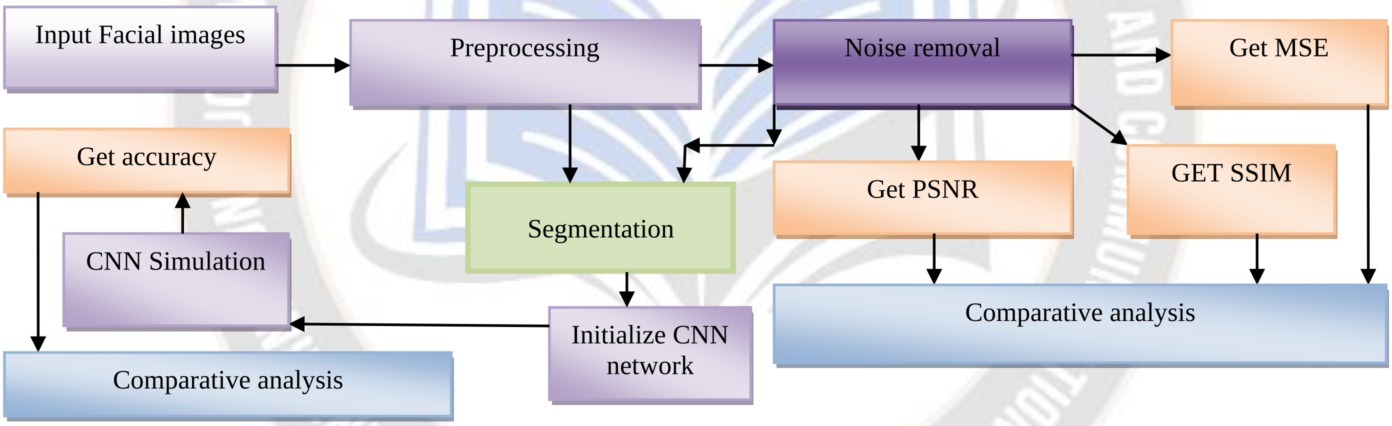


Figure 1. Proposed RESDEN System Architecture

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8322
***Article Received:** 12 July 2023 **Revised:** 31 August 2023 **Accepted:** 20 September 2023*

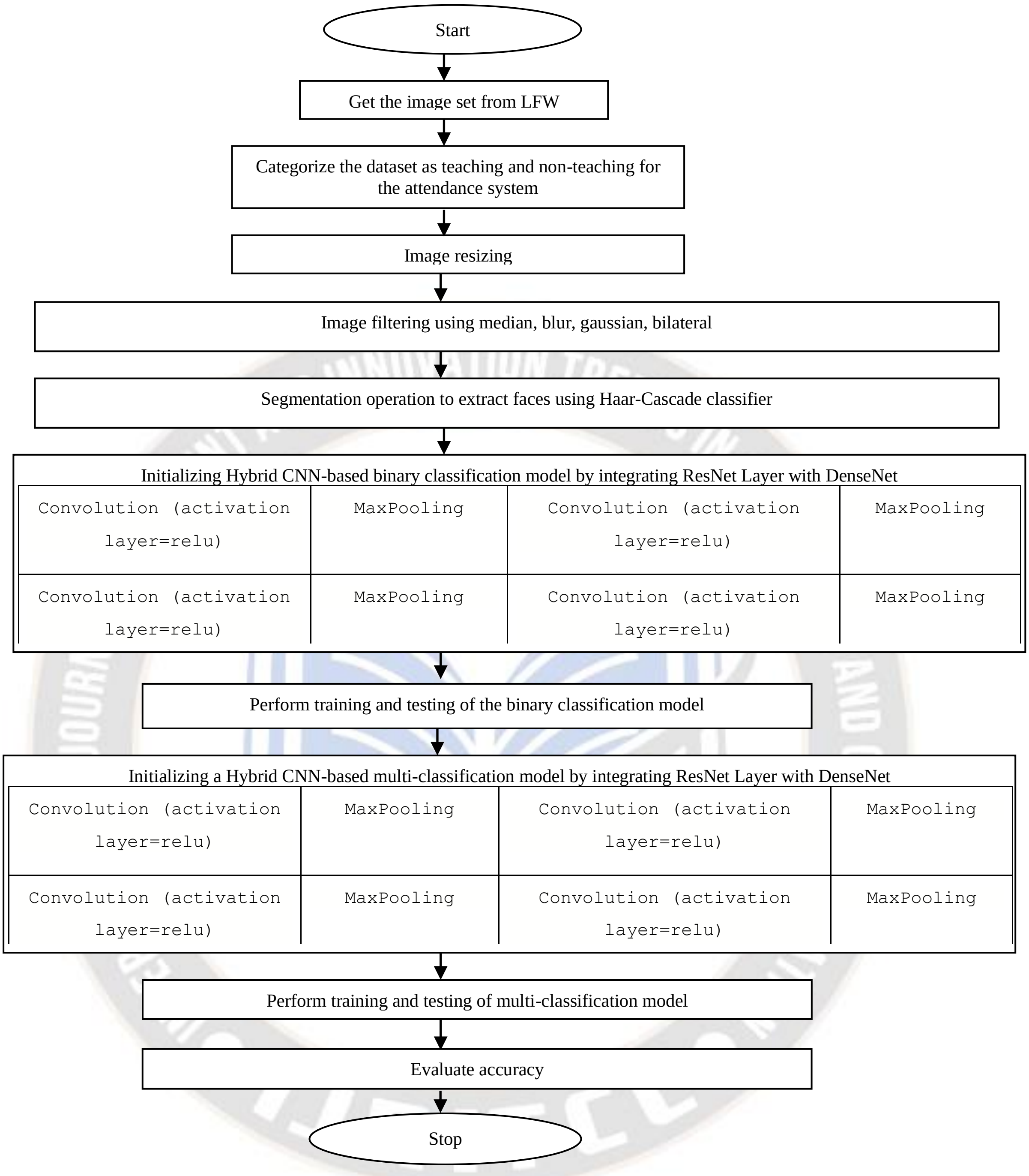


Figure 2. Flow chart of proposed work

The attendance of teaching and non-teaching staff has been used and the proposed system provides a solution with the LFW dataset. Fig. 2 explains the overall process flow of work. The image set of LFW has been considered for preprocessing and these images are compressed to get a dataset for teaching and non-teaching images. This system plays a significant role in image classification for the attendance system. Resizing helps to get consistent size of image and image filters such as median, blur, Gaussian, and bilateral remove noise from images. Haar cascade classifier is used to perform the segmentation to extract faces. Then this data is considered for training by a Hybrid CNN-based binary classification model that has been developed by integrating ResNet Layer with DenseNet. Initially, batch size, learning rate, image set, and epochs are configured for training. Then testing operation is processed to ensure the reliability of the proposed model.

D. Dataset and system requirements

For experiment purposes, a vast collection of face photographs known as the (Labeled Faces in the Wild) LFW

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8322
***Article Received:** 12 July 2023 **Revised:** 31 August 2023 **Accepted:** 20 September 2023*

dataset is considered. The LFW database is a collection of face images that have been used for the classification of teaching and non-teaching images for the attendance system. The sample of images from the LFW dataset is introduced in Fig.3. Further, the dimension of all images as a constant image resizing system has been performed. A noise filtering system has been applied over images to remove the noise from LFW images. The Haar-Cascade face detector is applied over LFW based on 13,233 images of 5,749 individuals.

Figure 3. Sample of Images from the LFW dataset [27]

The following Table I represents hardware and software requirements for the proposed hybrid model.

TABLE I. HARDWARE AND SOFTWARE REQUIREMENTS

Hardware Requirement		Software requirement	
Device	**Description**	**Software**	**Version**
CPU (GPU)	Above 1ghz	**Windows**	10,11
Hard disk	512 GB	**Python**	3.10
RAM	8 GB		

E. *Accuracy comparison*

To predict the accuracy, the process of classification is performed [21] on teaching and non-teaching for the facial Attendance System. The ResNet model results calculated during the detection of non-teaching are 21, non-teachings considered in non-teaching are 981 whereas non-teaching considered in teaching are 19 to find out the True Positive, True negative values correspondingly.

DenseNet result has been evaluated [26] on teaching and non-teaching. During classification teaching faces predicted in the teaching group are 978, and teaching faces predicted in the non-teaching are 22. In the same manner, non-teachings considered in non-teaching are 981 whereas non-teaching considered in teaching are 19 to find out the True Positive and true negative values correspondingly.

Finally, the proposed Unified Model of RESDEN is considered for the classification of teaching and non-teaching for the facial Attendance System. During classification, teaching faces predicted correctly in the teaching group are 988, and teaching faces predicted in the non-teaching are 12. In the same manner, non-teachings considered correctly in non-teaching are 987 whereas non-teaching considered in teaching are 13 to find out the True Positive and true negative values correspondingly.

IV. EXPERIMENTAL RESULTS

This part considers noise removal techniques such as bilateral, blur, median, and Gaussian filters. After preprocessing, MSE, PSNR, and SSIM are calculated to evaluate the quality of an image. Then deep learning models such as ResNet, DenseNet, and RESDEN have been applied to perform binary and categorical classification.

A. *Noise removal techniques*

To improve the quality of the image, noise removal operation has been performed with the Bilateral, Blur, Median, and Gaussian filters. The research observed that the Gaussian filter provides refined results than the median, blur, and bilateral filters in many perspectives. Filtered images are presented in the Fig. 4.

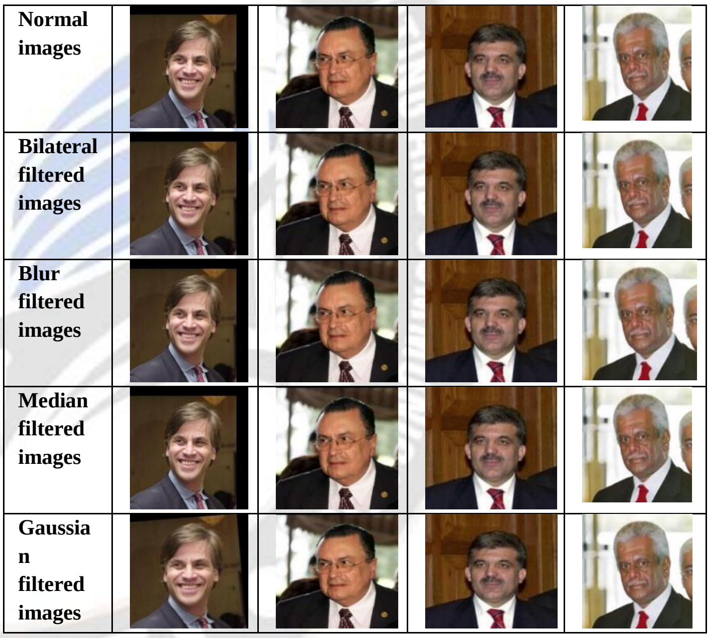


Figure 4. Filtered images of the LFW dataset

B. *Simulation for filtered images*

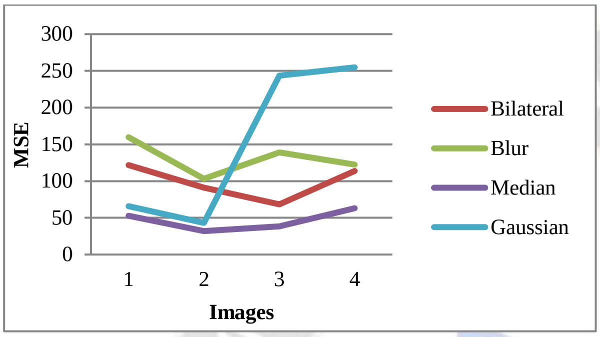
Firstly the simulation is performed for the Bilateral, Blur, Median, and Gaussian filters, and MSE, PSNR, and SSIM values are evaluated for the above filters. The overall result observed that the Gaussian filter provides better than other filters illustrated in Table II, and Fig. 5 respectively.

TABLE II. SIMULATION OF FILTERED IMAGES

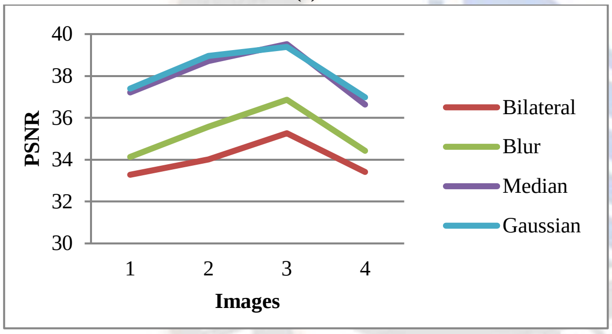
Filter	Image	MSE	PSNR	SSIM
Bilateral	1	121.765376	33.28158	0.903910
	2	91.167712	34.00903	0.902651
	3	68.321536	35.26316	0.913108
	4	**113.85552**	**33.4144**	**0.886784**
Blur	1	159.712912	34.13226	0.926720

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8322
***Article Received:** 12 July 2023 **Revised:** 31 August 2023 **Accepted:** 20 September 2023*

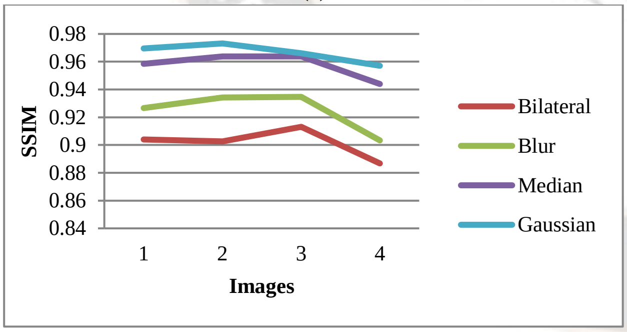
	2	103.128544	35.56713	0.934331
	3	139.090992	36.86374	0.934680
	4	**122.291488**	**34.42430**	**0.903340**
Median	1	52.530864	37.20478	0.95854
	2	31.946848	38.70812	0.96377
	3	38.521264	39.51798	0.96381
	4	**63.042310**	**36.6337**	**0.94396**
Gaussian	1	243.722032	39.38683	0.966050
	2	254.793072	36.98096	0.95710
	3	43.083712	38.96410	0.973154
	4	**66.00912**	**37.40066**	**0.969475**



(a)



(b)



(c)

Figure 5. (a), (b), and (c) Graphical representation of Comparative analysis of MSE, PSNR, and SSIM for different filters

C. Simulation for Segmentation

During the simulation, the dataset of LFW has been classified into teaching and non-teaching categories to simulate the accuracy of the attendance system. The image set for the teaching dataset is presented in Figure 6, and the image set used for the non-teaching dataset is represented in Fig. 7. HAAR Cascade has been applied over the image set to get the segmented image set. The results of segmented image datasets of teaching and non-teaching are illustrated in Fig. 8 and Fig. 9 relatively.

Figure 6. Image set of the teaching data set [27]

Figure 7. Image set of the non-teaching dataset [27]

Figure 8. Segmented Image set of the teaching data set

Figure 9. Segmented Image set of the non-teaching data set

D. Performance and Accuracy

During the training and testing of the proposed deep learning model, several parameters like epochs, validation step,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8322
***Article Received:** 12 July 2023 **Revised:** 31 August 2023 **Accepted:** 20 September 2023*

batch size, & optimizer are defined. These parameters are defined in Table III.

TABLE III. CONFIGURATION CHART

Parameters	Value
Epochs	100
Validation steps	800
Batch size	16
Target size	64x64
Optimizer	adam

Moreover, Simulation has been performed over Google Collaboratory. Five simulations have been performed for the ResNet model, DenseNet model, and proposed RESDEN model. During training, time taken is considered in hours for all three models. Here the proposed model has less training time as compared to Resnet and Densenet. Therefore, RESDEN performed better than simple ResNet and simple Densenet models which are shown with the help of Table IV and the performance evaluation graph represented in Fig. 10.

TABLE IV. PERFORMANCE COMPARISON OF DIFFERENT DATASETS

Simulation	ResNet	DenseNet	RESDEN
1	4:03	4:09	3:53
2	4:04	4:16	3:50
3	4:05	4:28	4:01
4	4:05	4:32	3:52
5	4:18	4:38	4:18

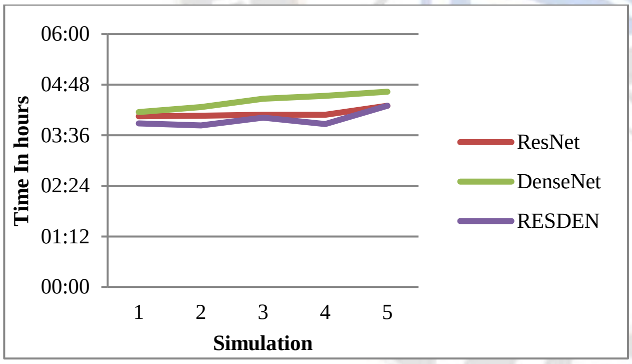


Figure 10. Performance evaluation

The accuracy of the model has been checked in binary and categorical class mode over a non-segmented image set where losses are considered as Binary_crossentropy and Mean_squared_error respectively in Table V. Binary Cross-Entropy and Mean Squared Error are two common loss functions used in machine learning, particularly in the context of neural networks and deep learning. They serve different purposes and are used for different types of tasks.

1) Binary Cross-Entropy (Log Loss):

Binary Cross-Entropy, also known as log loss or logistic loss, is primarily used for binary classification problems where you have two classes (e.g., 0 and 1). It measures the dissimilarity between the predicted probability distribution and the true binary labels. The formula for binary cross-entropy is as follows:

$$L(y,p) = -\frac{1}{N}\sum_{i=1}^{N}[y_i \cdot \log(p_i) + (1-y_i) \cdot \log(1-p_i)] \quad (4)$$

- y represents the true binary labels (0 or 1).
- p represents the predicted probabilities (usually output from a sigmoid activation function).
- N is the number of samples.

It penalizes large errors more heavily than small errors, making it sensitive to the confidence of the model's predictions.

2) Mean Squared Error (MSE):

Mean Squared Error is typically used for regression problems, where the goal is to predict continuous values. It measures the average of the squared differences between the predicted values and the true target values. The formula for mean squared error is as follows:

$$MSE(y,\hat{y}) = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2 \quad (5)$$

- y represents the true target values.
- $\hat{y}$ represents the predicted values.
- N is the number of samples.

It penalizes both small and large errors equally and is less sensitive to outliers compared to binary cross-entropy.

TABLE V. ACCURACY AFTER SIMULATION WITHOUT FILTERING AND SEGMENTATION

Loss	Class_mode	Validation Accuracy	Accuracy
Binary_cross entropy	Binary	0.4054	0.3391
Mean_squared_error	Categorical	0.6667	0.6667

The accuracy of the model has been checked in binary and categorical class mode over filtered and segmented images set where losses are considered as Binary_crossentropy and Mean_squared_error respectively in Table VI.

TABLE VI. ACCURACY AFTER SIMULATION, FILTERING, AND SEGMENTATION

Loss	Class_mode	Validation Accuracy	Accuracy
Binary_crossentropy	Binary	0.9129	0.8742
Mean_squared_error	Categorical	0.9235	0.9326

E. Comparative analysis of accuracy

ResNet, DenseNet, and RESDEN models are compared for accuracy. Here different classification models named ResNet, DenseNet, and RESDEN have been proposed with epochs of 30 and batch sizes of 16. Accuracy of 98.04%, 97.95%, and 98.75% has been obtained in the case of ResNet, DenseNet, and RESDEN respectively. On the other hand, Errors of 1.96%,

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
DOI: https://doi.org/10.17762/ijritcc.v11i9.8322
***Article Received:** 12 July 2023 **Revised:** 31 August 2023 **Accepted:** 20 September 2023*

2.05%, and 1.25% have been obtained in the case of ResNet, DenseNet, and RESDEN respectively illustrated in Table VII, and corresponding graphs are depicted in Figures 11 and 12.

TABLE VII. CLASSIFICATION OF DIFFERENT MODELS

Model	Epoch	Batch size	Accuracy (%)	Error Rate (%)
ResNet [25]	30	16	98.04	1.96
DenseNet [25]	30	16	97.95	2.05
RESDEN (Proposed)	30	16	98.75	1.25

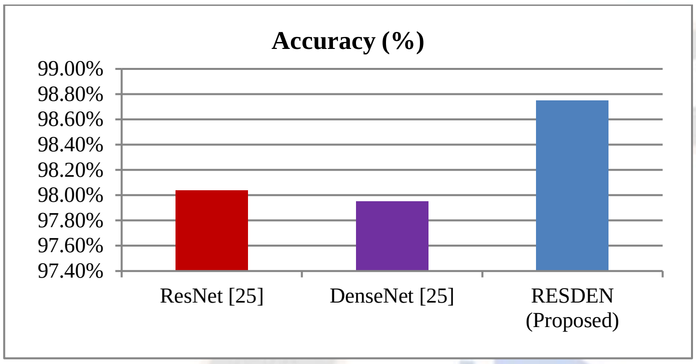


Figure 11. Comparison of Accuracy

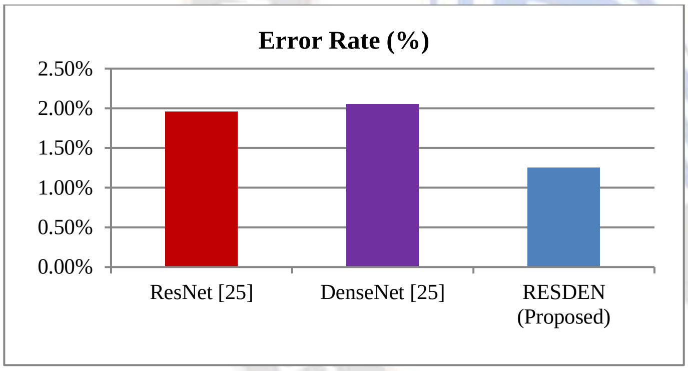


Figure 12. Comparison Error Rate

V. CONCLUSION

The findings of the simulation indicate that the Gaussian filter, when used with metrics such as MSE, PSNR, and SSIM, yields superior outcomes in the context of noise removal compared to other noise filters. The classification task is performed using the public banchmark LFW dataset. The RESDEN system incorporates a noise reduction method, a segmentation mechanism, and a CNN-based ResNet model integrated into DenseNet. The RESDEN architecture offers a higher level of precision compared to both ResNet and DenseNet. Ultimately, the hybrid model attain the accuracy of 98.7%. The simulation results demonstrate that the hybrid model presents a superior performance compared to the standard ResNet and DenseNet models.

VI. FUTURE SCOPE

A Deep Unified Model for Facial Attendance System will be used in government as well as private organizations. Such a model will be upgraded in the future using advanced classification mechanisms for the faculty or student attendance system. Appropriate noise reduction technology would be beneficial in upcoming research. Future models may use image sets with high image quality to check the consistency. The model can expand by adding multiple face detection and recognition concepts at a single time to make it more flexible.

REFERENCES

[1] K. Kavita and R.S. Chhillar, "Human Face Recognition and Age Estimation with Machine Learning: A Critical Review and Future Perspective", International Journal of Electrical and Computer Engineering Systems, Volume 13, Number 10, 2022 pp. 945-952.

[2] P. Tupe-Waghmare and R. R. Joshi, "Bibliometric Survey on Multipurpose Face Recognition System using Deep Learning," Libr. Philos. PractFebruary 2021.

[3] F. Kong, "Facial expression recognition method based on deep convolutional neural network combined with improved LBP features," Pers. Ubiquitous Comput., vol. 23, no. 3–4, pp. 531–539, 2019,

[4] Serign Modou Bah, Fang Ming, "An improved face recognition algorithm and its application in attendance management system", Array Elsevier Volume 5, March 2020, pp. 100014, 2019.

[5] Fahad P, M.d. Mahmudul, Md. Atiqur, Susan M, Moslehuddin M, Pandian V Face recognition based real-time system for surveillance. Intelligent Decision Technologies, IOS Press (2017), pp. 79-9211.

[6] Kirtiraj Kadam, Manasi Jadhav, Shivam Mulay and Tushar Indalkar, "Attendance Monitoring System Using Image Processing and Machine Learning", International Journal of Advance Engineering and Research Development, 2017.

[7] R. Zhang, L. Zhao, T. Lou, et al., "Automatic segmentation of acute ischemic stroke from DWI using 3-D fully convolutional DenseNets," IEEE Transactions on Medical Imaging2018, vol. 37, no. 9, pp. 2149–2160.

[8] Z. Lu, X. Jiang, and A. Kot, "Deep Coupled ResNet for Low-Resolution Face Recognition," in IEEE Signal Processing Letters, vol. 25, no. 4, April 2018, pp. 526-530.

[9] M. Z. Khan, S. Harous, S. U. Hassan, M. U. Ghani Khan, R. Iqbal, and S. Mumtaz, "Deep Unified Model for Face Recognition Based on Convolution Neural Network and Edge Computing," IEEE Access, vol. 7, no. c, pp. 72622–72633, 2019.

[10] X. Zhu et al., "Large-Scale Bisample Learning on ID Versus Spot Face Recognition," Int. J. Comput. Vis., vol. 127, no. 6–7, pp. 684–700, 2019.

[11] S. Almabdy and L. Elrefaei, "Deep convolutional neural network-based approaches for face recognition," Appl. Sci., vol. 9, no. 20, 2019.

[12] A. Elmahmudi and H. Ugail, “Deep face recognition using imperfect facial data,” Futur. Gener. Comput. Syst., vol. 99, pp. 213–225, 2019.

[13] Mohammed Ashik, Ramesh Patnaik Manapuram, Praveen B. Choppala. (2023). Observation Leveraged Resampling-Free Particle Filter for Tracking of Rhythmic Biomedical Signals. International Journal of Intelligent Systems and Applications in Engineering, 11(4s), 616–624. Retrieved from https://ijisae.org/index.php/IJISAE/article/view/2739

[14] B. Lodhi and J. Kang, “Multipath-DenseNet: a supervised ensemble architecture of densely connected convolutional networks,” Information Sciences, vol. 2019, 482, pp. 63–72.

[15] S. Sharma and V. Kumar, “37_2020 Voxel-based 3D face reconstruction and its application, Multimedia Tools and Applications”, Springer 2020, 79 pages–17330.

[16] X. Han and R. Jin, “A small sample image recognition method based on ResNet and transfer learning,” in Proceedings of the 2020 5th International Conference on Computational Intelligence and Applications (ICCIA), IEEE, Beijing, China, June 2020, pp. 76–81.

[17] L. Ke, Y. Deng, W. Xia, et al., “Development of a self-constrained 3D DenseNet model in automatic detection and segmentation of nasopharyngeal carcinoma using magnetic resonance images,” Oral Oncology, vol. 110,2020.

[18] Khawla Alhanaee, Mitha Alhammadi, Nahla Almenhali, Maad Shatnawi, “Face Recognition Smart Attendance System using Deep Transfer Learning,” Procedia Computer Science 192 (2021) 4093–4102.

[19] Y. Said and M. Barr, “Human emotion recognition based on facial expressions via deep learning on high-resolution images,” Multimed. Tools Appl., vol. 80, no. 16, 2021, pp. 25241–25253.

[20] M. K. Chowdary, T. N. Nguyen, and D. J. Hemanth, “Deep learning-based facial emotion recognition for human–computer interaction applications,” Neural Comput. Appl.2021, vol. 8.

[21] M. J. Khan, M. J. Khan, A. M. Siddiqui, and K. Khurshid, “An automated and efficient convolutional architecture for disguise-invariant face recognition using noise-based data augmentation and deep transfer learning,” Vis. Comput., vol. 38, no. 2,2021, pp. 509–523.

[22] Yohanssen Pratama, Lit Malem Ginting, Emma Hannisa Laurencia Nainggolan, Ade Erispra Rismanda, “Face recognition for presence system by using residual networks-50 architecture”, Vol. 11, No. 6, December 2021, pp. 5488~5496.

[23] G. Sunitha, K. Geetha, S. Neelakandan, A. K. S. Pundir, S. Hemalatha, and V. Kumar, “Intelligent deep learning based ethnicity recognition and classification using facial images,” Image Vis. Comput., vol. 121, 2022, p. 104-404.

[24] R. Ullah et al., “A Real-Time Framework for Human Face Detection and Recognition in CCTV Images,” Mathematical Problems in Engineering, vol. 2022. Hindawi Limited, pp. 1–12, Mar. 03, 2022. doi: 10.1155/2022/3276704.

[25] Z. Yu, Y. Dong, J. Cheng, M. Sun, and F. Su, “Research on Face Recognition Classification Based on Improved GoogleNet,” Security and Communication Networks, vol. 2022. Hindawi Limited, pp. 1–6, Jan. 05, 2022. doi: 10.1155/2022/7192306.

[26] 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. Walter de Gruyter GmbH, Jan. 01, 2023. doi: 10.1515/nleng-2022-0290.

[27] Md. F. Rabbi et al., “Autism Spectrum Disorder Detection Using Transfer Learning with VGG 19, Inception V3 and DenseNet 201,” Communications in Computer and Information Science. Springer Nature Switzerland, pp. 190–204, 2023. doi: 10.1007/978-3-031-23599-3_14.

[28] LFW - People (Face recognition). (2019, November 15). Kaggle. https://www.kaggle.com/datasets/atulanandjha/lfwpeople