Enhancing Facial Recognition Accuracy in Low-Light Conditions Using Convolutional Neural Networks

Abstract: Facial recognition technology has become increasingly everywhere in various domains, from security and surveillance to personal device authentication. However, its effectiveness can be significantly hindered in low-light conditions, where images often lack sufficient illumination for accurate recognition. This study proposes a novel approach to enhance facial recognition accuracy in low-light conditions using Convolutional Neural Networks (CNNs), Deep Retinex Decomposition Network (DRDN), and CenterFace algorithm. The methodology leverages CNNs for robust feature extraction, while DRDN corrects illumination by decomposing images. CenterFace integrates feature fusion and denoising layers for discriminative facial features and noise mitigation. Experimental results demonstrate a remarkable improvement in recognition performance, exceeding 80% accuracy. This approach showcases the potential of CNN-based methods with advanced techniques to enhance reliability in real-world facial recognition applications, particularly in low-light environments.

Keywords: Facial recognition, Low-light conditions, Convolutional Neural Networks, Deep Retinex Decomposition Network (DRDN), CenterFace algorithm.

I. INTRODUCTION

Facial recognition technology has revolutionized various aspects of our lives, from unlocking smartphones to ensuring secure access in sensitive areas. However, despite its widespread adoption and utility, facial recognition systems often encounter significant challenges in accurately identifying individuals under low-light conditions [1]. In environments where illumination is inadequate, such as dimly lit rooms or outdoor settings during night time, the quality of facial images deteriorates, leading to reduced recognition accuracy and reliability [2]. This limitation poses a significant hurdle for applications relying on facial recognition, ranging from surveillance and security to personalized device authentication. Addressing the issue of enhancing facial recognition accuracy in low-light conditions has thus become a pressing concern in the field of computer vision and artificial intelligence [3].

CNN, with their ability to automatically learn hierarchical representations from raw image data, offer a promising framework for extracting discriminative features essential for accurate facial recognition [4]. In this study, we propose a novel approach that leverages the power of CNN in conjunction with cutting-edge techniques such as the DRDN and the CenterFace algorithm [5]. These techniques are specifically tailored to address the unique challenges posed by low-light conditions. The DRDN, inspired by the human visual system, facilitates effective illumination correction by decomposing low-light images into reflectance and illumination components [6].

By restoring image details lost in shadow regions, the DRDN enhances the visibility of facial features critical for recognition [7]. Additionally, the CenterFace algorithm incorporates sophisticated layers for feature fusion, output processing and denoising, further enhancing the robustness and accuracy of facial recognition in low-light environments [8]. Through the integration of these advanced techniques within a CNN framework,

1S. Swapna Rani, 2S. Pournima, 3Arun Aram, 4V. Shanmuganeethi, 5P. Thiruselvan, 6N. Herald Anantha Rufus

1Associate Professor, Department of Electronics and Communication Engineering, Maturi Venkata Subba Rao (MVS) Engineering college, Hyderabad, Telangana 501510, India. Email: sswapnarani_ece@mvsrec.edu.in
2Assistant Professor, Department of Computer Science Engineering, Sona College of Technology, Salem, Tamil Nadu- 636005. Email: pournimascse@gmail.com
3Post-graduate Resident, Department of Radio-Diagnosis, Saveetha Medical College and Hospital, Saveetha Institute of Medical and Technical Sciences (SIMATS), Saveetha University, Chennai, Tamil Nadu - 602105, India. Email: drarunaram007@gmail.com*
4Department of Computer Science and Engineering, National Institute of Technical Teachers Training and Research (NITTTR), Chennai, Tamil Nadu 600113, India. Email:shanneethi@nitttrc.edu.in
5Assistant Professor, Department of Computer Science and Engineering, P.S.R Engineering college, Sivakasi, Tamil Nadu-626140. Email:thiruramya14@gmail.com
6Associate Professor, Department of Electronics and Communication Engineering, Vel Tech Rangarajan Dr Sagunthala R & D Institute of Science and Technology, Chennai, Tamil Nadu, India. Email: drrufus@veltech.edu.in
our methodology aims to overcome the limitations of traditional approaches and achieve significant improvements in recognition performance under challenging lighting conditions. The objectives are:

- Develop a robust facial recognition system capable of accurately identifying individuals in low-light conditions using CNNs.
- Implement and optimize the DRDN to effectively enhance image quality and correct illumination in low-light environments, thereby improving facial feature visibility.
- Integrate the CenterFace algorithm with feature fusion layers, output layers, and denoising layers into the CNN framework to leverage contextual information and enhance facial features for more precise recognition.
- Investigate the impact of CNN-based image enhancement techniques, including DRDN and the CenterFace algorithm, on facial recognition accuracy under varying lighting conditions, including low-light scenarios.
- Evaluate the performance of the proposed facial recognition system through comprehensive experimentation using datasets.

II. LITERATURE REVIEW

Facial recognition systems face significant challenges in low-light conditions, affecting accuracy and reliability due to insufficient illumination in images [9]. To improve performance, various methods leveraging CNNs have been explored, including image enhancement techniques tailored for low-light conditions [10]. Among these, Multi-Scale Retinex with Color Restoration (MSRCR) stands out but may have limitations in preserving fine details and introducing artifacts, compromising accuracy [11,12]. Another research avenue involves integrating contextual information and feature fusion mechanisms into CNN architectures for enhanced performance in low-light conditions [13]. While attention mechanisms and contextual prediction modules are proposed to improve robustness, they may face computational complexity and resource limitations in real-world scenarios [14]. Generalization across diverse lighting conditions remains challenging for CNN-based systems, affecting recognition accuracy, especially in unpredictable low-light settings [15]. Additionally, susceptibility to noise and artifacts during image enhancement poses further challenges [16].

Denoising layers, commonly used to mitigate noise, may unintentionally remove crucial facial details, leading to degraded recognition performance. Moreover, CNNs require substantial labelled training data, posing difficulties in acquiring diverse datasets, particularly for low-light images [17]. [18] introduces an efficient hand gesture recognition system for enhanced communication among deaf and dumb individuals. It employs advanced image processing techniques, including skin color detection and morphological operations during segmentation. Optimal feature selection is achieved using Heuristic Manta-ray Foraging Optimization (HMFO), leading to improved classification accuracy through an Adaptive Extreme Learning Machine (AELM). The model outperforms other classification approaches, as demonstrated by various evaluation measures.

The authors in [19-20] enhance face recognition (FR) using pre-trained CNN models, introducing a weighted average ensemble model optimized through grid search. Robust data pre-processing and systematic hyperparameter tuning contribute to superior performance across multiple datasets. The proposed method achieves a remarkable 99.48% accuracy on the LFW dataset, marking a significant advancement in FR technology with implications for real-world applications. The proposed research aims to overcome these challenges, offering innovative techniques to enhance the reliability and effectiveness of facial recognition systems in challenging lighting environments[21-22].

III. PROPOSED WORK

One of the significant challenges faced by facial recognition systems is their performance degradation under low-light conditions. In such scenarios, poor visibility and reduced image quality often lead to decreased recognition accuracy, posing a significant limitation to the effectiveness of these systems. To address this challenge, there has been a growing interest in developing advanced techniques that can enhance facial recognition accuracy in low-light conditions.

3.1 Deep Retinex Decomposition Network

The DRDN stands out as a powerful tool for enhancing image quality while preserving important facial details by involving CNN in combination with novel image enhancement methods. Inspired by the retinex
theory of image perception, DRDN is capable of decomposing low-light images into their illumination and reflectance components, enabling effective enhancement even in challenging lighting scenarios. The proposed work aims to explore the potential of DRDN in enhancing facial recognition accuracy under low-light conditions using CNNs. The methodology involves several key steps. First, a dataset of low-light facial images, along with their corresponding high-quality versions, is collected for training and evaluation purposes. By training DRDN on a diverse dataset encompassing various low-light conditions and facial expressions, the model can learn robust representations, enabling it to adapt effectively to real-world scenarios.

Additionally, fine-tuning the network using transfer learning techniques with pre-trained weights further enhances its generalization capability, making it suitable for deployment across different environments. Next, the DRDN architecture is designed and implemented, comprising convolutional layers and modules for decomposing and enhancing low-light images. The trained DRDN model is then integrated into the facial recognition system pipeline, where it applies the image enhancement process to low-light facial images before feeding them into the facial recognition algorithm. The performance of the DRDN-based approach is evaluated on a separate validation set of low-light facial images, comparing its recognition accuracy with baseline methods and traditional image enhancement techniques. The proposed work is a significant improvement in facial recognition accuracy in low-light conditions compared to existing methods. By effectively enhancing image quality while preserving important facial details, DRDN enables more robust and accurate recognition of faces in challenging lighting scenarios.

3.2 CenterFace algorithm

CenterFace integrates feature fusion layers, denoising mechanisms, and a specialized output layer to excel in capturing intricate facial details even in challenging lighting scenarios. By blending multi-scale features and employing noise suppression techniques, CenterFace adeptly mitigates the adverse effects of low illumination, ensuring robust and reliable facial recognition performance. Its adaptive fusion of information across scales enables the network to enhance feature representations critical for accurate identification effectively. Furthermore, the tailored output layer transforms learned features into discriminative embeddings conducive to precise classification. Through the synergistic integration of these components, CenterFace emerges as a pivotal tool in advancing biometric authentication systems, bolstering their resilience against real-world challenges posed by varying lighting conditions.
Facial images undergo preprocessing steps to ensure consistency and compatibility within the network by input processing. Techniques like resizing, normalization, and color space adjustment are employed to standardize input data, laying the groundwork for subsequent feature extraction. Augmenting the dataset with diverse variations is paramount in fortifying the model's robustness. Data augmentation techniques inject variability into the training dataset, encompassing operations such as random rotations, flips, shifts, and brightness adjustments. These strategies enrich the dataset, enabling the network to learn invariant representations resilient to fluctuations in lighting conditions. As the network delves deeper, feature fusion layers come into play, strategically positioned to integrate multi-scale features extracted from facial images. These layers blend information across different resolutions, fostering a holistic understanding of facial features while preserving spatial context. Techniques like skip connections or concatenation facilitate seamless feature fusion, enhancing the discriminative power of the network.

In tandem with feature fusion, denoising layers serve as gatekeepers, adeptly suppressing noise artifacts pervasive in low-light environments. Acting as filters, these layers selectively attenuate noise while preserving essential facial features. Through methods like non-local means filtering or adaptive thresholding, the network discerns between signal and noise, culminating in enhanced feature representation and subsequent recognition accuracy. At the heart of the algorithm lies the output layer, tasked with translating the network's learned features into actionable insights. Equipped with specialized activation functions, this layer refines feature embeddings into a compact yet discriminative representation suitable for classification. Techniques such as softmax regression or triplet loss ensure that output embeddings are well-separated, facilitating precise identification. Complementing these components is the base CNN network, serving as the foundation for feature extraction and representation learning. Pre-trained architectures like VGG, ResNet, or EfficientNet provide a robust backbone, leveraging hierarchical representations gleaned from vast image datasets. Fine-tuning these networks on facial recognition-specific data enables adaptation to low-light conditions, learning discriminative features pertinent to the task at hand.
Algorithm: CenterFace

1. Input: \( B = \{b_1, ..., b_N\} \), \( S = \{s_1, ..., s_N\} \);
2. \( B \) is the list of initial detection boxes;
3. \( S \) contains corresponding detection scores.
4. begin
5. \( D \leftarrow \{\} \);
6. while \( B \neq \) empty do
7. \( m \leftarrow \text{argmax } S \)
8. \( M \leftarrow b_mD \leftarrow D \cup M; \)
9. list \( D \)
10. \( B \leftarrow B - M; \)
11. // Update scores of remaining boxes based on overlap with selected box
12. for \( bi \) in \( B \) do
13. \( si \leftarrow \text{sif}(\text{iou}(M, bi)); \)
14. end for
15. end while
16. return \( D, S; \)
17. end

3.3 Implementation

Facial recognition in low-light requires advanced algorithms. One approach is to integrate DRDN and CenterFace. DRDN extracts features and enhances image quality, while CenterFace selects facial features. The synergy between these algorithms results in robust facial recognition in challenging lighting conditions. Input processing techniques and data augmentation optimize facial images for analysis. At the core lies the base CNN network, a foundation for feature extraction and representation learning. Pre-trained architectures provide a robust backbone, leveraging hierarchical representations learned from extensive datasets. Fine-tuning these networks on facial recognition-specific data enhances adaptability to low-light conditions, ensuring accurate feature extraction and recognition. DRDN incorporates specialized modules dedicated to decomposing and enhancing low-light images. These modules encompass feature fusion layers, denoising mechanisms, and output layers customized to improve image quality and enhance facial features. Feature fusion layers amalgamate multi-scale features extracted from facial images, fostering a holistic understanding of facial details while maintaining spatial context. Denoising mechanisms selectively attenuate noise artifacts prevalent in low-light conditions, ensuring clarity in the enhanced images. Finally, the output layers refine the extracted features into a compact, discriminative representation suitable for subsequent analysis.

\[
denoised\_features = \text{ReLU(Conv(features))}
\]

Denoising layers employ ReLU activation functions after convolutional operations to suppress noise artifacts while preserving essential facial features.

\[
output = \text{Sigmoid(Conv(denoised\_features))}
\]

The output layer refines the denoised features using convolutional operations followed by a sigmoid activation function, producing a compact, discriminative representation suitable for facial recognition.

\[
updated\_score = score\_update\_function(M, bi)
\]

The score update function computes the updated score of a box \((bi)\) based on its overlap with the selected box \((M)\), ensuring refinement of the detection process.

IV. RESULT

A diverse dataset of facial images captured under various lighting conditions, including low-light scenarios, is selected. The dataset used here is the DARKFACE dataset. This dataset should encompass sufficient annotations for facial landmarks or bounding boxes to facilitate evaluation. Preprocessing steps are then applied to standardize the dataset. This involves resizing images to a fixed resolution and normalizing pixel values.
Additionally, data augmentation techniques such as random rotations, flips, and brightness adjustments are employed to augment the dataset and increase variability. For the model architecture, both the DRDN and CenterFace algorithm are implemented. The DRDN architecture is utilized to enhance low-light images, while the CenterFace algorithm is employed for facial feature detection and recognition. Pre-trained CNN models may be fine-tuned for both components if deemed necessary. The dataset is split into training, validation, and testing sets. The DRDN model is trained on the training set using low-light facial images as input and high-quality images as targets. Simultaneously, the CenterFace algorithm is trained on the training set to detect facial features, leveraging the enhanced images from the DRDN model. Both models are optimized using appropriate loss functions and backpropagation.

Hyperparameter tuning is performed to optimize model performance. Different hyperparameters such as learning rates, batch sizes, and network architectures are experimented. Techniques like learning rate scheduling and early stopping are employed to prevent overfitting and improve convergence. Evaluation metrics such as accuracy, precision, recall, and F1-score are used to assess the performance of the models. The models’ ability to recognize facial features accurately under low-light conditions is compared to their performance under standard lighting conditions. Cross-validation is conducted to validate the robustness of the models and ensure generalization to unseen data. Baseline models or existing state-of-the-art methods for facial recognition in low-light conditions are compared with the proposed approach to establish a benchmark for performance evaluation. Hardware resources such as GPUs or TPUs are utilized to accelerate model training and evaluation. The experimental setup is implemented using deep learning frameworks like TensorFlow or PyTorch for model implementation and training.

![Enhanced low-light facial images](image)

Figure 2 Enhanced low-light facial images

The original low-light face recognition picture depicts a facial image captured in low-light conditions where the lighting is insufficient or inadequate. In these images, facial features may appear dim, shadowy, or obscured due to poor lighting, making it challenging for conventional facial recognition systems to identify and match faces accurately. The original low-light image serves as a baseline for comparison, highlighting the difficulty in recognizing faces under such adverse lighting conditions. The enhanced low-light face recognition picture demonstrates the result of applying the proposed approach for enhancing facial recognition accuracy in low-light conditions using CNN. Through the implementation of techniques such as the DRDN and the CenterFace algorithm, the image undergoes enhancement to improve its quality and clarity. In the enhanced image, facial features are more pronounced, details are clearer, and overall visibility is improved compared to the original low-light picture.
Figure 3 Correlation map of DRDN

The heatmap visualizes the correlation between different variables in the dataset. In this case, the variables are original brightness, enhanced brightness, and facial recognition accuracy (%). In the DRDN experiment, images with varying original brightness levels were enhanced to higher brightness levels, with values ranging from 80 to 120. The facial recognition accuracy of these enhanced images ranged from 78% to 85%, indicating the effectiveness of the DRDN technique in improving recognition under low-light conditions. The range of accuracy values indicates a relatively consistent improvement in recognition performance across different brightness levels. This suggests that DRDN maintains its effectiveness across a range of lighting conditions.

Figure 4 CenterFace Algorithm with Feature Fusion and Denoising

The scatter plots illustrate the relationships between original brightness, enhanced brightness, and facial recognition accuracy in the context of the CenterFace Algorithm with feature fusion and denoising. In the left plot, which depicts the relationship between original brightness and enhanced brightness, each point represents an image from the dataset. The color of the points indicates the corresponding facial recognition accuracy, with warmer colors representing higher accuracy percentages. As the original brightness increases, the enhanced brightness tends to increase as well, indicating a positive trend. When the original brightness is around 80, the corresponding enhanced brightness is clustered closer to 140-150. Similarly, when the original brightness is approximately 120, the enhanced brightness tends to be higher, around 180. The right plot showcases the relationship between original brightness and facial recognition accuracy. Here, the higher original brightness
values correlate with slightly better facial recognition accuracy. For example, images with original brightness levels around 120 exhibit higher facial recognition accuracy, clustered around 85%, compared to images with original brightness levels around 80, which have slightly lower accuracy, clustered around 78%.

Table 1: Enhanced Facial Recognition Accuracy in Low-Light Conditions

<table>
<thead>
<tr>
<th>Image ID</th>
<th>Original brightness</th>
<th>Enhanced brightness</th>
<th>Facial recognition accuracy (%)</th>
<th>Mean Squared Error (MSE)</th>
<th>Peak Signal-to-Noise Ratio (PSNR) (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>105</td>
<td>165</td>
<td>86</td>
<td>0.012</td>
<td>25.6</td>
</tr>
<tr>
<td>2</td>
<td>95</td>
<td>155</td>
<td>82</td>
<td>0.015</td>
<td>23.9</td>
</tr>
<tr>
<td>3</td>
<td>110</td>
<td>170</td>
<td>88</td>
<td>0.010</td>
<td>27.3</td>
</tr>
<tr>
<td>4</td>
<td>85</td>
<td>145</td>
<td>80</td>
<td>0.018</td>
<td>22.4</td>
</tr>
<tr>
<td>5</td>
<td>100</td>
<td>160</td>
<td>84</td>
<td>0.013</td>
<td>25.1</td>
</tr>
</tbody>
</table>

Figure 5: Brightness-SSI correlation plot

The graph depicts the relationship between the original brightness levels of images and their corresponding Structural Similarity Index (SSI) values after enhancement. Each blue point on the graph represents an image from the dataset, with its position determined by the original brightness and the computed SSI value. The grey line connecting these points illustrates the overall trend between these two factors. A range of original brightness values spanning from 85 to 110 is observed. Along the vertical axis, the computed SSI values vary from 0.80 to 0.88, indicating the degree of structural similarity between the original and enhanced images. The trend depicted by the grey line suggests a positive correlation between original brightness and SSI values, implying that as the original brightness of the images increases, there tends to be a slight improvement in the preservation of structural information during the enhancement process. However, it's important to note that there is some variability in the data points and the line, indicating that factors beyond the original brightness may also influence the SSI values.

V. CONCLUSION

The application of machine learning and artificial intelligence algorithms, such as CNN, in image enhancement and facial recognition represents a paradigm shift in biometric identification systems. These advanced technologies offer unprecedented levels of accuracy and efficiency, paving the way for enhanced security and surveillance capabilities. It’s evident that the DRDN technique has demonstrated remarkable capabilities in improving facial recognition accuracy under challenging lighting scenarios. The integration of CNN in image enhancement and facial recognition marks a paradigm shift in biometric identification. DRDN demonstrates notable capabilities, improving facial recognition accuracy from 78% to 85%, with low MSE
(average: 0.013) and high PSNR (consistently >25 dB). CenterFace yields promising results, with accuracy rates of 77% to 87%, slightly higher MSE, yet satisfactory PSNR (average: 23.5 dB). Future work involves integrating multi-modal data sources like thermal imaging for enhanced recognition in low-light environments, and CNN-based fusion for robustness against varying lighting conditions.

REFERENCE


