Image Enhancement using Convolutional Neural Network for Low Light Face Detection

Antonius Filian Beato Istianto[1], Gede Putra Kusuma[2]
[1]Computer Science Department, Binus Graduate Program – Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia
[2]Computer Science Department, Binus Graduate Program – Master of Computer Science, Bina Nusantara University, Jakarta, Indonesia
E-mail: [1]antonius.istianto@binus.ac.id, [2]inegara@binus.edu

Abstract

This research aims to combine the study of face detection with improvement of image quality in low-light conditions. In this research, we introduce a method that combines Convolutional Neural Networks for image processing to enhance face detection performance in low-light conditions. The proposed method involves pre-processing the images using three image enhancement methods: Deep Lightening Network, Deep Retinex Net, and Signal-to-Noise Ratio Aware. Each of these methods is combined with the face detection method, RetinaFace. The experiment is evaluated using the DARKFACE Dataset, and the performance of each combination is assessed using Average Precision (AP). The combination that yields the best AP value will be determined as the best approach for low-light face detection. The best combination, which utilizes Signal to Noise Ratio Aware for image enhancement and RetinaFace for face detection, achieves an AP score of 52.92%. This result surpasses the face detection performance using the original images from the DARKFACE Dataset, which scored 7.12% in AP. Thus, this experiment demonstrates that image enhancement using Convolutional Neural Networks can significantly improve face detection in low-light conditions.

Keywords: *Image Enhancement, Face Detection, Low Light Condition, Convolutional Neural Network, Deep Learning*

1. Introduction

Face detection has become a popular topic in machine learning research. Faces have been used as a means of human identification for criminal and identification purposes since 1850. This was marked by the introduction of standard mug-shot photos proposed by Alphonse Bertillon for the police and the use of passports during World War I. Research on face detection has continued, and in 1964, researcher Woodrow Wilson Bledsoe developed man-machine technology related to face detection and recognition. Subsequently, other neural network-based approaches emerged in the 1980s. Soon after, Principal Component Analysis (PCA)-based systems were introduced by Kirby and Sirovich and have since evolved.

Face detection has rapidly advanced due to its widespread use in everyday life. It serves as an initial step for face recognition, estimating crowd density, focusing cameras, and more. The concept of face detection involves identifying and locating human faces in images, regardless of their size, position, or condition. According to Ferik et al., face detection involves distinguishing between two labels: face and non-face. However, many assume that available face data share the same size, background, and conditions, whereas in reality, there are various conditions and sizes. For instance, many images feature minimal lighting conditions, unlike typical normal photos.

Face detection has seen many advancements due to the existence of specific conditions that affect the quality of detected face images. Factors such as noise, low image quality, and poor lighting can hinder face detection performance. This necessitates image

ISSN: 2720-992X

enhancement processes to improve the image quality. Image enhancement techniques include sharpening, lighting adjustment, and noise reduction. This research focuses on developing image enhancement techniques specifically for low-light images.

Previous studies have explored image enhancement in low-light conditions. Wang et al. proposed a method called Deep Lightening Network [1], which was evaluated using SSIM, PSNR, and NIQE, with respective scores of 80.7, 21.94, and 3.65. Wei et al. introduced the Deep Retinex Net method [2], but its evaluation was not measured. Xu et al. proposed the Signal to Noise Ratio Aware method [3], which was evaluated using SSIM and PSNR, with scores of 0.842 and 24.61, respectively.

Several previous studies have also been conducted on face detection. Deng et al. introduced the RetinaFace method, which was evaluated using AP and achieved AP scores of 88.67% for the easy dataset, 87.09% for the medium dataset, and 80.99% for the hard dataset [4].

Based on the discussions of previous studies, several research works have been conducted to improve image quality in low-light conditions and face detection. These previous studies will be chosen and implemented in this research as they can be easily translated into code. This research has the potential to enhance existing solutions by building upon the advancements made in previous studies.

2. Related Works

2.1. Related Works Related to Low Light Image Enhancement

Several studies on image quality improvement have been conducted before. Zhang et al. proposed a deep learning-based method that simultaneously enhances image contrast and reduces noise [5]. The method consists of two sub-networks: Image Contrast Enhancement Network (ICE-Net) and Re-Enhancement and Denoising Network (RED-Net) [5]. The evaluation of this research utilized SSIM, PSNR, and NIQE, with respective scores ranging from 76 to 4.01.

Wang et al. proposed a method called Deep Lightening Network (DLN) to improve image quality [1]. DLN is an extension of Convolutional Neural Network (CNN) [1]. The evaluation of this research used SSIM, PSNR, and NIQE, with scores of 80.7, 21.94, and 3.65, respectively.

Patil et al. proposed a method for image quality improvement using dehazing, followed by histogram equalization (HE), and concluded with denoising [6]. Dehazing is used to eliminate haze, histogram equalization enhances contrast, and denoising eliminates noise in the image. The results of this research were not evaluated.

Park et al. proposed image quality improvement using a modified Retinex Model with adaptive L2 spatial [7]. Additionally, this research simplified Retinex using FFT methods and quantized weight values for consumer devices [7]. The evaluation of this research used SSIM and PSNR, with scores ranging from 0.79 to 0.83 and 19.24 to 20.78, respectively.

Guo et al. proposed a method called Regularizer Illumination Optimization combined with Deep Noise Suppression [8]. This research used three images as the dataset and was evaluated using NIQE and BMTQI, with scores of 2.78 and 3.03, respectively.

Guo et al. proposed a method for image quality improvement using Illumination Map Estimation [9]. The dataset used in this research was the HD Dataset, and it was evaluated using LOE, with a score of 2.39.

Ren et al. proposed the Trainable Hybrid Network method for image quality improvement [10]. The Trainable Hybrid Network contains of two streams simultaneously trained on global content in the image. The dataset used in this research was the MIT-Adobe FiveK Dataset. It was evaluated using SSIM and PSNR, with scores of 0.96 and 28.43, respectively.

Shen et al. proposed a CNN method called MSR-Net combined with Retinex to enhance image quality [11]. This research used the MEF Dataset, NPE Dataset, and VV Dataset. It was evaluated using SSIM and NIQE, with scores of 0.92 and 3.46, respectively.

Singhal et al. proposed the modified LRD-Net method to improve image quality in low-light conditions [12]. This research used the LOL Dataset and evaluated it using SSIM, PSNR, and NIQE, with scores of 0.78, 16.02, and 5.97, respectively.

Guo et al. proposed a novel method called Zero-Reference Deep Curve Estimation (Zero-DCE) to enhance image quality in low-light conditions [13]. Several datasets were used, including the SICE Dataset, DARK FACE Dataset, and WIDER Dataset. The evaluation included PSNR, SSIM, and MAE, with scores of 16.57, 0.59, and 98.78, respectively.

Wei et al. proposed the Deep Retinex Net method to improve image quality in low-light conditions [2]. This research used the Low of Light (LOL) Dataset. The evaluation of this research was not measured.

Xu et al. proposed the Signal to Noise Ratio Aware method to improve image quality in low-light conditions [3]. It was evaluated using SSIM and PSNR, with scores of 0.842 and 24.61, respectively.

2.2. Related Works Related to Face Detection

Meanwhile, in face detection, research is also continuously being conducted to improve evaluation. Nath et al. proposed face detection using Histogram of Oriented Gradient (HOG) [14]. This research was not evaluated, and the conclusion was that HOG provides better and more accurate results compared to other learning methods such as Haar Cascade, although it requires longer processing time.

Vimal and Chamandeep conducted research to test various methods for face detection [15]. The algorithms tested in this research include Principal Component Analysis (PCA) and Skin Color Modeling (SCM), Haar-Like Feature, High-Level Language Face Detection, and Facial Features Face Detection. This research was not evaluated, and the conclusion was that Haar-Like Feature Extraction is the best method for face detection and recognition [15].

Jadhav et al. conducted research to test various methods for face detection [16]. The algorithms tested in this research include Cascade Classifier, Dlib CNN, Dlib HOG, and MTCNN [16]. This research was not evaluated, and the conclusion was that MTCNN is the most efficient algorithm and yields the highest accuracy.

Shamrat et al. proposed a method combining Convolutional Neural Network with Max Pooling for face detection in images [17]. This research used the LFW dataset consisting of 13,000 photos [17]. The evaluation of this research used accuracy and achieved scores of 95.72% for training accuracy and 96.27% for validation accuracy.

Jiang, Huaizu & Miller, Erik proposed a new method called Faster R-CNN, which is an improvement of R-CNN [18]. The dataset used in this research was the WIDER Face dataset. This research was not quantitatively evaluated. The research concluded that the effectiveness of face detection comes from the Region Proposal Network (RPN) because the convolutional layer between the RPN and Faster R-CNN models does not incur additional computational burden.

Li et al. conducted research on face detection using the Convolutional Neural Network Cascade [19]. This research used the Wild Dataset, Annotated Facial Landmarks in the Wild (AFLW) Dataset, and Annotated Faces in the Wild. The evaluation of this research used AP with a score of 87.48.

Obaida et al. conducted research to compare face detection methods using Viola Jones and YOLO v3 [20]. In this research, a private dataset was used, and the conclusion was that YOLO v3 achieved better accuracy than Viola Jones, although YOLO v3 required slower processing time compared to Viola Jones [20]. The accuracy obtained using the YOLO v3 method was 98%, while the accuracy obtained using Viola Jones was 86%.

Li et al. proposed the YOLO v3 method with improvements for face detection [21]. The improvements made to YOLO v3 included using a layer to detect small faces, selecting SoftMax as the loss function, and reducing the dimension of feature layers in the detection process accelerate the process [21]. This research used the WIDER FACE Dataset, CelebA Dataset, and FDDB Dataset. The evaluation of this research used accuracy with a score of 93.9%.

Li et al. proposed a method called DBCFace, which is a modified CNN with NMS [22]. This research used datasets such as AFW, PASCAL Face, FDDB, and WIDER FACE. The research was measured using Average Precision (AP). The AP scores for each dataset were as follows: AFW Dataset: 99.87%, PASCAL Face: 99.23%, FDDB: 98.7%, and WIDER FACE: 90.34%.

Zhang et al. introduced a technique named RefineFace, comprising five components: Selective Two-step Classification (STC), Selective Two-step Regression (STR), Feature Supervision Module (FSM), Scale-aware Margin Loss (SML), and Receptive Field Enhancement (RFE) [23]. This research used datasets such as WIDER FACE, AFW, PASCAL FACE, MAFA, and FDDB. The research was measured using AP and Recall Rate. Using the WIDER FACE dataset, the research achieved AP scores of 96.6%, 95.8%, and 91.4% for easy, medium, and hard datasets, respectively. On the AFW dataset, it achieved an AP score of 99.9%. On the PASCAL Face dataset, the research obtained a score of 99.45%. On the FDDB dataset, it achieved an AP score of 99.11%, and on the MAFA dataset, it obtained an AP score of 83.9%.

Deng et al. proposed a method called RetinaFace [4]. This method was evaluated using AP and achieved AP scores of 88.67% for the easy dataset, 87.09% for the medium dataset, and 80.99% for the hard dataset.

2.3. Related Works Related to Low Light Face Detection

Several studies have been conducted on low light face detection. Yu et al. proposed Single-Stage Low Light Face Detection as a face detection method [24]. The study utilized Multi-scale Retinex with Color Restoration for image enhancement and employed the PyramidBox algorithm for face detection. The DARKFACE dataset was used for evaluation. The study measured the Mean Average Precision (mAP) for image enhancement, which resulted in a score of 76.2. The face detection performance was also evaluated using mAP, achieving a score of 82.3.

Liang et al. proposed the Recurrent Exposure Generation (REG) module and paired it with the Multi-Exposure Detection (MED) module [25]. The study used the low light DARK FACE dataset and measured the mAP, obtaining a score of 77.69.

Wang et al. proposed the Joint-High Low Adaption (HLA) Framework [26]. HLA was used to enhance the image quality before performing face detection. The DARK FACE dataset was utilized, and the evaluation was conducted using mAP, resulting in a final score of 44.4.

Theoffilus et al. proposed a combination of Retinex and RetinaFace for low light face detection [27]. Retinex was used for image enhancement, while RetinaFace was employed for face detection. The LOL dataset was utilized, and the final mAP score in this study was 43.

Wang et al. proposed a combination of the MSRCR method for image enhancement and the Cascade R-CNN method with DetectoRS for face detection [28]. The DARKFACE dataset was used, and the evaluation was conducted using mAP. The study achieved a mAP score of 84.3.

Asadzadeh, Mehdi, and Rikhtegar proposed a combination of non-local means, non-local adaptive means, and Retinex for image enhancement, and PCA for face detection [29]. The study utilized the dataset from Extended Yale B. The evaluation focused on detection speed, achieving a time of 2.55ms, with a maximum detection rate of 97.5% in the 360-dimensional PCA space.

3. Reseach and Methodology

3.1. Image Enhancement

Image enhancement is one way to improve image quality. According to Nur, image enhancement is method, technique, or operation to provide more details in an image [30]. The process of image quality enhancement aims to deliver a better image that provides relevant information according to the purpose or interest of the image enhancement itself. According to Nur, the process of image enhancement can be demonstrated in Figure 1.

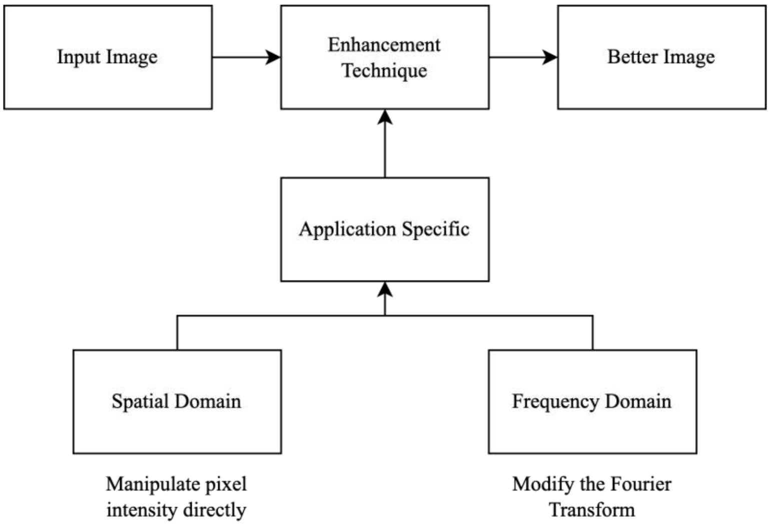

Figure 1. Image enhancement Process

Generally, the domain of image enhancement can be divided into two types: spatial and frequency domains. The spatial domain involves manipulating the pixel values directly, influenced by the spatial values of other pixels. On the other hand, the frequency domain involves manipulating the values of the frequency spectrum of the image.

In spatial domain, image enhancement performed by finding the average value, median, or other statistical measures for each grayscale or RGB level. The spatial domain is suitable for enhancing image quality with relatively low irregularity.

Broadly speaking, the spatial domain can be divided into two types: point processing and mask processing. In point processing, image enhancement focuses on individual pixels, meaning it does not involve neighboring pixels. The masking size in point processing is 1x1. On the other hand, in mask processing, we process a neighborhood window within an image. The next step in mask processing is the application of a mask to that window, which is also known as convolving the mask with the pixel window.

In the frequency domain, image quality enhancement is achieved using Fourier Transformation, where the matrix values of the image, previously in spatial domain, are converted to the frequency domain. In the frequency domain, due to the use of Fourier Transformation, rotation, scaling, or shifting can be applied, providing more accurate information in the image compared to the spatial domain. Image quality enhancement is commonly performed to improve images affected by noise, blur, or poor lighting conditions [31].

3.2. Deep Lightening Network

The Deep Lightening Network algorithm is a new algorithm proposed by Li Wen Wang in 2020 [1]. The architecture of the Deep Lightening Network can be seen in Figure 2.

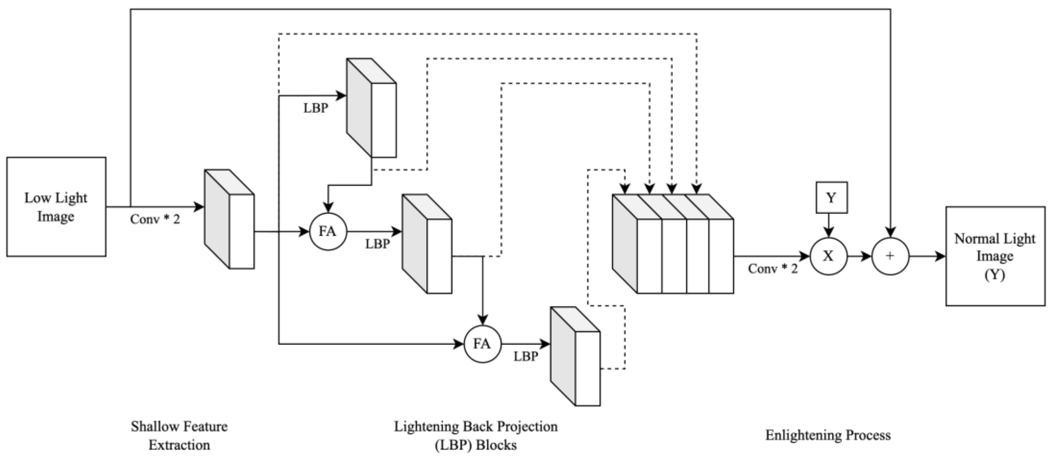

Figure 2. Architecture of Deep Lightening Network

There is an improvement with the use of this Deep Lightening Network. The updates made by Li Wen Wang in this research include: Interactive Low Light Enhancement, Deep Lightening Network, Lightening Back Projection (LBP), and Feature Aggregation. Using Interactive Low Light Enhancement, this research addresses low-light enhancement by utilizing a residual learning model between low-light and normal-light images. This research proposes a new Deep Lightening Network approach based on the above-mentioned residual model to enhance quality of images with low light conditions. The Deep Lightening Network algorithm utilizes several brightening blocks, as seen in Figure 2. Deep Lightening Network has been compared with other similar algorithms for image enhancement, particularly for images with low light conditions, and the results indicate that Deep Lightening Network outperforms all other methods in both subjective and objective measures.

This research introduces the concept of enhancing low-light images through the Lightening Back Projection block. This block iteratively adjusts the brightness and darkness of low-light images to learn the residual enhancement. Notably, this is the first study to incorporate a novel back-projection structure specifically for low-light enhancement. Additionally, the proposed Feature Aggregation block combines the outcomes of diverse brightening procedures, resulting in more informative features that can be used in subsequent brightening processes.

3.3. Deep Retinex Net

The Deep Retinex Net algorithm is an algorithm proposed by Wei et al. in 2018 [2]. The implementation of the Deep Retinex Net algorithm to improve image quality in low-light conditions can be seen in Figure 3.

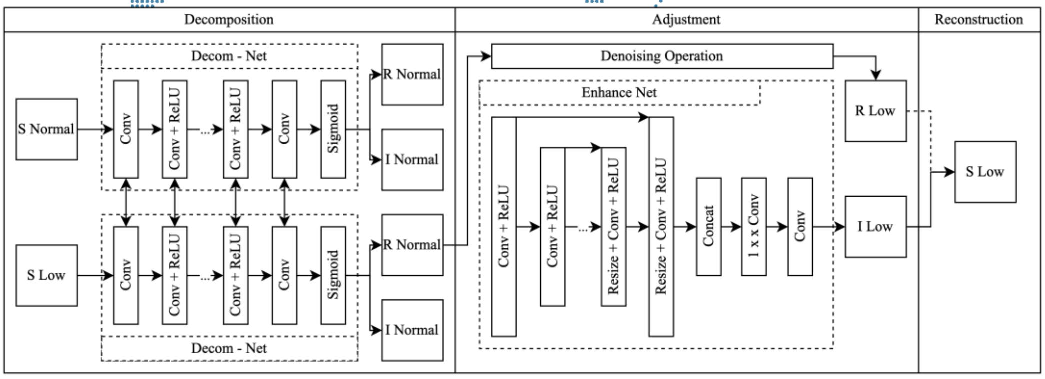

Figure 3. Architecture of Deep Retinex Net

The updates made in this research include the introduction of the Deep Retinex Net algorithm, which is a neural network approach in image enhancement to improve the light quality of images captured in low-light conditions. This method aims to address issues such as low contrast, poor lighting quality, and loss of details commonly caused by low-light conditions. The process of Deep Retinex Net works similarly to the human retina, decomposing the image into three components: reflectance, natural illumination, and dark component. The reflectance represents important information about the texture and object details in the image, while the natural illumination represents a smoother lighting component with a smooth spatial distribution. The dark component represents the image with unwanted noise or artifacts.

The steps involved in the Deep Retinex Net process are as follows: Decomposition, Enhancement, Fusion, and Reconstruction. In Decomposition step, the original image is divided into three components: reflectance, natural illumination, and dark component. This can be achieved using a specially trained deep network for this purpose. In Enhancement step, the separated reflectance from the previous step is used to enhance the original image. This process involves noise removal, image enhancement, and the retrieval of lost image details. In Fusion step, the enhanced image is combined with the separated natural illumination from the first step. The goal is to maintain the original lighting information and readjust the lighting in the enhanced image. In Reconstruction step, the enhanced images are recombined to produce the final image with improved quality from low-light conditions.

3.4. Signal to Noise Ratio Aware

The Signal to Noise Ratio Aware algorithm is an algorithm proposed by Xu et al. in 2022 [3]. The implementation of the Signal to Noise Ratio Aware algorithm to improve image quality in low-light conditions can be seen in Figure 4.

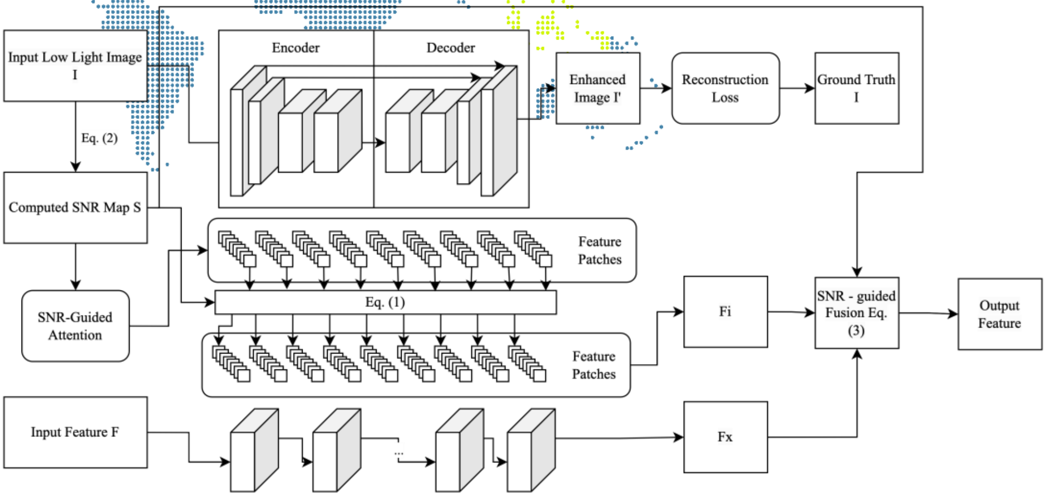

Figure 4. Architecture of Signal to Noise Ratio Aware

The update made in this research conducted by Xu et al. in 2022 is the introduction of the Signal to Noise Ratio Aware method to enhance low light images captured. This method aims to improve the image quality in low-light conditions by considering the Signal to Noise Ratio as an important factor in the image enhancement process. The Signal to Noise Ratio Aware method aims to enhance image quality under low-light conditions by following a specific working principle. Initially, the method involves measuring the Signal to Noise Ratio (SNR) of the original image, which represents the ratio between the signal power and noise power in the image, offering insights into its noise level.

Subsequently, the original image undergoes enhancement while considering the previously measured SNR value. This enhancement process incorporates noise reduction, contrast enhancement, and detail restoration, specifically tailored to address low-light conditions. A range of techniques, including adaptive contrast enhancement and histogram-based adaptive contrast enhancement, are employed to achieve these improvements.

Once the image is enhanced, the next crucial step is to recover the details that may have been lost due to the challenging lighting conditions. This detail recovery phase may involve edge restoration and texture recovery to regain the fine nuances in the image.

Finally, the results obtained from both the enhancement and detail recovery stages are skillfully fused together to generate the final enhanced image. As a result, the merged output exhibits superior light quality, enhanced contrast, and improved details, effectively addressing the challenges posed by low-light conditions.

3.5. Face Detection using RetinaFace

The updates made by Deng et al. (2019) introduce a method called RetinaFace for accurate, efficient face detection in various (wild) conditions [4]. This method is designed to overcome the challenges in face detection, such as variations in pose, face size, diverse lighting conditions, and facial expression variations. The modeling illustration in RetinaFace can be seen in Figure 5.

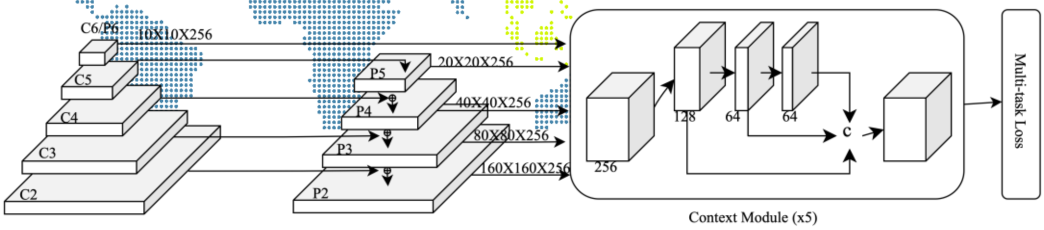

Figure 5. Architecture of RetinaFace

RetinaFace, a sophisticated face detection method, follows a series of stages to accurately identify human faces in images. Firstly, the method employs a Convolutional Neural Network architecture to detect relevant features in the input image. This network, consisting of deep convolutional layers, excels at learning robust feature representations of human faces.

Next, RetinaFace adopts a single-stage-dense detection approach, directly analyzing each pixel in the image to determine whether it corresponds to a face or not. By eliminating separate proportional steps, this approach enhances detection accuracy and efficiency.

To handle variations in human face sizes, RetinaFace employs a multi-scale and multi-level strategy. By utilizing image pyramids and convolutional layers with different perspectives, the method effectively detects faces with varying scales and resolutions.

After the initial detection, RetinaFace refines and filters the bounding boxes surrounding the detected faces. Refinement techniques, such as bbox regression, and filtering techniques, like bbox classification, are applied to enhance detection precision and reduce false positives.

Moreover, in addition to face detection, RetinaFace can also detect facial landmarks, such as eyes, nose, and mouth. This feature allows for more comprehensive face recognition and enables various applications, including face verification tasks.

3.6. Dataset

This research uses two datasets named DARKFACE Dataset and LOL Dataset. The LOL Dataset, or Low Light Dataset, is a dataset specifically designed to train models to improve image quality under low-light conditions. This dataset is suitable for developing algorithms to enhance image quality in low-light conditions.

The LOL Dataset consists of pairs of images: the original image with low lighting and the same reference image with good lighting. The LOL Dataset consists of 500 pairs of dark and bright images. The dataset is divided into 485 images for training and 15 images for testing. In the LOL Dataset, the reference image with good lighting is used to train the model. Performance measurement can be done by calculating the SSIM or PSNR by comparing the two images (the original image with the reference image). Figure 6 show the LOL Dataset sample.

Figure 6. LOL Dataset Sample

KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)
Terakreditasi Nomor 204/E/KPT/2022 | Vol. 5, No. 1, Januari (2024), pp. 71-85

The DARKFACE Dataset consists of 6000 low light images captured. These images were taken in educational buildings, streets, bridges, parks, and other locations. The six thousand images have been labeled with bounding boxes for human faces as the main training or validation set. In addition, the DARKFACE Dataset also provides 9000 unlabeled images taken from the same locations and settings. The DARKFACE Dataset also provides a unique set of 789 pairs of low-light images that do not contain faces. These can be used as training variations if needed. For testing, the dataset contains 4000 usable images. Figure 7 show the DARKFACE Dataset sample.

Figure 7. DARKFACE Dataset Sample

3.7. Image Enhancement Method

In this research, image quality enhancement experiments are conducted. The image enhancement methods used in this research are Deep Lightening Network, Deep Retinex Net, and Signal to Noise Ratio Aware. Explanation for each method can be found in Section 3, "Theory and Methods."

To perform training and validation, the LOL dataset is used. This dataset is divided into two parts, one for training and the other for testing. Additionally, hyperparameter tuning is carried out for each experiment and tested method. The parameters subjected to hyperparameter tuning are the optimizer and learning rate.

The best combinations of hyperparameter tuning for each method will result in a model. These models, derived from each method, will be utilized on the DARKFACE dataset to generate bright image outputs that can be used for training and testing face detection.

3.8. Face Detection

After the image quality enhancement process is completed, the resulting enhanced images are used for training and testing face detection. The face detection method used is RetinaFace. Explanation for RetinaFace method can be found in Section 3, “Theory and Methods”.

To achieve the best performance, hyperparameter tuning is also conducted during the face detection training. The parameters subjected to tuning are the optimizer and learning rate. By fine-tuning these parameters, the face detection model can be optimized for optimal performance.

3.9. Experimental Design

The experimental design in this research is divided into three stages: training, validation, and testing, with respective proportions of 60:20:20. This means that the 60% of the data will be used for training, the 20% for validation, and the remaining the 20% for testing.

The training phase, also known as the learning phase, involves training the designed model to obtain the appropriate parameters. Proper parameters will lead to better evaluation results. This process requires labeled data.

During the training phase, hyperparameter tuning is conducted to optimize the model and achieve the best evaluation results. Hyperparameter tuning involves adjusting the model's parameters to significantly impact its performance, as emphasized by Minarno et al. in the context of machine learning and deep learning. In this study, hyperparameter tuning is applied to determine the optimizer parameter, dropout after the pooling layer, dropout in the fully connected layer, and dense layer.

The validation stage follows the training phase. In this stage, the researcher still uses labeled datasets to test whether the training results are satisfactory. If there are areas for improvement, further training can be performed.

The final stage is testing, where the researcher evaluates the trained model using testing data to assess its performance. The purpose of this testing stage is to verify the accuracy of the model in both validation and prediction. The data used for testing is different from the training and validation data.

3.10. Performance Metrics

In this research, performance metrics can be categorized into two groups: performance metrics used to measure image enhancement results and performance metrics used to assess face detection.

To measure the image enhancement results, two performance metrics are utilized: Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM). PSNR measures the quality of enhanced image by evaluating the signal-to-noise ratio, while SSIM quantifies the structural similarity between the enhanced image and the original image.

On the other hand, to evaluate the face detection performance, the Average Precision (AP) metric is used. AP calculates the precision of the face detection model, which assesses the accuracy of detecting faces in the images.

By using these performance metrics, the researchers can effectively measure and compare the quality of image enhancement and the accuracy of face detection in the experiment.

4. Result and Discussion

An image enhancement experiment was conducted on the LOL Dataset. This experiment aimed to assess the performance of each method in improving images under low-light conditions. Each method underwent training and validation processes to measure its performance in enhancing images in low-light conditions.

Each method was tested using the LOL Dataset, and the evaluation metrics used were PSNR and SSIM. Table 1 provides a summary of each model's ability to enhance low-light image quality in the LOL Dataset.

Table 1. Training and Validation for Image Enhancement

Method	Training		Validation	
	PSNR	SSIM	PSNR	SSIM
Original	5.68	0.16	7.74	0.17
Deep Lightening Network	14.42	0.62	17.72	0.66
Deep Retinex Net	15.84	0.41	17.55	0.43
Signal to Noise Ratio Aware	20.98	0.787	24.61	0.84

From Table 1, it can be observed that the Signal to Noise Ratio Aware method outperforms the other two methods, Deep Lightening Network and Deep Retinex Net. Both in terms of PSNR and SSIM, the Signal to Noise Ratio Aware method yields higher values compared to Deep Lightening Network and Deep Retinex Net. The Signal to Noise Ratio Aware method achieves a PSNR of 20.98 and SSIM of 0.787 in the training process, and a PSNR of 24.61 and SSIM of 0.842 in the validation process.

Hyperparameter tuning was also performed for each model to achieve optimal performance. Several configurations were tuned, including batch size, epoch, optimizer, and learning rate. Table 2 presents the configurations used for each method to obtain the performance values in Table 1.

Table 2. Hyper Parameter Tuning Configuration for Image Enhancement Method

Method	Batch Size	Epoch	Optimizer	Learning Rate
Deep Lightening Network	32	500	Adam	1.00E-05
Deep Retinex Net	32	500	Adam	1.00E-04
Signal to Noise Ratio Aware	32	1200	Adam	1.00E-07

From Table 2, it can be observed that the Signal to Noise Ratio Aware method achieved its best performance with a batch size of 32, 1200 epochs, Adam optimizer, and learning rate of 1.00E-07.

Each training and validation process will result in a model that will be used for testing. The configuration used for each training and validation process to achieve the best performance is the same, with 250 epochs, a batch size of 32, a learning rate of 1.00E-02, and SGD optimizer. Table 3 provides a summary of the training and validation results for each tested method.

Table 3. Training and Validation Result

Method	AP Training	AP Validation
Original DARKFACE Dataset + RetinaFace	6.89%	7.19%
Deep Lightening Network + RetinaFace	50.31%	47.92%
Deep Retinex Net + RetinaFace	30.10%	35.91%
Signal to Noise Ratio Aware + RetinaFace	49.90%	50.71%

After obtaining the best models from the training and validation processes, these models are used for testing. The summary of AP scores as the evaluation results from the testing process can be seen in Table 4.

Table 4. Testing Result

Method	AP Testing
Original DARKFACE Dataset + RetinaFace	7.73%
Deep Lightening Network + RetinaFace	51.98%
Deep Retinex Net + RetinaFace	36.99%
Signal to Noise Ratio Aware + RetinaFace	52.92%

From Table 4, it can be observed that the method with the best AP score is the Signal to Noise Ratio Aware method. This aligns with the PSNR and SSIM values of the Signal to Noise Ratio Aware method, where the PSNR and SSIM values are higher compared to the other two methods, Deep Lightening Network and Deep Retinex Net.

There is a theory behind why the combination of Signal to Noise Ratio and RetinaFace is the best approach for face detection in low-light conditions. The Signal to Noise Ratio method is effective in improving images compared to the Deep Lightening Network and Deep Retinex Net methods because it is specifically designed to enhance the signal-to-noise ratio in an image. This helps reduce the level of noise in the image and improve overall image quality. On the other hand, RetinaFace is a strong method for face detection due to its deep convolutional neural network architecture. This architecture consists of multiple convolutional and merging layers that enable the model to learn complex features from image data.

The research also demonstrates that the best combination of image quality enhancement and face detection methods is achieved by combining the best methods from their respective fields. For example, the Signal to Noise Ratio method excels in image quality enhancement, which consequently contributes to superior face detection results in low-light conditions.

5. Conclusion

The research on Image Quality Enhancement with Convolutional Neural Network for Face Detection in Low-Light Conditions has been successfully conducted. The combination of face detection with image enhancement has been proven to enhance the face detection capabilities in low-light conditions compared to face detection without image enhancement. In this study, the combination of image quality enhancement methods with face detection methods that yielded the best AP evaluation score was the Signal to Noise Ratio Aware method for image quality enhancement, combined with the RetinaFace method for face detection. This combination resulted in an AP score of 52.92%.

Although this research has been completed, there are still many areas for further development. The dataset used in this study is limited to the DARKFACE dataset, and future experiments or research can explore different datasets. Additionally, further research can evaluate the quality of face detection, considering factors such as detection stability, ability to detect faces with different poses, facial expressions, or faces wearing accessories like masks, in addition to detection accuracy.

References

[1] Wang, L.-W., Liu, Z.-S., Siu, W.-C., & Lun, D. P. K. (2020). Lightening Network for Low-Light Image Enhancement. Dalam IEEE Transactions on Image Processing (Vol. 29, hlm. 7984–7996). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/tip.2020.3008396

[2] Wei, C., Wang, W., Yang, W., & Liu, J. (2018). Deep retinex decomposition for low-light enhancement. arXiv. http://arxiv.org/abs/1808.04560.

[3] Xu, X., Wang, R., Fu, C.-W., & Jia, J. (2022). SNR-Aware Low-light Image Enhancement. In 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. https://doi.org/10.1109/cvpr52688.2022.01719

[4] Deng, J., Guo, J., Zhou, Y., Yu, J., Kotsia, I., & Zafeiriou, S. (2019). RetinaFace: Single-stage Dense Face Localisation in the Wild (Version 2). arXiv. https://doi.org/10.48550/ARXIV.1905.00641

[5] Zhang, Y., Di, X., Zhang, B., Li, Q., Yan, S., & Wang, C. (2021). Self-supervised Low Light Image Enhancement and Denoising (Versi 1). arXiv. https://doi.org/10.48550/ARXIV.2103.00832

[6] Patil, A., Chaudhari, T., Deo, K., Sonawane, K., & Bora, R. (2020). Low Light Image Enhancement for Dark Images. Dalam International Journal of Data Science and Analysis (Vol. 6, Issue 4, hlm. 99). Science Publishing Group. https://doi.org/10.11648/j.ijdsa.20200604.11

[7] Park, S., Yu, S., Moon, B., Ko, S., & Paik, J. (2017). Low-light image enhancement using variational optimization-based retinex model. IEEE Transactions on Consumer Electronics, 63(2), 178-184.

[8] Guo, Y., Lu, Y., Liu, R. W., Yang, M., & Chui, K. T. (2020). Low-Light Image Enhancement With Regularized Illumination Optimization and Deep Noise Suppression. In IEEE Access (Vol. 8, pp. 145297–145315). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/access.2020.3015217

[9] Guo, X., Li, Y., & Ling, H. (2017). LIME: Low-Light Image Enhancement via Illumination Map Estimation. In IEEE Transactions on Image Processing (Vol. 26, Issue 2, pp. 982–993). Institute of Electrical and Electronics Engineers (IEEE). https://doi.org/10.1109/tip.2016.2639450

[10] Ren, W., Liu, S., Ma, L., Xu, Q., Xu, X., Cao, X., ... & Yang, M. H. (2019). Low-light image enhancement via a deep hybrid network. IEEE Transactions on Image Processing, 28(9), 4364-4375.

[11] Shen, L., Yue, Z., Feng, F., Chen, Q., Liu, S., & Ma, J. (2017). Msr-net: Low-light image enhancement using deep convolutional network. arXiv preprint arXiv:1711.02488.

[12] S. Singhal, S. Nanduri, Y. Raghav and A. S. Parihar, "LRD-Net: A Lightweight Deep Network for Low-light Image Enhancement," 2021 3rd International Conference on Signal Processing and Communication (ICPSC), 2021, pp. 647-651, doi: 10.1109/ICSPC51351.2021.9451681.

[13] Guo, C., Li, C., Guo, J., Loy, C. C., Hou, J., Kwong, S., & Cong, R. (2020). Zero-reference deep curve estimation for low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 1780-1789).

[14] Nath, Raktim & Kakoty, Kaberi & Bora, Dibya & Welipitiya, Udari. (2021). Face Detection and Recognition Using Machine Learning. 43. 194-197.

[15] Vimal, C. (2022). Face Detection's Various Techniques and Approaches: A Review. Dalam International Journal for Research in Applied Science and Engineering Technology (Vol. 10, Issue 1, hlm. 839–843). International Journal for Research in Applied Science and Engineering Technology (IJRASET). https://doi.org/10.22214/ijraset.2022.39890

[16] Jadhav, A., Lone, S., Matey, S., Madamwar, T., & Jakhete, S. (2021). Survey on face detection algorithms. International Journal of Innovative Science and Research Technology, 6(2).

[17] Shamrat, F. J. M., Al Jubair, M., Billah, M. M., Chakraborty, S., Alauddin, M., & Ranjan, R. (2021). A Deep Learning Approach for Face Detection using Max Pooling. In 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) (pp. 760-764). IEEE.

[18] Jiang, H., & Learned-Miller, E. (2017, May). Face detection with the faster R-CNN. In 2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017) (pp. 650-657). IEEE.

[19] Li, H., Lin, Z., Shen, X., Brandt, J., & Hua, G. (2015). A convolutional neural network cascade for face detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 5325-5334).

[20] Obaida, T. H., Hassan, N. F., & Jamil, A. S. (2022). Comparative of Viola-Jones and YOLO v3 for Face Detection in Real time. Iraqi Journal Of Computers, Communications, Control And Systems Engineering, 22(2), 63-72.

[21] Li, C., Wang, R., Li, J., & Fei, L. (2020). Face detection based on YOLOv3. In Recent Trends in Intelligent Computing, Communication and Devices (pp. 277-284). Springer, Singapore.

[22] Li, X., Lai, S., & Qian, X. (2021). Dbcface: Towards pure convolutional neural network face detection. IEEE Transactions on Circuits and Systems for Video Technology, 32(4), 1792-1804.

[23] Zhang, S., Chi, C., Lei, Z., & Li, S. Z. (2020). Refineface: Refinement neural network for high performance face detection. IEEE transactions on pattern analysis and machine intelligence, 43(11), 4008-4020.

[24] Yu, J., Hao, X., & He, P. (2021). Single-stage Face Detection under Extremely Low-light Conditions. Dalam 2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). 2021 IEEE/CVF International Conference

on Computer Vision Workshops (ICCVW). IEEE. https://doi.org/10.1109/iccvw54120.2021.00392.

[25] J. Liang et al., "Recurrent Exposure Generation for Low-Light Face Detection," in IEEE Transactions on Multimedia, vol. 24, pp. 1609-1621, 2022, doi: 10.1109/TMM.2021.3068840.

[26] Wang, W., Yang, W., & Liu, J. (2021). HLA-Face: Joint High-Low Adaptation for Low Light Face Detection. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 16190-16199.

[27] Theoffilus, R. B., Dharmadinata, O. J., & Kusuma, G. P. (2022). Low-Light Face Detection using Deep Learning. Journal of Theoretical and Applied Information Technology, 100(10).

[28] Wang, P., Ji, L., Ji, Z., Gao, Y., & Liu, X. (2021). 1st Place Solutions for UG2+ Challenge 2021 - (Semi-)supervised Face detection in the low light condition. ArXiv, abs/2107.00818.

[29] Asadzadeh, M., & Rikhtegar, A. (2018). Face Detection at the Low Light Environments. Journal of Artificial Intelligence in Electrical Engineering, 6(24), 29-37.

[30] Wakhidah, N. (2011). Perbaikan Kualitas Citra Menggunakan Metode Contrast Stretching. Jurnal Transformatika, 8(2), 78-83.

[31] Sugiarti, S. (2018). Peningkatan Kualitas Citra Dengan Metode Fuzzy Possibility Distribution. ILKOM Jurnal Ilmiah, 10(1), 100-104.