Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

IMPROVING FACE RECOGNITION IN LOW ILLUMINATION CONDITION USING COMBINATION OF IMAGE ENHANCEMENT AND FACE RECOGNITION METHODS

OWEN JACKSON DHARMADINATA[1], GEDE PUTRA KUSUMA[2]

Computer Science Department, BINUS Graduate Program - Master of Computer Science,

Bina Nusantara University, Jakarta, Indonesia, 11480

E-mail: [1]owen.dharmadinata@binus.ac.id, [2]inegara@binus.edu

ABSTRACT

Some face recognition models are constantly challenged by low illumination. However, we can address this issue by employing a low illumination image enhancement filter on the face recognition model. This study compares four low illumination image enhancement methods: MIRNet, Adaptive Gamma Correction, RetinexNet, and Retinex. SSIM evaluates these four models using the LOL dataset. The assessment findings are applied to both the ArcFace model and Facenet for face recognition, and they are evaluated with a combination of the Essex Faces96 and Essex Grimace datasets. The final model based on our suggested approaches is a combination of MIRNet and ArcFace models and MIRNet and Facenet models. MIRNet and Facenet combination perform better in higher illumination percentages, whereas MIRNet and ArcFace combination performs better in lower illumination percentages. At both 20 and 10% illumination percentage, the combined approach of low illumination image enhancement MIRNet method and Facenet model obtained Rank-1 Accuracy of 0.96 and 0.94. While the combination of MIRNet and ArcFace achieved a 0.84 Rank-1 Accuracy for a 05% illumination percentage.

Keywords: *Face Recognition, Low Illumination Face Recognition, Deep Learning, Facenet Model, ArcFace Model*

1. INTRODUCTION

Face Recognition has become a widespread issue in recent years since it allows the government and the general public to identify a person's identity by identifying their features. The majority of Face Recognition technology has been utilized in workplaces for face recognition to register their attendance time, where they are now working, and when they are leaving work.

In certain situations, the authorities employed Face Recognition technology to identify a wanted criminal or an escaped criminal fugitive by identifying their characteristics [1]. Some security cameras do not have built-in infrared or night vision features since a security camera capable of infrared or night vision is rather expensive. In that instance, the performance of Face Recognition may suffer degradation caused by low illumination conditions. In this study, we aimed to evaluate the effectiveness of face recognition in low illumination conditions by using a combination of the Essex Faces96 [2] and Essex Grimace [3] datasets and a newly created Low Illumination Image Enhancement (LOL) [4] dataset.

State-of-the-art methods in Face Recognition with normal illumination conditions have proven to perform at an extraordinary rate, with an accuracy of 99.83 percent in both ArcFace [5] and MagFace [6] using LFW as the dataset, while it is true that both ArcFace and MagFace perform an astounding performance in normal illumination condition dataset. It has not been demonstrated that the same result can be obtained in a dataset with low light. Rahman *et al*. introduced a solution to the low illumination condition dataset by adding the CLAHE Technique, an image enhancement method that changes the contrast of an image to be more evenly spread, which resulted in 76.92% accuracy; however, the accuracy drops dramatically at feature extraction steps in images with extremely low illumination conditions [7]. While Huang *et al.* suggest a CNN-based image improvement with PSNR and SSIM values of 20.37 and 0.73,

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

respectively, for the best-performing Image Enhancement in a Low Illumination scenario [8].

The Essex Faces96 [3] and Essex Grimace [4] datasets have limitations for use in low illumination face recognition due to the level of illumination in the images being insufficiently low. To address this issue, we used the Pillow Library to synthesize low illumination images from a set of indoor images with normal illumination to further decrease the level of image illumination in the datasets. The synthesized dataset was used to train ArcFace and FaceNet models

Image Enhancement improves Face Recognition in Low Illumination Conditions by utilizing a method that increases the quality and information an image provides, often an image enhancement that includes spatial filtering, contrast enhancement, density slicing, and many more. Some image enhancements concentrate on enhancing an image with poor light. Several research has offered state-of-the-art methods in boosting a low illumination image, such as [9], which employ adaptive gamma correction and histogram equalization to accomplish contrast enhancement while simultaneously increasing the image's color.

To improve the visibility of the images, contain the synthesized dataset, we used image enhancement methods that is trained using LOL dataset [4], We then applied these models to the enhanced images from the Synthesized combination of Essex Faces96 and Essex Grimace datasets, which were cropped to 112x112 and 160x160 respectively.

Overall, the goal of this research is to investigate the potential of image enhancement techniques to improve the performance of face recognition in low illumination conditions and to provide a better understanding of the impact of illumination in Face Recognition Performance, which in the research conducted in [10] and [11] used regular illumination and have not been proven to achieve the same performance at other levels of illumination. While also suggest a method that utilizes image enhancement to rectify the low illumination condition in Face Recognition, which is a state-of-the-art method with the purpose of increasing security given by Face Recognition that uses non-built-in night vision cameras. The state-of-the-art image enhancement models that will be explored in this work include Adaptive Gamma Correction (AGC) [12], Retinex [13], RetinexNet [14], and MirNet [15]. Face Recognition Models that will be explored include ArcFace [5] and FaceNet [16]. The combination of these models and methods is then evaluated using the Essex Faces96 [2] and Essex Grimace [3] datasets.

The contributions of this paper are as follows:

- Finding the best combination of the Image Enhancement method and Face Recognition Model for Low Illumination Face Recognition.
- Evaluate the impact of low illumination levels on the performance of Face Recognition in Accuracy.

2. RELATED WORKS

This section divides related works into four sub-sections: face recognition in normal illumination, face Recognition in low illumination, low illumination image enhancement, and finally, a summary of related works. These related works aim to analyze the best performance model for this research.

2.1 Low Illumination Image Enhancement

Yang *et al.* suggested a technique for picture enhancement in low illumination circumstances that employs both retinex decomposition and adaptive gamma correction [17]. To begin, the suggested technique employs a pyramid decomposition network to extract multi-scale characteristics that increase the efficacy of retinex decomposition. The deconstructed illuminated image was further enhanced by adjusting the image illumination level using adaptive gamma correction. Experimental performance was assessed using PSNR, SSIM, NIQE, BIQI, and BRISQUE, yielding 5.05, 0.59, 0.48, 18.89, 0.75 in the LOL dataset, and 4.29, 0.69, 0.66 in the MEF dataset for NIQE, BIQI, and BRISQUE, respectively. In contrast, the LIME dataset yielded NIQE at 3.80, BIQI at 0.65, and BRISQUE at 0.70. Finally, the VV dataset yields NIQE values of 2.6967, BIQI values of 0.6635, and BRISQUE values of 0.6791.

Zhuang and Guan suggested a solution for weakly illuminated images based on a deep learning network, which is inspired by the derived image and Retinex theory [18]. The proposed method employs image reconstruction to improve the low-illuminated image, which is obtained by training the deep decomposition network with both low-illuminated and normal-illuminated images. In

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

addition, the method employs an enhancement strategy to improve image brightness and contrast by using derived images that employ Adaptive Gamma Correction with Weighting Distribution (AGCWD). The experimental performance is evaluated and quantified using PSNR, RMS (Root Mean Squared), and DE datasets, yielding a score with a maximum of 18.71, 7.82, and 87.12 for PSNR, RMS, and DE scores, respectively.

Ma *et al*. proposed employing an adaptive contrast enhancement based on Retinex Model and Guided Filter to solve a low illuminated lane line acquired by a captured image [19]. Experimental performance is evaluated using Tusimple and Culanedataset datasets, which are chosen randomly, with PSNR and Epoch scores ranging from 18.1 to 7.73 and from 7.61 to 6.83.

Huang *et al*. presented a method for image enhancement in low illumination images utilizing a convolutional neural network and loss function from Peak-Signal-to-Noise (PSNR) loss, multi-scale Structural Similarity (MS-SSIM), and a combination of both PSNR loss and MS-SSIM loss [8]. Using the Multi-Exposure picture dataset, experiment outcomes are measured using PSNR, SSIM, and HS with scores of 20.370, 0.732, and 0.594, respectively.

Tang and Li suggested an image enhancement approach based on merging images that have been enhanced using gamma correction and Contrast Limited Adaptive Histogram Equalization (CLAHE), which creates two new weight maps and employs them for multi-scale fusion [20]. The findings of the experiments were measured using the Natural Image Quality Evaluator (NIQE) and the Integrated Local NIQE (ILNIQUE), which scored an average of 3.73 and 26.20, respectively.

Zamir *et al.* presented an updated method for image enhancement in low illumination images replacing mirnet with MIRNet-v2 [21]. Which uses non-local attention mechanism for capturing contextual information rather than using spatial and channel attention mechanism which in MIRNet [15] uses. The findings of the experiments is conducted using LOL dataset for low light image enhancement, which uses both PSNR and SSIM method. Result shows that MIRNet-v2 resulted in 24.74 in PSNR and 0.851 in SSIM.

2.2 Face Recognition in Normal Illumination Condition

William *et al*. proposed a study to compare and assess the performance of the FaceNet face recognition module for face detection while utilizing a pre-trained model, both of which are VGGFace2 and CASIA- WebFace [16]. Experimental performance was examined in both models VGGFace2 and Casia- WebFace using a variety of datasets, including Yale, JAFFE, AT&T, Georgia Tech, Essex faces94, Essex faces95, Essex faces96, and Essex grimace. VGGFace2 Model received scores of 100, 100, 100, 100, 99.37, 100, 77.67, and 100, whereas Casia-WebFace received scores of 98.9, 100, 97.5, 100, 99.37, 99.65, 76.86, and 100.

Deng *et al*. proposed Additive Angular Margin Loss (ArcFace), a novel approach for Face Recognition that combines similarity learning by employing distance metric learning for the classification task with Angular Margin Loss [5]. Experimentation performance was evaluated using several datasets. However, the LFW dataset performed best, with a verification performance of 99.83 percent.

Wang *et al*. presented CosFace, a Large Margin Cosine Loss (LMCL) for Deep Face Recognition, which employs LMCL as an alternative to the softmax loss function to normalize features and weight vectors to enhance further decision margin [22]. Experimentation performance was evaluated using LFW, YTF, and other datasets, with LFW scoring 99.73 and YTF scoring 97.6.

Huang *et al*. suggested a Face Recognition approach based on adaptive curriculum learning loss that employs both easy and hard samples in the early and late phases [23]. Experiment performance was evaluated using different datasets, including LFW, CFP-FP, CPLFW, AgeDB, and CALFW, with results of 99.80, 98.37, 93.13, 98.32, and 96.20, respectively.

Meng *et al*. introduced a technique for Face Recognition that uses MagFace, a universal learning feature extracted from the diverse quality of the same image, to learn the face embedding for each loss, hence preventing models from overfitting [6]. Experiment performance was evaluated using many datasets, including LFW, CFP-FP, AgeDB-30, CALFW, and CPLFW, with results of 99.83, 98.46, 98.17, 96.15, and 92.87, respectively.

Dang TV introduced a smart home management system for face recognition that uses a Deep Convolutional Neural Network that is based on Arcface Model [10]. Experiment performance was evaluated using 10 individuals each with their own 60 images, result shows that ArcFace achieve accuracy at 94 to 97 percentage.

Asmara *et al.* conduct a study using both ArcFace and FaceNet for Face Recognition in google cloud platform as an attendance system mobile application, this research uses model from Deepface library [11]. Study shows that accuracy resulted in 78% and 19% for both ArcFace and FaceNet respectively without additional artifical lightning.

2.3 Face Recognition in Low Illumination Condition

Huang and Chen proposed a method for low illumination face recognition using a feature reconstruction network that creates an image by merging both the raw face image and the improved illuminated face image. Face Recognition is performed in this study using ArcFace as the face feature extractor [24]. Experimental performance testing yields a 2.5% improvement in verification accuracy when utilizing the ArcFace and OpenFace models in the Specs on Faces (SOF) dataset.

Rahman *et al.* proposed a solution for Face Recognition under low illumination conditions that use the adaptive image enhancement CLAHE technique to improve image illumination quality and Fisherface as the face recognition method that uses both Principal Component Analysis (PCA) and Fisher's Linear Discriminant (FLD) to reduce input data and speed up the classification and recognition process [7]. The experimentation performance is evaluated using self-collected data that is processed to produce new types of images with brightness levels ranging from -10 to -100. The outcome of adding an image enhancement up to -80 brightness level with accuracy at 76.92, while the face recognition model could distinguish faces with -70 brightness level with accuracy at 77.69 without using any image enhancement and solely by utilizing fisherface.

Agarwal *et al.* proposed a solution to low illumination conditions in Face Recognition by combining Thermal Images, which were utilized to detect faces further and extract them using the blob approach, and Visual Images from the original dataset, which were transferred using Gradient Transfer. This work employs a CNN-based model called MobileNet to conduct Face Recognition, which performs quicker inferencing than a standard CNN [25]. Experimentation is carried out with a self-gathered dataset of 100 individuals, each with ten images. Performance was evaluated using three models: NasnetMobile (75.4% accuracy), MobilenNet v2 (93.3% accuracy), and MobileNet v2 (95.7% accuracy). The highest performing model, MobileNet v2, increased the accuracy of its initial visual picture by 5%.

Ren *et al.* presented a method for Face Recognition in low illumination conditions by performing image enhancement utilizing homomorphic filtering to minimize image noise and improve image quality and image multiplication to brighten the image. The Face Recognition approach used in this study is Two Dimensional Principal Component Analysis (2DPCA), which employs a characteristic matrix to categorize the identities [26]. This experiment uses the YALE database, which has 165 individuals with 11 images associated with their identity. Experimentation performance resulted in a 10% improvement in accuracy to the initial performance by adopting the provided approach, which resulted in a 90% accuracy.

Kang and Pan suggested a low illumination Face Recognition method that employs an image denoising filter with histogram equalization to improve the picture contrast value, followed by a logarithmic adjustment to augment the image details. Finally, image features may be recovered using high-pass filtering. The proposed technique performs face recognition using Principal Component Analysis; the overall Low Illumination Face Recognition algorithm is known as DELFN [27]. Experiment performance is evaluated using three databases: Yale Face Database B, Extended Yale Face Database B, and CMU PIE Face Database, with a 97.27% accuracy.

2.4 Summary of Related Works

Based on a review of the above papers, the best Face Recognition model for regular dataset would be both ArcFace [5] and MagFace [6], with each performing in 99.83 percent at LFW dataset, while regular illumination performance is excellent. Although ArcFace and MagFace function admirably in a regular illumination dataset, they

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

cannot be guaranteed to behave similarly in a low illumination dataset.

FaceNet [16] has demonstrated an outstanding result in verification accuracy for both Essex faces96 at 77.67 percent and Essex grimace at 100 percent, which have relatively low illumination images than others. Rahman *et al*. present methods for Face Recognition in Low Illumination Conditions with 76.92% accuracy. This is because, unlike Rahman *et al*., other methods demonstrated the usage of a database with many levels of illumination rather than simply low illumination images, which can be perceived as inaccurate in a low illumination environment-only database.

Aside from Face Recognition in regular illumination and Face Recognition in low illumination, low illumination image enhancement plays a significant role in achieving a low illumination Face Recognition model, which MIRNet-v2 [21] has successfully achieved with their proposed solution by using a Learning Enriched Features approach with a PSNR and SSIM of 24.74 and 0.851 respectively.

In the more recent studies of Face Recognition that utilizes both FaceNet and ArcFace uses a dataset that amounted to 600 images [11] which is too little to be proven accurate with large dataset, while in another recent study uses a self-gathered dataset.

Concluding through reading and experimenting with several models, this study proposed utilizing FaceNet and ArcFace as the Face Recognition model, which in William *et al*. study has been tested in a dataset with a low illumination setting, which is a mix of both Essex faces96 and Essex Grimace. For low-light image improvement, this study offered four methods: Retinex, MirNet (which employs a CNN methodology), AGC, and RetinexNet. These models will be tested using three dataset LOL dataset for image enhancement, while Essex faces96 and Essex Grimace will be used to test Face Recognition performance.

3. THEORY AND METHOD

This section divides theory and method into two sub-sections, both method and models used in this study, consisting of image enhancement and face recognition.

3.1 Image Enhancement Methods

Image Enhancement is utilized to improve Face Recognition and is mostly employed in this study as a low-light image enhancement. It improves the face recognition module's recognition rate. Mirnet, Retinex, RetinexNet, and Adaptive Gamma Correction are some of the Image Enhancement methods that will be explored in this study.

3.1.1 Mirnet

MirNet is a Low Illumination Image Enhancement method that's based on a Convolutional Neural Network that combines contextual information from multi-scale residual blocks through learning enriched set of features, while al-so maintaining the spatial details from the high-resolution image [15]. There are various keys in the multi-scale residual block, including parallel multi-resolution convolution streams, information transfer across multi-resolution streams, spatial and channel attention approaches, and attention-based multi-scale. The overall architecture of MIRNet are illustrated at Figure 1.

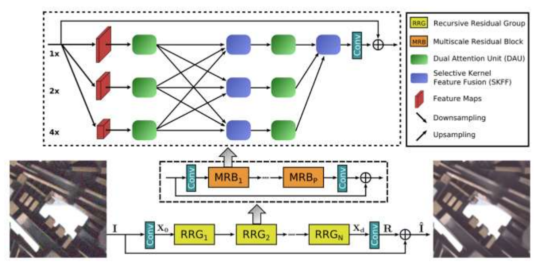


Figure 1. MIRNet Architecture

MirNet's architecture is built on recursive residual design, which is used to break down the learning process in order to minimize inputted signal information and develop deep networks. The architecture's core is located on the multi-scale residual block to give greater contextual characteristics in the parallel branches and main branch that are committed to maintaining spatially-precise high-resolution and to deliver. It can help to improve the low-resolution feature by conducting information exchange using the high-resolution feature and vice versa by applying selective kernel feature fusion (SKFF).

3.1.2 Retinex

Retinex is a long-standing concept or hypothesis that outlines how humans see an object using their Retina and Cortex [13]. Researchers are

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

still exploring this concept or hypothesis in order to improve image quality so that the image obtained by the computer is as good as human visual acuity. Images with underexposed, noisy, and foggy illumination were improved by the researchers. They accomplish this by distinguishing image reflectance from light in a given image. The equation that depicts multi-scale retinex use is shown in equation (1) below.

$$R_{MSRi} = \sum_{n=1}^{N} \omega_n R_{ni} \qquad (1)$$

In the equation R_{ni} represents the i^{th} color on the n^{th} scale, while for N represents number of scales, R_{MSRi} stand for MSR output's i^{th} spectral component , and lastly ω_n is the n^{th} scale weight.

3.1.3 Deep retinex decomposition (retinexnet)

RetinexNet is a Retinex-developed approach. The RetinexNet structure decomposes the image using Decom-Net and Enhance-Net to separate the observed image into two forms of smooth illumination: lighting-independent reflectance and structure-aware smooth illumination and changes the lighting map to provide stability over large regions while modifying local distributions using multi-scale concatenation [12]. RetinexNet architecture is illustrated in Figure 2.

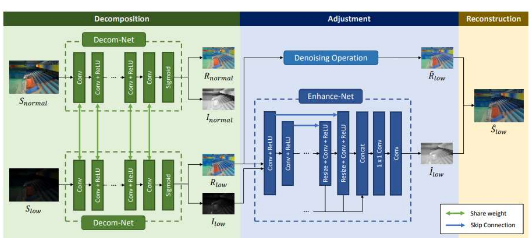


Figure 2. RetinexNet Architecture

The architecture of RetinexNet is organized into three steps: decomposition, adjustment, and reconstruction. The inputted image is separated into reflectance and illumination images in the decomposition process, which determines whether the network has low or normal illumination. While the subnetwork Enhance-Net serves as the illumination picture enhancer in the next step, a denoising operation is performed on the low reflectance image, while the low illumination image is enhanced. Finally, the reconstruction stage combines the final improved and denoised low illumination image as well as the low reflectance image.

3.1.4 Adaptive gamma correction (agc)

Adaptive Gamma Correction is an image enhancement method that enhances a low illumination image by applying a gamma correction calculated for each image depending on its gamma value, which is separated into two values: low illumination image and moderate-contrast image. If the mean intensity value is less than 0.5, the image is categorized as dark; if the mean intensity value is more than 0.5, the image is classified as bright [10]. The equation (2) used in gamma correction is shown below.

$$O = \left(\frac{I}{255}\right)^{\gamma} x\ 255 \qquad (2)$$

The symbol γ stands for gamma value, in which if $\gamma < 1$ then the dark pixel's area in the image will be brighter than the original, while the opposite of it $\gamma > 1$ makes it darker in bright pixels area, while I stand for the input value from the image [28].

3.2 Face Recognition Models

Face Recognition is utilized to recognize a person's identity through face embedding using a person's extracted face features by performing face detection before. In this study Face Recognition models that is going to be examined are both Facenet and ArcFace models.

3.2.1 Additive angular margin loss (arcface)

ArcFace is a Face Recognition Model that employs a loss function inspired by the softmax approach, with the aim of improving it due to a lack of feature embedding optimization, which causes a performance issue in deep face recognition that employs a substantial intra-class variance [5]. The whole process of the Architecture is depicted in Figure 3.

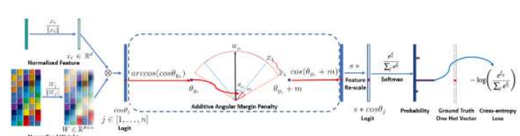


Figure 3. ArcFace Architecture

The training process of the ArcFace is divided into Normalization with W as weight and feature as x_i getting $cos\theta j$(logit) for every $W_j^T xi$. Next, we calculate $arccos\theta_{yi}$ and attaining the angle using feature x_i and ground truth W_{yi} additionally adding

Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Additive Angular Margin Penalty m which is the main solution to the ground truth angle θ_{yi}. Next, calculating the result of it using $\cos(\theta yi + m)$ and using all logits multiplying it by Feature Rescaling s, next softmax classification, and lastly cross-entropy loss.

ArcFace enforces feature embedding by increasing both the compactness and the difference in the intra-class on the hypersphere surface. At the same time, unlike FaceNet, which measures face similarity using euclidean distance, ArcFace employs geodesic distance. The ArcFace equation may be observed in equation (3) and is further explained below.

$$L_3 = -\frac{1}{N}\sum_{i=1}^{N} log \frac{e^{s\left(cos\left(\theta_{y_i}+m\right)\right)}}{e^{s\left(cos \quad y_i+m\right)\right)}+\sum_{j=1,j\neq y_i}^{n} e^{scos \quad j}} \qquad (3)$$

In ArcFace loss, the formula is transformed from the original softmax method which is x_i represents i-th deep feature sample, that belongs to y_i-th class, while θ_j represents the angle between the weight W_j and the feature x_i, s as the feature scale that help adjust the distance between class, which is the radius of the hypersphere, and lastly m as the angular margin penalty.

3.2.2 Facenet

FaceNet is a state-of-the-art Face Recognition method that employs a unified embedding from Euclidean space in accordance with face similarity, which equates to squared L2 distance and is trained using a deep convolutional neural network (DCNN). Figure 4. depicts the FaceNet Architecture.

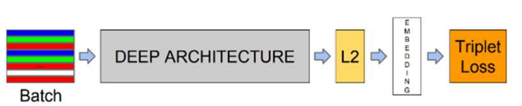


Figure 4. FaceNet Architecture

FaceNet employs Inception in its deep architecture, which was inspired by GoogleNet, and employs a 1x1 filter to decrease the channel parameters while preserving the size of the input size. L2 normalization is used to normalize the model's output before computing the loss using the triplet loss function, which distinguishes the face feature from other identities.

The triplet loss function, which is anchor, positive, and negative each with their own shared weights. Anchor as the input image of a person needs to be closer in distance with positive as dataset of the person in the image than of with the negative, which is another image of the individual person. Triplet loss equation is described as below in equation (4).

$$Loss = \sum_i^N \left[\left\|f(x_i^a) - f(x_i^p)\right\|_2^2 - \|f(x_i^a) - f(x_i^n)\|_2^2 + \alpha\right]_+ \qquad (4)$$

In Triplet loss, xi represents the image input, while $f(xi)$ represents the embedding in the image input, and α represents the positive and negative margins, which essentially used for threshold value with the minimum value of 0.5 [14].

4. PROPOSED METHOD

This study proposes a solution by experimenting with multiple combinations of state-of-the-art Low Illumination Image Enhancement methods. The Low Illumination Image Enhancement methods are used to improve Face Recognition Models and thus improve the accuracy rate in identifying individuals in images with low illumination quality from both the Essex Faces96 and Essex grimace datasets. The overview of the proposed solution is illustrated below in Figure 5.

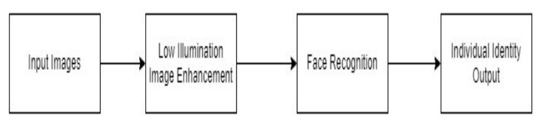


Figure 5. Overview

Retinex, RetinexNet, Adaptive Gamma Correction, and MirNet are the Low Illumination Image Enhancement methods employed in this study. Those Low Illumination Image Enhancement must undergo a training procedure that includes 485 matched images with both low and normal illumination. Following that, the trained Low Illumination Image Enhancement methods are used to improve the level of illumination of the low illumination images from the dataset, hence recovering hidden elements of the image and lowering the noise included in those images due to the low illumination.

Before conducting Face Recognition, dataset images previously processed using Low Illumination Image Enhancement methods are downsized into two dimensions: 160x160 and

112x112. Face Recognition models utilized in this study include ArcFace from a study that used Normal Illuminated datasets and FaceNet from a study that used Low Illuminated datasets. ArcFace employs dataset images with a dimension of 112x112, whereas FaceNet uses a dimension of 160x160. Both ArcFace and FaceNet use RetinaFace as the Feature detector in face detection.

This solution experiments with eight combinations of Face Recognition models and Low Illumination Image Enhancement methods, as shown in Table 1 below.

Table 1. Face Recognition + Image Enhancement Combination

No	Combination
1	ArcFace + Retinex
2	ArcFace + RetinexNet
3	ArcFace + AGC
4	ArcFace + MirNet
5	FaceNet + Retinex
6	FaceNet + RetinexNet
7	FaceNet + AGC
8	FaceNet + MirNet

5. EXPERIMENTS

This section divides experiments into three sub-sections, that consists of dataset, experimental design, and lastly experimental results.

5.1 Dataset

This study used two datasets for training, testing, and validating low illumination image enhancement methods and low illumination face recognition models: the LOL dataset [4] and a combination of both Essex Faces96 [2] and Essex Grimace [3]. LOL dataset was used to determine the best performing method at low illumination conditions, while the combination of both Essex Faces96 and Essex Grimace was used for low illumination face recognition.

5.1.1 Lol dataset

LOL dataset [4] was used to train and test our low illumination image enhancement models to validate their performance. LOL Dataset consists of 500 paired images labeled as low-illumination and normal-illumination. From the 500 paired images, 485 are utilized as a training dataset, while the remaining 15 are used for testing. Images with Low Illumination in comparison to the Normal Illuminated images, the dataset has a large number of images with a reduced quantity of information contained in the image due to a large number of noises caused by the image's low illumination. Each image in the dataset has a resolution of 400 x 600. An example of the low and Normal Illumination in the LOL dataset is shown in Figure 5.

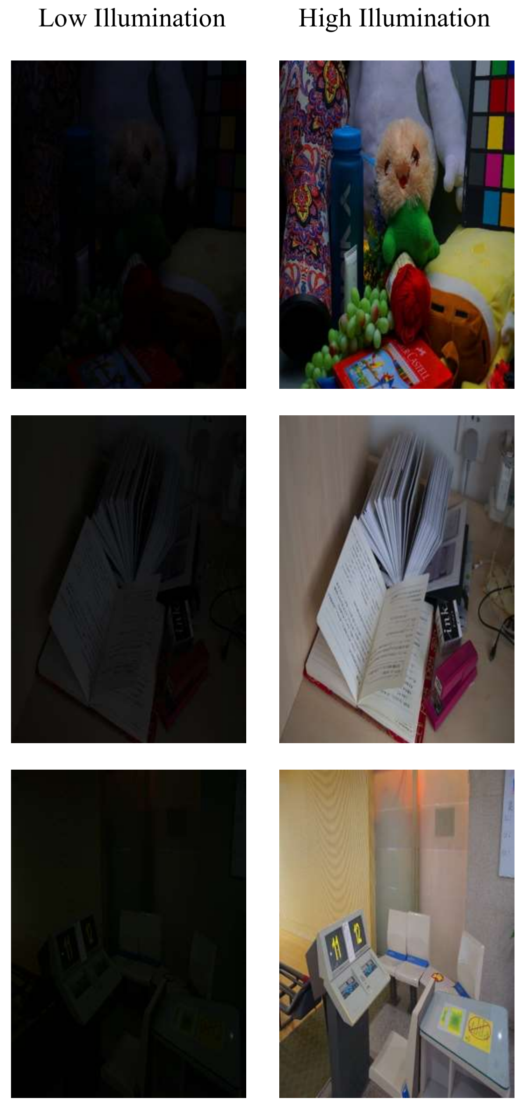


Figure 5. Pair of images in LOL dataset.

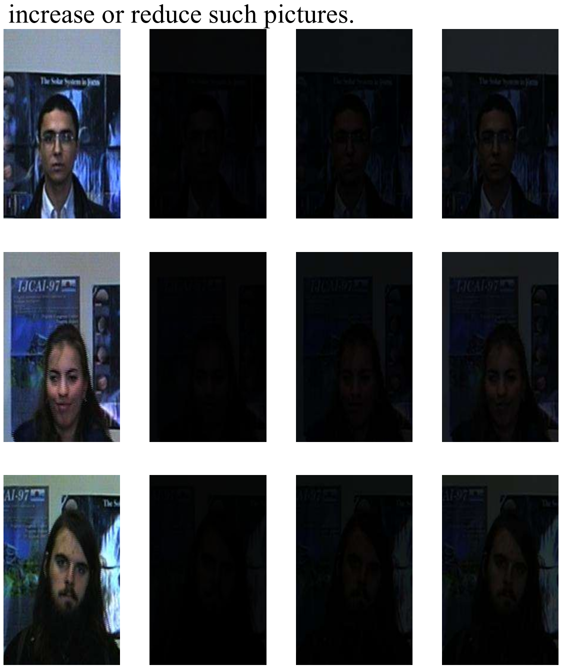
5.1.2 Essex faces96 dataset

The Essex Faces96 dataset [2] captures a sequence of 20 images in each participant using a fixed camera. Essex Faces96 has 152 individuals with a 196 × 196 image resolution. The illumination of the images in this dataset varies due to the person

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

moving, causing the image's Illumination state to change.

Due to the illumination in the Essex Faces96 dataset is not relatively low enough, image synthesis is necessary by lowering the image illumination value to a level that is difficult to perceive with the naked eye. Such image synthesis is carried out using the pillow library and the ImageEnhance module offered by the library to increase or reduce such pictures.



(a) Original | (b) 5% | (c) 10% | (d) 20%

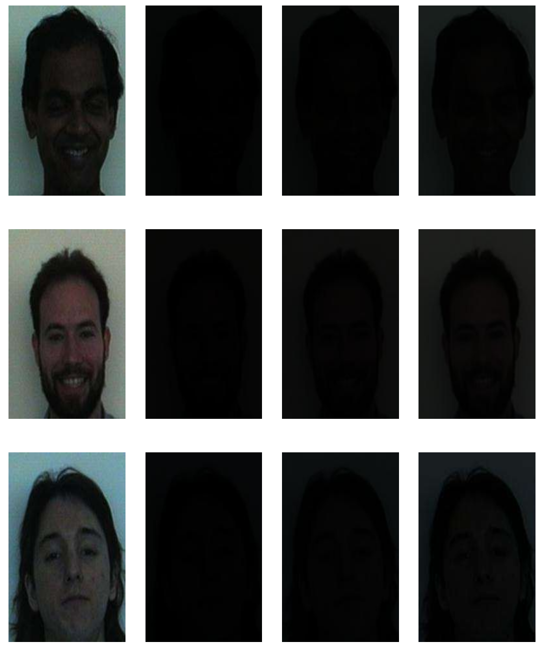
Figure 6. Essex Faces96 example (a) Essex Faces96 Original example (b) Essex Faces96 Synthesized 5% Illumination example (c) Essex Faces96 Synthesized 10% Illumination example (d) Essex Faces96 Synthesized 20% Illumination example

The image illumination in this dataset is reduced using the pillow library to an illumination of 5%, 10%, and 20% of the original image, as illustrated in Figure 3.4.

5.1.2 Essex grimace dataset

The Essex Grimace dataset [3] captures a sequence of 20 images in each individual using a fixed camera as the subject's head moves. Essex Grimace consists of 18 persons with a 180 x 200 image resolution. The image brightness in this dataset is adjusted to very low.

The same procedure is used for this dataset as it is for Essex Faces 96, where the illumination condition is insufficient, and image synthesis is performed by lowering the illumination in the images contained in the dataset to 5%, 10%, and 20% of their original images, as shown in Figure 3.4.



(a) Original | (b) 5% | (c) 10% | (d) 20%

Figure 7. Essex Faces96 example (a) Essex Faces96 Original example (b) Essex Faces96 Synthesized 5% Illumination example (c) Essex Faces96 Synthesized 10% Illumination example (d) Essex Faces96 Synthesized 20% Illumination example

5.2 Experimental Design

This study, experimental design is divided into two parts, which are experimental design for image enhancement and experimental design for face recognition.

5.2.1 Experimental design for image enhancement

The LOL dataset is utilized in the Image Enhancement experiment to perform training and testing, which is separated into 485 images for training and 15 images for testing with several Low Illumination Image Enhancement methods, including AGC, Retinex, RetinexNet, and MirNet.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Following that, we perform Low Illumination Image Enhancement by applying the trained models to several dataset combinations of both Essex Faces 96 and Essex Grimace datasets, whose illumination levels have already been reduced to several percentages, namely five, ten, and twenty percent to the original illumination level of the image by performing image synthesis using Pillow Library in Python3 programming language, and finally the original illumination level itself.

5.2.2 Experimental design for face recognition

In the Face Recognition experiment, the enhanced dataset from image enhancement is divided into training, testing, and validation in the following proportions: 60% for training, 20% for testing, and 20% for validation. Following that, 60 percent of the training is conducted on ArcFace and FaceNet Face Recognition models. Furthermore, the testing utilizes 20% of the increased dataset for both Face Recognition models. Lastly, we do a further validation utilizing the remaining 220 percent of the enhanced dataset to validate the results.

Experimentation is carried out by first extracting the face in the image using the face detection method by looping through all the dataset subfolders, obtaining the individual's name, and saving it into labels in both the training and testing datasets. After labeling both the training and testing datasets, the experiment proceeds to execute face embedding using the specified Face Recognition methods, which are ArcFace and FaceNet. The output of the face feature embedding is stored in an array. Next, the embedded array is normalized using L2 normalization, and lastly, the Multi-Layer Perceptron (MLP) Classifier is used to categorize the individual's identity.

The final result of this experiment is in an accuracy that is could rank-1 accuracy.

5.3 Experimental Result

This study, experimental result is divided into two parts, which are low illumination image enhancement results, low illumination face recognition results.

5.3.1 Low illumination image enhancement results

After conducting some quantitative experiments on the low illumination image enhancement methods, Table 2 shows that MIRNet outperforms other methods.

Table 2. Evaluation Result on Low Illumination Image Enhancement Methods..

	SSIM		
Model	5%	10%	20%
Raw	0.05	0.11	0.25
MIRNet	0.63	0.67	0.73
Retinex	0.54	0.55	0.56
Retinex Net	0.61	0.62	0.59
AGC	0.38	0.45	0.53

5.3.2 Low illumination face recognition results

Since we have completed the evaluation of low illumination image enhancement, we will proceed to test low illumination face recognition using a combination dataset of both essex faces96 and essex grimace to obtain face embedding using arcface and facenet models, which will be used to predict an individual's identity using mlp classifier.

Table 3. Evaluation Results on Face Recognition models in original illumination.

Combination	Testing Accuracy	Training Accuracy	Validation Accuracy
Original Facenet	0.95	0.99	0.99
Original ArcFace	0.94	0.99	0.93

As shown in Table 3, the facenet model is superior to the arcface in original illumination with an accuracy of 0.95.

Next, we jump to low-illumination face recognition that uses the help of image

Journal of Theoretical and Applied Information Technology
28th February 2023. Vol.101. No 4

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

enhancement methods to ease the recognition procedure.

Table 4. Evaluation Results on Low Illumination Face Recognition models in Training Accuracy.

Combination	Training Accuracy		
	5%	10%	20%
Original Facenet	0.92	0.99	0.99
Facenet+MIRNet	0.98	0.99	0.99
Facenet+RetiNexnet	0.89	0.96	0.98
Facenet+Retinex	0.95	0.99	0.99
Facenet+AGC	0.44	0.96	0.99
Original ArcFace	0.94	0.98	0.99
ArcFace+MIRNet	0.98	0.99	0.99
ArcFace+RetinexNet	0.95	0.97	0.98
ArcFace+Retinex	0.96	0.98	0.99
ArcFace+AGC	0.79	0.97	0.99

Table 4 results reveal that the training accuracy produces an outstanding percentage in most cases, with the exception of some results that is generated in five percentage, such as Facenet+AGC and ArcFace+AGC.

Table 5. Evaluation Results on Low Illumination Face Recognition models in Validation Accuracy.

Combination	Validation Accuracy		
	5%	10%	20%
Original Facenet	0.66	0.89	0.96
Facenet+MIRNet	0.80	0.94	0.99
Facenet+RetiNexnet	0.34	0.63	0.78
Facenet+Retinex	0.50	0.86	0.97
Facenet+AGC	0.38	0.60	0.91
Original ArcFace	0.75	0.79	0.89
ArcFace+MIRNet	0.86	0.91	0.93
ArcFace+RetinexNet	0.52	0.65	0.79
ArcFace+Retinex	0.56	0.74	0.89
ArcFace+AGC	0.50	0.71	0.85

Table 5 demonstrates that in higher illumination percentages, validation accuracy performance is comparable to training accuracy performance, but in lower illumination percentages, the gap is significantly greater.

Table 6. Evaluation Results on Low Illumination Face Recognition models in Testing Accuracy.

Combination	Testing Accuracy		
	5%	10%	20%
Original Facenet	0.66	0.82	0.94
Facenet+MIRNet	0.78	0.94	0.96
Facenet+RetiNexnet	0.32	0.62	0.79
Facenet+Retinex	0.49	0.82	0.94
Facenet+AGC	0.22	0.52	0.81
Original ArcFace	0.68	0.78	0.89
ArcFace+MIRNet	0.84	0.91	0.92
ArcFace+RetinexNet	0.48	0.64	0.79
ArcFace+Retinex	0.49	0.74	0.89
ArcFace+AGC	0.44	0.63	0.83

Table 6 result shows that in twenty percentage illumination due to the original relatively high accuracy, the only useable combination of method and model are both FaceNet+MIRNet and ArcFace+MIRNet, which resulted in a higher accuracy percentage of 0.96 and 0.92, respectively.

While, in ten percentage illumination result shows that due to the original accuracy resulting in a good performance, the only useable combination of method and model are both FaceNet+MIRNet and ArcFace+MIRNet, which resulted in a better accuracy percentage in 0.94 and 0.91, respectively.

Lastly, in five percentage illumination result shows that due to the original accuracy resulting in a mediocre performance, the only useable combination of method and model are both FaceNet+MIRNet and ArcFace+MIRNet, which resulted in a far better accuracy percentage in 0.78 and 0.84, respectively.

To further assist in visualizing the results of low illumination face recognition, Figure 8. Is generated

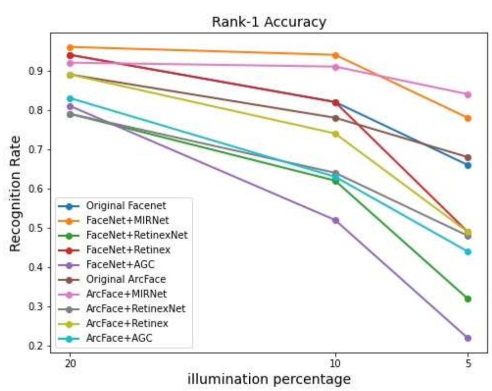


Figure 8. Rank-1 Accuracy on Low Illumination Face Recognition percentages.

5.4 Discussions

From our experiments, we can see that the facenet model in 10 and 20 percentages of illumination perform better than arcface, while in lower percentages in 05 percentage of illumination, arcface perform better than facenet. Furtherly the only combination that is useable is facenet with mirnet and arcface with mirnet. These experimentation output correlate with the performance in [23], which describes ArcFace to be proven more useable in low illumination images in 78% accuracy and FaceNet in 18% accuracy.

6. CONCLUSION AND FUTURE WORK

In this study, we explored the effectiveness of four methods for enhancing low-illumination images MIRNet, RetinexNet, Retinex, and AGC when combined with two face recognition models ArcFace and Facenet. To assess the performance of the combined models, we used Rank-1 Accuracy as the evaluation metric. Results showed that MIRNet consistently performed the best according to SSIM. In terms of the best-performing combination for each illumination percentage, we found that at 20% illumination, the combination of Facenet and MIRNet had the highest Rank-1 Accuracy of 0.96. At 10% illumination, the combination of Facenet and MIRNet had the highest Rank-1 Accuracy of 0.94. At 5% illumination, the combination of ArcFace and MIRNet had the highest Rank-1 Accuracy of 0.84.

These findings suggest that the use of MIRNet Image Enhancement method can improve the performance of face recognition models in low-illumination conditions, and that the ArcFace model performs better as the illumination level decreases, while the Facenet model's performance decreases significantly. Other than that, Results in both original ArcFace and Facenet models shows that there is a big impact of performance in Face Recognition Accuracy in different low illumination levels with at least a 10% reduction in performance.

In future work, it would be beneficial to investigate the use of a dataset that captures true low-illumination conditions rather than artificially synthesizing low illumination on existing datasets, as well as conduct research using cutting-edge image enhancement methods rather than outdated image enhancement methods, excluding the MIRNet image enhancement method used in this study.

REFERENCES

[1] Valentino-devries J. How the police use facial recognition, and where it falls short [Internet]. The New York Times. The New York Times; 2020 [cited 2022Nov6]. Available from: https://www.nytimes.com/2020/01/12/technology/facial-recognition-police.html
[2] Spacek L. [Internet]. Face recognition data. [cited 2022Nov6]. Available from: https://cmp.felk.cvut.cz/~spacelib/faces/faces96.html

[3] Spacek L. [Internet]. Face recognition data. [cited 2022Nov6]. Available from: https://cmp.felk.cvut.cz/~spacelib/faces/faces96.html
[4] BMVC2018 deep Retinex decomposition. [cited 2022Nov6]. Available from: https://daooshee.github.io/BMVC2018website/
[5] Deng J, Guo J, Xue N, Zafeiriou S. Arcface: Additive angular margin loss for deep face recognition. InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2019 (pp. 4690-4699).
[6] Meng Q, Zhao S, Huang Z, Zhou F. Magface: A universal representation for face recognition and quality assessment. InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 2021 (pp. 14225-14234).
[7] Rahman MF, Sthevanie F, Ramadhani KN. Face recognition in low lighting conditions using fisherface method and clahe techniques. In2020 8th International Conference on Information and Communication Technology (ICoICT) 2020 Jun 24 (pp. 1-6). IEEE.
[8] Huang H, Tao H, Wang H. A convolutional neural network based method for low-illumination image enhancement. InProceedings of the 2nd International Conference on Artificial Intelligence and Pattern Recognition 2019 Aug 16 (pp. 72-77).
[9] Veluchamy M, Subramani B. Image contrast and color enhancement using adaptive gamma correction and histogram equalization. Optik. 2019 Apr 1;183:329-37.
[10] Dang TV. Smart home Management System with Face Recognition Based on ArcFace Model in Deep Convolutional Neural Network. Journal of Robotics and Control (JRC). 2022 Dec 4;3(6):754-61.
[11] Asmara RA, Sayudha B, Mentari M, Budiman RP, Handayani AN, Ridwan M, Arhandi PP. Face Recognition Using ArcFace and FaceNet in Google Cloud Platform For Attendance System Mobile Application. In2022 Annual Technology, Applied Science and Engineering Conference (ATASEC 2022) 2022 Dec 29 (pp. 134-144). Atlantis Press.
[12] Cao G, Huang L, Tian H, Huang X, Wang Y, Zhi R. Contrast enhancement of brightness-distorted images by improved adaptive gamma correction. Computers & Electrical Engineering. 2018 Feb 1;66:569-82.
[13] Parthasarathy S, Sankaran P. An automated multi scale retinex with color restoration for image enhancement. In2012 National Conference on Communications (NCC) 2012 Feb 3 (pp. 1-5). IEEE.
[14] Wei C, Wang W, Yang W, Liu J. Deep retinex decomposition for low-light enhancement. arXiv preprint arXiv:1808.04560. 2018 Aug 14.
[15] Zamir SW, Arora A, Khan S, Hayat M, Khan FS, Yang MH, Shao L. Learning enriched features for real image restoration and enhancement. InEuropean Conference on Computer Vision 2020 Aug 23 (pp. 492-511). Springer, Cham.
[16] William I, Rachmawanto EH, Santoso HA, Sari CA. Face recognition using facenet (survey, performance test, and comparison). In2019 fourth international conference on informatics and computing (ICIC) 2019 Oct 16 (pp. 1-6). IEEE.
[17] Yang J, Xu Y, Yue H, Jiang Z, Li K. Low-light image enhancement based on Retinex decomposition and adaptive gamma correction. IET Image Processing. 2021 Apr;15(5):1189-202.
[18] Zhuang L, Guan Y. Image enhancement by deep learning network based on derived image and Retinex. In2019 IEEE 3rd Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) 2019 Oct 11 (pp. 1670-1673). IEEE.
[19] Ma H, Lv W, Li Y, Liu Y. Image Adaptive Contrast Enhancement for Low-illumination Lane Lines Based on Improved Retinex and Guided Filter. Applied Artificial Intelligence. 2021 Dec 15;35(15):1970-89.
[20] Tang S, Li C. Low illumination image enhancement based on image fusion. In2020 6th International Conference on Robotics and Artificial Intelligence 2020 Nov 20 (pp. 74-78).
[21] Zamir SW, Arora A, Khan SH, Munawar H, Khan FS, Yang MH, Shao L. Learning Enriched Features for Fast Image Restoration and Enhancement. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022 Apr 13.
[22] Wang H, Wang Y, Zhou Z, Ji X, Gong D, Zhou J, Li Z, Liu W. Cosface: Large margin cosine loss for deep face recognition. InProceedings of the IEEE conference on computer vision and pattern recognition 2018 (pp. 5265-5274).
[23] Huang Y, Wang Y, Tai Y, Liu X, Shen P, Li S, Li J, Huang F. Curricularface: adaptive curriculum learning loss for deep face

recognition. Inproceedings of the IEEE/CVF conference on computer vision and pattern recognition 2020 (pp. 5901-5910).

[24] Huang YH, Chen HH. Face recognition under low illumination via deep feature reconstruction network. In2020 IEEE International Conference on Image Processing (ICIP) 2020 Oct 25 (pp. 2161-2165). IEEE.

[25] Agarwal S, Sikchi HS, Rooj S, Bhattacharya S, Routray A. Illumination-invariant face recognition by fusing thermal and visual images via gradient transfer. InScience and Information Conference 2019 Apr 25 (pp. 658-670). Springer, Cham.

[26] Ren D, Ma H, Sun L, Yan T. A novel approach of low-light image used for face recognition. In2015 4th International Conference on Computer Science and Network Technology (ICCSNT) 2015 Dec 19 (Vol. 1, pp. 790-793). IEEE.

[27] Kang Y, Pan W. A novel approach of low-light image denoising for face recognition. Advances in Mechanical Engineering. 2014 Apr 14;6:256790.