Ministry of Science and Higher Education of the Republic of Kazakhstan

SDU University

Abdrakhim Diana

Denoising face recognition system

THESIS

Presented in Partial Fulfilment for the

Degree of Master of Technical Science in Computer Science

(degree code: 7M06102)

Department of Computer Science

Faculty of Engineering and Natural Sciences

Supervisor: **Baimukashev Rashid**

Kaskelen, June 2024

SDU University
Faculty of Engineering and Natural Sciences
Department of Computer Science

Dean of Faculty of Engineering and Natural Sciences

Assistant Professor, PhD. Akhmedov Ramis

____

« _____» _________________________________ 2024

Topic of the thesis:
Denoising face recognition system

Thesis submitted as part of the requirements for the award of the MSc in "7M06102 - Computer Science" , SDU University, 2024

Head of Department Mukash Zhanar

Academic Supervisor Baimukashev Rashid

Master student Abdrakhim Diana

Kaskelen, 2024

Declaration

I confirm that this is my own work and the use of all material from other sources has been properly and fully acknowledged.

Abdrakhim Diana

June 2024

Acknowledgements

With great appreciation, I would like to thank my supervisor for all of their help and advice during my master's program. I also owe my mother a huge debt of gratitude for her unwavering encouragement and support. Without their tremendous help and confidence in me, my thesis would not have been feasible.

Dedication

This thesis is dedicated to my family for their unwavering support and belief in my dreams, and to my mentors, whose guidance has been instrumental in shaping my academic journey.

Abstract

This thesis investigates the influence of sophisticated denoising techniques on the efficacy of face recognition systems, particularly in environments characterized by substantial image noise. Considering the dependency of face recognition algorithms on the quality of input images, this research conducts a comprehensive evaluation of various denoising strategies, ranging from conventional filters like Median, Bilate-ral, and Gaussian, to advanced deep learning approaches, exemplified by the Deep Convolutional Neural Network (DnCNN). The Extended Yale B dataset, augmented with synthetically introduced noise, provides the basis for this empirical study. Employing quantitative metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM), coupled with qualitative evaluations, this dissertation quantifies the enhancements in image quality and recognition precision afforded by each denoising method. The findings affirm that the integration of advanced denoising algorithms markedly improves the accuracy of face recognition systems, highlighting the efficacy of adaptive deep learning solutions in addressing the complexities introduced by noisy visual environments.

Аннотация

В этой диссертации исследуется влияние сложных методов шумоподавления на эффективность систем распознавания лиц, особенно в средах, характеризующихся значительным шумом изображения. Учитывая зависимость алгоритмов распознавания лиц от качества входных изображений, в этом исследовании проводится всесторонняя оценка различных стратегий шумоподавления, начиная от обычных фильтров, таких как Медианный, Гауссовский и Двусторонний, и заканчивая передовыми подходами глубокого обучения, примером которых является Глубокая сверточная нейронная сеть. (DnCNN). Расширенный набор данных Yale B, дополненный синтетически введенным шумом, обеспечивает основу для этого эмпирического исследования. Используя количественные показатели, такие как пиковое отношение сигнал/шум (PSNR) и индекс структурного сходства (SSIM), в сочетании с качественными оценками, в этой диссертации количественно оцениваются улучшения качества изображения и точности распознавания, обеспечиваемые каждым методом шумоподавления. Результаты подтверждают, что интеграция передовых алгоритмов шумоподавления заметно повышает точность систем распознавания лиц, подчеркивая эффективность адаптивных решений глубокого обучения в решении сложностей, связанных с шумной визуальной средой.

Аңдатпа

Бұл диссертация бетті тану жүйелерінің тиімділігіне, әсіресе кескіннің айтарлықтай шуылымен сипатталатын орталарда күрделі деносизация әдістерінің әсерін зерттейді. Бетті тану алгоритмдерінің кіріс кескіндерінің сапасына тәуелділігін ескере отырып, бұл зерттеу Median, Gauss және Bilateral сияқты кәдімгі сүзгілерден бастап терең конволюционды нейрондық желі мысалында терең оқытудың жетілдірілген тәсілдеріне дейінгі әртүрлі деноизация стратегияларын жан-жақты бағалауды жүргізеді. (DnCNN). Синтетикалық түрде енгізілген шумен толықтырылған кеңейтілген Yale B деректер жинағы осы эмпирикалық зерттеуге негіз береді. Сигнал-шудың шыңына қатынасы (PSNR) және құрылымдық ұқсастық индексі (SSIM) сияқты сандық көрсеткіштерді, сапалы бағалаулармен бірге пайдалана отырып, бұл диссертацияда әр деноэзия әдісімен қамтамасыз етілген кескін сапасы мен тану дәлдігіндегі жақсартулар сандық сипатталады. Нәтижелер дыбыссыздандырудың жетілдірілген алгоритмдерін біріктіру шулы визуалды орталар енгізетін күрделіліктерді шешуде бейімделген терең оқыту шешімдерінің тиімділігін көрсете отырып, тұлғаны тану жүйелерінің дәлдігін айтарлықтай жақсартатынын растайды.

Abbreviations

DnCNN - Deep Convolutional Neural Network

PSNR - Peak Signal-to-Noise Ratio

SSIM - Structural Similarity Index

CNNs - Convolutional Neural Networks

BM3D - Block-matching 3D Filtering

DSDSA - Deep Stacked Denoising Sparse Autoencoders

DeLFN - Denoising of Low-Frequency Noise

GANs - Generative Adversarial Networks

RNNs - Recurrent Neural Networks

LDA - Linear Discriminant Analysis

EBGM - Elastic Bunch Graph Matching

SAE - Stacked Autoencoder

PCA - Principal Component Analysis

SVM - Support Vector Machines

dNS - Nearest Subspace

SRC - Sparse Representation Classification

LBP - Local Binary Patterns

Table of Contents

Declaration **i**

Acknowledgements **ii**

Dedication **iii**

Abstract **iii**

List of Abbreviations **vii**

1 Background and motivations **1**
1.1 Introduction 2
1.2 Problem Statement 2

2 Literature Review **4**
2.1 Face Recognition Techniques 11
2.1.1 Classical approaches 11
2.1.2 Modern approaches 13
2.1.3 Hybrid Approaches 14

3 Denoising Methods **16**
3.1 Classical denoising methods 16
3.2 Transform techniques in image denoising 22
3.3 BM3D 23
3.4 Deep learning-based denoising methods 24

4 Experiments **26**
4.1 Dataset 26
4.2 Evaluation Metrics 27

5 Results and Analysis **29**
5.1 Performance Comparison 29
5.2 Impact of Denoising on Recognition Accuracy 30

6 Conclusion **33**
6.1 Conclusion 33

Chapter 1

Background and motivations

In recent years, face recognition systems have grown in significance. They serve the purposes of biometrics, security, surveillance, and human-computer interaction. These systems use a person's facial features to automatically identify or validate them. Managing undesired changes in facial pictures is a major problem. Numerous factors, such as dim illumination, subpar cameras, motion blur, obstructions, and sensor problems, can cause noise in face recognition. The quality of facial photographs can be lowered by this noise, which can raise false positives or negatives and reduce the accuracy of recognition. Thus, denoising techniques are essential for boosting overall face recognition accuracy and the quality of facial images [1].

There are several main reasons why it is vital to research denoising in face recognition systems. The main objective is to lessen or completely remove the influence of noise, which is essential for efficient recognition in difficult circumstances. Face recognition systems can perform better in real-world scenarios by strengthening and resolving noise-related problems.Secondly, denoising assists in reducing undesired fluctuations in facial photographs, including variations in lighting, shadows, and image capture or transmission artefacts [2]. It helps to improve facial feature alignment and matching, which enhances identification accuracy, by minimising these variances [2]. Furthermore, low-quality photos can occasionally cause face recognition to struggle. Denoising techniques assist in resolving this issue and enhance the identification system's overall effectiveness. In real life, we frequently see pixelated, distorted, or fuzzy photos. These images can be improved with denoising algorithms to improve recognition precision. They "clean up" the picture by bringing back the details and eliminating visual noise. Also a tonne of visual data is captured by contemporary cameras. Excessive detail is produced by high-resolution photos and 3D scans, but noise is also increased [3]. Therefore, specific denoising techniques are essential for facial recognition technology. They preserve crucial facial information while filtering out unwanted noise.

1.1 Introduction

The main goal of a denoising face recognition system is to improve the quality of facial images before the recognition process. By removing unwanted flaws and distortions caused by factors like camera noise or environmental conditions, denoising algorithms aim to make the facial features more accurate and reliable for the recognition system to work with.

Face recognition systems use two types of denoising methods: non-learning and learning-based. Non-learning methods use signal processing like filtering, smoothing, and image restoration algorithms. They assume properties about noise and use filters such as Gaussian, median, or bilateral to reduce noise while keeping facial details [4]. In contrast, learning-based methods rely on machine learning models to map noisy and clean facial images. Deep learning models like convolutional neural networks (CNNs) have shown promising denoising results. They capture complex image structures and noise patterns, and are trained on large datasets to remove noise while preserving facial features [5].

Face recognition systems can greatly be improved by denoising techniques. These techniques enhance the precision and dependability of essential facial feature extraction. As a result, recognition performance is more reliable and consistent, even when noise or other environmental influences are present. Denoising methods contribute to more consistent facial representations by mitigating variances caused by noise. Thus, the accuracy of matching faces is improved [6]. Furthermore, in real-world deployment scenarios with fluctuating image quality, effective denoising enables face recognition systems to generalise better and perform more consistently.In conclusion, improving the clarity of face pictures and lowering ”noise” are essential for face recognition systems to function well. The issues brought on by noise can be helped by denoising algorithms that are both learning-based and non-learning-based. The accuracy of facial recognition systems is much increased when such denoising techniques are incorporated. It also makes them more reliable and effective across different industries and situations. When denoising is used, the systems get stronger and are able to handle more situations. Thus, denoising algorithms improve overall performance of face recognition systems [7].

1.2 Problem Statement

The paper investigates the issue of face recognition performance deteriorating due to noise and unwanted variations in facial images. While face recognition systems have demonstrated remarkable effectiveness in numerous applications, factors like poor lighting, motion blur, obstructions, and sensor noise often compromise their accuracy and reliability [8]. These elements introduce variations into facial images, which can impact the overall performance of face recognition systems by increasing the rates of false acceptance or rejection.

The current denoising methods have provided useful insights for addressing image processing issues connected to noise. Facial images, however, have certain characteristics and needs. Therefore, applying these methods straight to face recognition algorithms might not produce the optimum results [9, 10, 11]. As a result, specific denoising techniques created to satisfy the particular requirements of face recognition scenarios are required.

This thesis looks into, assesses, and improves methods for reducing noise and unwanted fluctuations in face photos used for facial recognition. Enhancing face recognition's accuracy and resilience is the aim. The unique characteristics of facial photographs, such as facial features, lighting fluctuations, and the variability of facial expressions, will be taken into account by the suggested denoising algorithms. Standard datasets and pertinent indicators will be utilised to evaluate the efficacy of the proposed methods. To assess the efficacy and efficiency of the suggested methods in boosting face recognition performance and increasing facial picture quality, a comparison with existing denoising techniques will be made. By tackling the denoising issues in face recognition systems, the results of this work will improve biometric technologies and their practical applications, such as identity verification, security systems, and surveillance [12]. The proposed facial image denoising algorithms have the potential to enhance the accuracy and reliability of face recognition systems, hence enabling their successful use in real-world scenarios characterised by varying noise levels and unwanted alterations in facial representations.

Chapter 2

Literature Review

In the realm of digital imaging, noise is characterized as random variations in brightness or color information that appear in images, serving as an unwanted byproduct that obscures the true content. This visual distortion degrades the quality of the data, impacting the accuracy of subsequent analyses in applications ranging from basic photo editing to advanced computer vision tasks [13].

Noise can originate from a variety of sources, each contributing differently to the degradation of image quality. Electronic noise is primarily caused by the camera's sensor and circuitry, which is more pronounced under low light conditions or when the sensor overheats during extended use. This includes read noise, thermal noise, and quantization noise, all of which are intrinsic to the electronic nature of image capturing devices [13].

Photon noise, also known as shot noise, arises from the quantum nature of light itself and the inherent variations in the number of photons that hit the sensor [14]. This type of noise is fundamentally unavoidable and is tied to the statistical nature of light photons and their interaction with the image sensor [15].

Environmental factors also contribute to noise through external conditions that affect light capture, such as atmospheric interference or fluctuations in lighting conditions. Additionally, compression noise may occur during the image compression process, particularly with lossy algorithms like JPEG, where important information is discarded to reduce file size[16].

Processing noise may be introduced during image enhancements such as sharpening, where the intent to clarify image details can inadvertently amplify existing noise or generate new artifacts.

The manifestation of noise in images can vary, presenting several common types that affect image processing differently. Gaussian noise, often modeled by a normal distribution, affects every pixel in the image with a constant variance, typically introduced by the electronic components of the camera. Salt-and-pepper noise, or impulse noise, randomly turns pixels black or white, resulting in sharp and sudden disturbances in the image signal [17].

Speckle noise is another variant, most commonly encountered in modalities like radar, ultrasound, and laser imaging, where the coherent nature of the imaging process leads to interference patterns that degrade image clarity [18].

The presence of noise reduces detail and texture clarity, affects the sharpness, and complicates tasks such as image segmentation and recognition in computer vision applications. To mitigate these effects, various noise reduction techniques are employed. Spatial filtering methods like median filtering and Gaussian blurring are commonly used to reduce noise while preserving edges. Frequency domain filtering involves applying specific filters in the frequency domain, such as Wiener or Gaussian filters, to clean up the image [19].

Wavelet transforms offer a means to separate noise from image content by distinguishing between the frequency characteristics, effectively isolating and removing noise components [20]. In more advanced applications, deep learning models such as Deep Convolutional Neural Networks (DnCNN) are employed [21]. These models use machine learning to adaptively learn how to remove noise from images based on extensive training datasets, providing a powerful tool for improving image quality.

In the comprehensive survey "Brief review of image denoising techniques"[22] authors elaborate on a wide array of methodologies for image denoising. The paper is structured to first define the problem of image denoising where the objective is to recover a clean image from its noisy counterpart, described mathematically as

$$y = x + n \tag{2.0.1}$$

, n representing additive white Gaussian noise.

The discussion extends to spatial domain methods where the focus is on spatial filtering, encompassing both linear approaches like mean and Wiener filters and non-linear techniques such as median and bilateral filters. These methods aim to mitigate noise while striving to preserve crucial image features like edges and textures. Additionally, variational denoising techniques are explored, which involve the minimization of an energy function that incorporates predefined image priors, with methods such as total variation regularization and non-local means highlighted for their ability to maintain image integrity while reducing noise.

By the same token , the paper delves into transform domain methods, notably wavelet transforms and block-matching 3D filtering (BM3D), which process images in a domain where noise characteristics are distinctly separable from the signal, facilitating more effective noise reduction.

Figure 1 - Visual comparisons of denoising results on Lena image corrupted by additive white Gaussian noise with standard deviation 30: a. Wiener filtering; b. Bilateral filtering; c. PCA method; d. Wavelet transform domain method; e. BM3D [22]

The narrative progresses to cover machine learning approaches, with a particular emphasis on the use of convolutional neural networks (CNNs) for image denoising. These networks are trained to learn optimal denoising filters directly from data, showcasing robustness across various levels of noise contamination.

The document synthesizes findings across these techniques, noting that methods like BM3D and CNN-based denoising often outperform others in both quantitative metrics such as PSNR and SSIM, and in visual quality assessments. It acknowledges, however, that real-world applications continue to pose significant challenges due to the complex nature of noise beyond laboratory settings.

Briefly to conclude, the review underscores the evolution from traditional filtering techniques to sophisticated machine learning models in image denoising. It emphasizes the need for continued research aimed at developing methods tailored to specific, real-world noise types and suggests a potential shift towards more adaptive and unsupervised learning models that do not rely on clean data for training. This perspective points towards future directions where image denoising can adapt more dynamically to the varying conditions encountered in practical applications.

In this work [23], the authors developed a framework to improve face recognition performance by denoising probe face images using quality assessment-based parameter selection. The paper starts by identifying the common issue in face recognition systems where probe face images may be contaminated with noise due to various factors like environmental conditions, sensor errors, and transmission errors. Such noise significantly degrades the performance of face recognition algorithms.

The authors propose a framework that selects optimal parameters for denoising techniques based on quality assessment of the images. The idea is that by improving the quality of the image through denoising, the recognition performance can be enhanced.

They employed quality assessment algorithms to evaluate the level of noise and other distortions in the images. These assessments help in determining the most suitable parameters for denoising each specific image. A Support Vector Machine (SVM) classifier is used within the framework to learn the relationship between the quality assessment scores and the optimal denoising parameters. This method allows for an automated and intelligent selection of denoising parameters that are most likely to improve image quality and thus recognition accuracy.

The proposed framework was tested on the AR face dataset. The authors conducted experiments to show how the denoising parameters selected by their framework affect the recognition performance. They demonstrated that their method improves both the accuracy and computational efficiency of face recognition systems under noisy conditions.

Their results include detailed analyses of how different denoising parameters influenced the recognition rates. They also presented a correlation study to ascertain the relationship between the quality scores and the recognition rate, confirming that their quality assessment-based parameter selection can effectively enhance face recognition performance.

The authors of the research created a facial recognition system that makes use of autoencoders and deep learning techniques. They designed a face recognition system based on deep stacked denoising sparse autoencoders. This system incorporates deep neural network technologies, sparse autoencoders, and a denoising task within the autoencoder framework. The system uses autoencoders to learn a

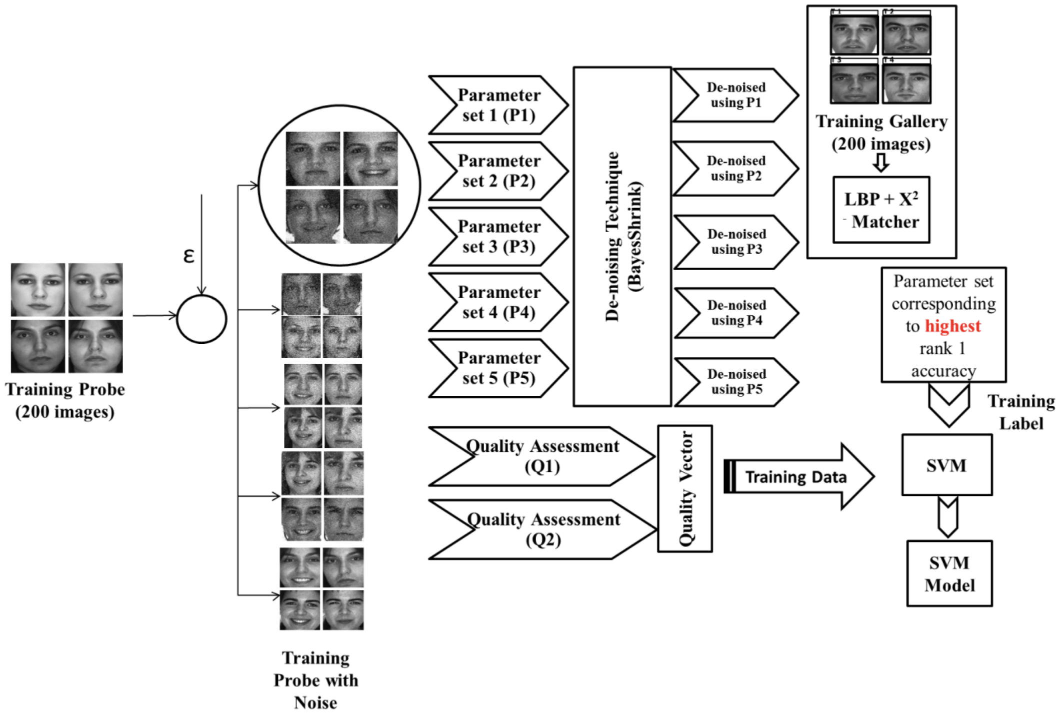


Figure 2 - The training scheme of the proposed assessment based denoising framework.[23]

representation of face images by extracting salient features automatically. These features are intended to reduce dimensionality and improve the representativeness of raw face data. For the classification task, they utilized two types of classifiers: a multi-class Support Vector Machine (SVM) and a Softmax classifier, to enhance the decision-making process in recognizing different faces.

They conducted extensive experiments on several known face databases, such as ORL, Yale, Caltech, and a subset of PubFig. These experiments demonstrated the effectiveness of the DSDSA system in face recognition tasks, showing promising performance and competitive accuracy. Their system is designed to handle various challenges associated with face recognition, such as differences in poses, illumination, expressions, and occlusion. By employing a denoising process in the autoencoder, their method enhances the robustness of the feature extraction process, enabling the system to perform well even with noisy or imperfect data inputs. They meticulously evaluated and optimized the system's parameters through experiments to find the best settings for achieving high accuracy in face recognition.

Generally, their work represents a significant contribution to the field of face recognition by integrating advanced deep learning techniques with traditional face recognition challenges. This integration aims to enhance the accuracy and reliability of face recognition systems in diverse and challenging environments.

Yimei Kang and Wang Pan, the authors of the research paper [24], created a novel technique for enhancing facial recognition efficacy through the process of denoising low-light photos.

Extensive experiments were carried out to analyze the distribution of noise frequencies in low-light images, revealing a significant presence of low-frequency noise. A new denoising approach named DeLFN (Denoising of Low-Frequency Noise) (Fig. 3 -) was introduced, functioning on three distinct levels. First, applies histogram equalization to boost overall image contrast and reduce assorted noise forms. Second, uses logarithmic transformation to refine image details, particularly targeting low-frequency noise. And finally, employs high-pass filtering to lessen remaining very low-frequency noise, thus enhancing the retrieval of features from the genuine images.

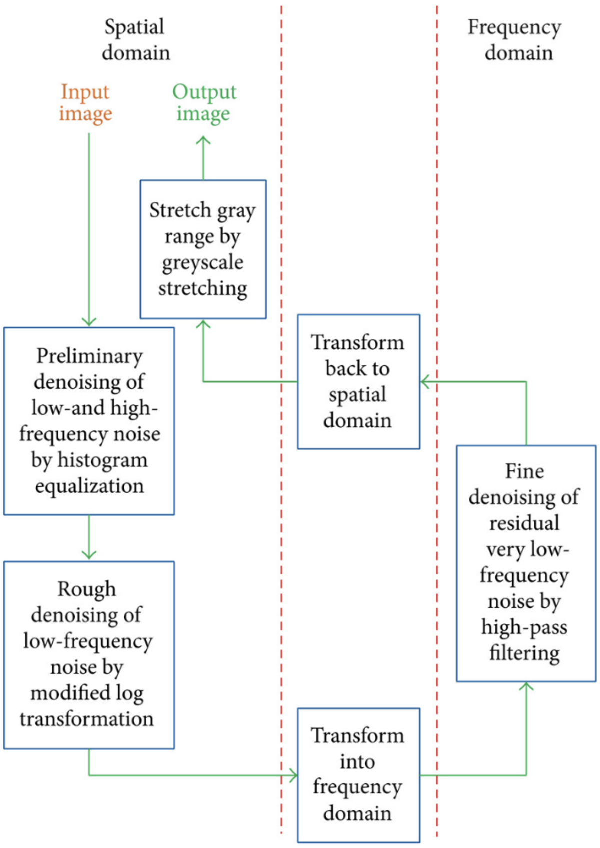


Figure 3 - Flowchart of proposed method DeLFN.[24]

The method was implemented and tested using Principal Component Analysis (PCA) on several face databases, including the Yale Face Database B, Extended Yale Face Database B, and CMU PIE Face Database.

The authors compared DeLFN against several traditional denoising and illumination preprocessing methods. Their findings demonstrated that DeLFN not only enhanced visual quality and face recognition rates but was also simpler and faster, making it suitable for real-time applications.

The paper concludes that the DeLFN method significantly improves the performance of face recognition systems under low-light conditions by effectively reducing noise without compromising essential image details necessary for accurate recognition.

To improve facial recognition systems, the researchers in the study[25] combine Local Binary Pattern (LBP) and Stacked Autoencoders (SAE).

The authors utilize the basic LBP technique to encode the texture features of face images, leveraging its robustness for effective initial feature extraction. To address the challenge of high-dimensional LBP features, they implement a unified LBP approach, which simplifies the feature set and improves computational efficiency.

The extracted features from the LBP process are then input into a stacked autoencoder. The SAE is specifically designed to denoise and further encode these features, aiming to enhance classification and recognition performance. This deep learning architecture uses multiple layers of encoding to abstract features at different levels, which enhances the discriminative capabilities of the system.

The integration of LBP with SAE harnesses the strengths of both techniques: LBP's effectiveness in capturing critical texture details and SAE's ability to compress and denoise features. The workflow involves processing face images with the LBP method initially, followed by encoding through the SAE. This sequence ensures the system's robustness against environmental variables such as noise and lighting variations. This method improves facial recognition technology's accuracy and environmental adaptability, which is a major breakthrough in the field of biometric security.

The LBP algorithm achieved a recognition rate of 90.05% on the Yale database, which is indicative of its high accuracy and effectiveness in handling real-world variations in face images. The system demonstrated strong robustness against changes in lighting, showcasing its utility in diverse and challenging conditions. When compared to other standard algorithms such as PCA and SVM, the LBP-SAE approach outperformed these methods, particularly in environments affected by noise and lighting variations.

The authors of the research [26] suggest a way to strengthen facial recognition algorithms by making them more resilient to different kinds of noise that are frequently encountered in practical situations.

In the study [26] the writers developed a method to enhance the accuracy of face recognition systems by improving their ability to handle noisy data. The method centers on recovering a clean, low-rank subspace from noisy training images. Although face images from a single subject are expected to lie within a low-rank subspace, practical challenges such as occlusions, shadows, and specularities often corrupt this subspace.

The approach involves formulating the challenge of denoising and subspace recovery as a rank-minimization problem. This process decomposes a data matrix into a low-rank matrix, which represents the clean images, and a sparse error matrix, which accounts for noise and outliers.

To address the computational challenges associated with solving the rank-minimization problem directly, the authors adopt a convex surrogate. This surrogate simplifies the problem by replacing the rank function with a nuclear norm—sum of the singular values—and substituting the ℓ_0 norm with the ℓ_1 norm to promote sparsity. These modifications enable the use of advanced convex optimization techniques to solve the problem effectively.

The results of the study showed notable improvements in the robustness of face recognition techniques against various noises affecting both training and testing images. The method adeptly managed real-world image distortions including occlusions and specular reflections by restoring the fundamental low-rank structure of the face images.

The utility of the proposed method was verified using the Extended Yale B Face Database, where it was effective in markedly decreasing noise impacts and enhancing recognition accuracy, even when up to 50% of the pixels in the face images were corrupted.

Furthermore, when the proposed low-rank recovery method was combined with established subspace methods such as Nearest Subspace (NS) and Sparse Representation Classification (SRC), it exhibited enhanced performance. This was particularly evident in scenarios characterized by substantial levels of noise and corruption, demonstrating its superiority over traditional methods.

The paper[27] conducted a thorough quantitative analysis to explore how different image processing techniques influence the performance of deep learning-based face recognition systems. The study primarily focused on the application of Gaussian filtering as a denoising technique and the self-snake model for image enhancement, utilizing real-case forensic data involving 33 face image sets to assess these methods' efficacy.

The methodology employed in this research was rigorous, involving both denoising and enhancement techniques to determine their impact on the accuracy of face recognition. The data used comprised images from actual forensic cases, providing a robust foundation for evaluating the practical implications of the findings. The analysis centered on the hypothesis that improving image quality through processing techniques can significantly enhance recognition accuracy [28].

The findings from this study were revelatory in several aspects such as impact of image quality, denoising versus enhancement, comparative effectiveness.

The investigation substantiated that the quality of face images is a critical determinant of recognition performance. Enhanced image quality, achieved through sophisticated processing techniques, was directly linked to improved accuracy in recognizing faces.

Among the techniques evaluated, Gaussian filtering—a method aimed at reducing noise—proved more effective than the self-snake model [29], which focuses on enhancing image features. The research highlighted that denoising is particularly advantageous for deep face recognition systems used in forensic scenarios, suggesting a pivotal role for denoising in contexts where precision is crucial.

The comparative analysis further demonstrated that the benefits of Gaussian filtering not only exceed those of the self-snake model but also contribute significantly to the precision of face recognition systems. This suggests a clear preference for denoising over enhancement in applications where the integrity and clarity of facial features are paramount.

The experimental validation of these methods used the Extended Yale B Face Database, a well-recognized benchmarking tool in the field. This validation confirmed that denoising, particularly through Gaussian filtering, leads to a substantial improvement in recognition rates, markedly more so than enhancement techniques, especially in environments afflicted by real-world noise [30].

This research offers substantial evidence that advancing image processing techniques, particularly denoising, can significantly enhance the effectiveness of deep learning-based face recognition systems in forensic and security applications. The study not only reinforces the importance of high-quality image data for accurate recognition but also underscores the need for further exploration into image processing techniques that can optimize performance in real-world scenarios [31]. These findings are instrumental for future developments in forensic technology, providing a direction for refining image processing tools to support legal and security frameworks effectively.

2.1 Face Recognition Techniques

2.1.1 Classical approaches

Classical approaches to face recognition primarily rely on feature extraction and matching techniques that analyze facial features and their spatial relationships

1) Principal Component Analysis (PCA)

The Principal Component Analysis (PCA) approach for face recognition is a well-established classical technique that has been extensively studied and utilized within the field. The underlying premise of this methodology is to represent facial imagery within a lower-dimensional ”face space” by identifying the principal components, or eigenfaces, that capture the predominant variability observed across a training set of face images [32].

The PCA-based face recognition framework operates as follows: During the training phase, the eigenfaces are computed by performing principal component analysis on the covariance matrix of the training face images [33]. Each training face is then projected onto this eigenface subspace, and the corresponding projection coefficients are stored in the training database.

For the recognition of a novel facial input, the face image is first preprocessed to normalize factors such as size, pose, and illumination. It is then projected onto the eigenface subspace, and the resulting projection coefficients are compared to those stored in the training database. The face is recognized as the individual whose stored coefficients best match those of the input image.

The advantages of the PCA-based approach lie in its relative simplicity and computational efficiency. However, it can be susceptible to variations in lighting, pose, and expression, as the eigenfaces may not necessarily capture the most discriminative features for effective recognition [34]. Extensions and variations of the PCA technique, such as Fisherfaces and 2D-PCA, have been proposed to address some of these limitations and enhance the robustness of the face recognition system.

2) Linear Discriminate Analysis (LDA)

Linear Discriminant Analysis (LDA) is a refined approach to face recognition that enhances the foundational techniques of Principal Component Analysis (PCA). Unlike PCA, which primarily aims to maximize the variance across the data, LDA seeks to identify a linear combination of features that distinctly separates different classes, each corresponding to individual identities in the context of face recognition [35].

At the heart of LDA's methodology lies the objective to compute a subspace that maximizes the scatter between classes while minimizing the scatter within each class. This dual focus yields a feature representation that is significantly more discriminative, offering enhanced capability to distinguish between individual identities compared to the PCA approach, which does not inherently account for class structures [36].

The process of implementing LDA in face recognition mirrors that of PCA to some extent. Initially, input face images are subjected to preprocessing steps to standardize factors like size, pose, and lighting conditions. Subsequently, the LDA subspace is calculated by identifying the linear discriminant vectors that optimize the ratio of the between-class scatter matrix to the within-class scatter matrix [36]. Each image in the training set is then projected onto this LDA subspace, and the projection coefficients are preserved in the training database.

When recognizing a new facial input, the image is projected onto the LDA subspace. The projection coefficients of this image are then compared with those stored in the training database. Recognition occurs when the identity associated with the closest matching coefficients is identified.

The principal advantage of using LDA lies in its ability to derive a subspace representation that is inherently more discriminative by explicitly incorporating class (identity) information during the computation of the subspace. This capability often translates into superior recognition performance, particularly in scenarios where the dataset exhibits a well-defined class structure.

LDA is not without its challenges; it is particularly susceptible to the curse of dimensionality, requiring a larger number of training samples as the complexity of the face recognition task increases. To mitigate this issue, hybrid techniques such as Fisherfaces have been developed, which integrate PCA and LDA in a sequential two-stage process.

In essence, LDA offers a robust, statistically grounded framework for learning a discriminative representation of facial images, establishing it as a critical technique in the classical arsenal of face recognition methods [37].

3) Elastic Bunch Graph Matching (EBGM)

Elastic Bunch Graph Matching (EBGM)[38] is a well-established facial recognition approach that utilizes a graph-based representation of facial features to achieve robust and accurate face identification. This technique was developed to address the limitations of earlier face recognition methods, which often exhibited sensitivity to variations in pose, illumination, and expression.

The EBGM approach represents a face as a labeled graph, where the nodes correspond to fiducial facial landmarks (e.g., eyes, nose, mouth), and the edges represent the spatial relationships between these landmarks [39]. Each node in the graph is associated with a set of Gabor wavelet coefficients, which capture the local texture and shape information around the corresponding landmark. The system maintains a generic face model, referred to as a "bunch graph," which encodes the prototypical facial structure and appearance across a population of faces.

To recognize a novel face, EBGM first locates the fiducial landmarks on the input face image and constructs a corresponding landmark graph. This input graph is then compared to the bunch graph using an elastic graph matching algorithm, which allows for deformations and variations in the facial structure. The recognition is performed by finding the best match between the input graph and the graphs stored in the bunch graph, typically by minimizing an energy function that captures the similarity between the graphs [39].

The key advantages of the EBGM approach lie in its robustness to variations, its discriminative power, and its localized feature representation. The elastic graph matching mechanism allows EBGM to handle various types of facial variations, such as changes in pose, illumination, and expression, by accounting for deformations in the facial structure. The Gabor wavelet coefficients used to describe the facial landmarks provide a rich and discriminative representation, which can effectively capture the unique facial features of different individuals. Moreover, the graph-based representation and the use of local Gabor features enable EBGM to leverage both the global and local facial information for improved recognition performance [29].

EBGM has been widely studied and applied in the field of face recognition, and it has demonstrated competitive performance compared to other classical techniques, especially in scenarios with challenging variations in facial appearance. However, the computational complexity of the elastic graph matching process can be a limiting factor for some real-time applications.

These classical approaches have been widely studied and used in face recognition research and applications. More recent developments in deep learning have led to significant advancements in face recognition accuracy and robustness, but these traditional techniques still provide a useful foundation for understanding the problem and its evolution.

2.1.2 Modern approaches

In recent years, the field of face recognition has seen significant advancements, driven by the rapid progress in deep learning and computer vision. While classical approaches, such as Linear Discriminant Analysis (LDA) and Elastic Bunch Graph Matching (EBGM), have provided foundational techniques, modern face recognition methods have leveraged the power of deep neural networks to achieve unprecedented levels of accuracy and robustness.

Deep Learning

Deep learning, a branch of machine learning, transforms feature extraction for classification tasks by removing the need for manual feature extraction steps by automatically identifying relevant features during training. Neural networks are forced by this process to learn more discriminating features in order to differentiate between various facial attributes. Deep learning has revolutionized face recognition, with applications that cut across multiple industries [40].

The landscape of face recognition technology has undergone significant transformation with the advent of Convolutional Neural Networks (CNNs)[41]. Notable among these advancements are techniques such as DeepFace, developed by Facebook AI Research, which utilizes a deep CNN trained on an extensive dataset of facial images. Similarly, Google's FaceNet represents another pioneering approach, where a deep learning model directly maps facial images to a compact Euclidean space, with distances in this space directly correlating to facial similarity [42]. Another innovative technique is SphereFace, which employs an angular softmax loss function to cultivate a discriminative face representation, ensuring that faces of the same identity are closely clustered in the feature space.

Beyond CNNs, the exploration of other neural network architectures like Generative Adversarial Networks (GANs) and Recurrent Neural Networks (RNNs) has also been instrumental in addressing more complex challenges such as variations in pose, illumination, and expression [43, 44]. The rapid evolution of computational hardware, particularly the development and accessibility of powerful GPUs, has significantly facilitated the training and practical deployment of these sophisticated deep learning-based face recognition systems. These systems are increasingly being integrated into a variety of real-world applications, ranging from mobile devices and surveillance systems to access control mechanisms.

Despite the remarkable progress and the high levels of performance achieved by modern face recognition techniques, several challenges remain prevalent. Key areas of ongoing research include tackling issues related to bias and fairness, enhancing the ability of these systems to generalize to new, unseen domains, and improving their interpretability and explainability.

Comprehensive, the field of face recognition has witnessed extraordinary advancements, largely propelled by deep learning methodologies. These developments have not only set new performance benchmarks but are also paving the way towards more precise and reliable face recognition technologies in the future.

2.1.3 Hybrid Approaches

The robustness of face recognition systems in noisy environments is significantly enhanced by employing hybrid denoising strategies. These strategies synergistically combine multiple denoising techniques to tackle different types and intensities of noise, thus ensuring the preservation of essential facial features necessary for accurate recognition. This section elaborates on the various hybrid combinations used, their implementation, and the resultant improvements in image quality and recognition accuracy [45].

The study categorizes the denoising combinations based on the number of algorithms combined, distinguishing between dual and triple combinations. Dual combinations involve two different denoising methods applied sequentially, leveraging their complementary strengths to enhance image quality. Conversely, triple combinations involve the sequential application of three denoising methods, providing a more comprehensive noise reduction by addressing various noise characteristics more effectively.

We employed core denoising algorithms in our study, including Median Filtering, Gaussian Filtering, Bilateral Filtering, Non-Local Means (Adaptive) Filtering and DnCNN. Dual combinations, such as Median followed by Gaussian and Median followed by DnCNN, target specific noise types effectively—median filtering tackles salt-and-pepper noise, while Gaussian filtering addresses Gaussian noise. Triple combinations, such as Median, Gaussian, followed by Adaptive, provide a robust approach by smoothing out noise progressively while preserving important textural details essential for recognition.

Chapter 3

Denoising Methods

3.1 Classical denoising methods

Spatial domain filtering

Classical denoising methods in the spatial domain have played a significant role in addressing the problem of image and signal denoising. These techniques aim to remove or reduce unwanted noise while preserving the essential features and details of the original data. The spatial domain refers to the direct manipulation of pixel values or signal samples in the original coordinate space [46, 47].

One of the cornerstone methods for denoising in the spatial domain is the Gaussian filter. This method employs a weighted averaging technique, where each pixel is adjusted based on a Gaussian kernel. The kernel provides higher weights to the central pixel while progressively reducing weights for surrounding pixels according to their spatial distance. This process effectively smooths the image or signal, mitigating high-frequency noise components and maintaining the overall structural integrity.

The median filter is another pivotal spatial domain denoising strategy [48]. It works by substituting each pixel with the median value among its neighboring pixels. This technique is particularly adept at removing impulsive noise, such as salt-and-pepper noise, and is known for its capability to preserve edges and sharp image features.

Enhancing the realm of spatial domain denoising, the bilateral filter integrates both spatial and photometric (intensity) data in its operation. It performs a weighted averaging where the weights are influenced by both the spatial proximity and the intensity differences between the central pixel and its neighbors. This dual consideration facilitates superior edge and detail preservation while effectively diminishing noise [49].

Beyond these foundational techniques, more elaborate methods like non-local means (NLM) and sparse coding-based approaches have been introduced. These advanced strategies leverage the intrinsic self-similarity and sparsity observed in natural images and signals, resulting in enhanced denoising outcomes.

While spatial domain filtering techniques offer computational efficiency and straightforward implementation, their effectiveness can vary with different types of noise, particularly those that are correlated or structured. For such noise types, transform domain methods, which focus on the frequency domain representation of the data, might yield more robust denoising results.

Traditional spatial domain denoising techniques, including Gaussian, median, and bilateral filtering, remain integral to image and signal processing. These methods strike a practical balance between simplicity in execution and effectiveness in noise reduction, thereby serving as essential tools across various applications.

Variational denoising methods

Variational denoising methods have become a prominent approach in the field of image and signal processing. These techniques approach the denoising problem from a variational optimization perspective, where the goal is to find an optimal solution that minimizes a well-defined energy or cost function [22].

The general framework of variational denoising methods involves the encoding of desired properties, such as smoothness, sparsity, or fidelity to the input data, into an energy function. This function typically includes terms that balance the trade-off between noise reduction and the preservation of important features. Regularization techniques, such as Total Variation (TV) regularization and sparse regularization, are often employed to impose additional constraints or priors on the denoised output. The energy function can be expressed as[22]:

$$\hat{x} = \arg\min_{x} \left\{ \frac{1}{2}\|y - x\|_2^2 + \lambda R(x) \right\} \tag{3.1.1}$$

where $\hat{x}$ is the estimated clean image, x represents potential estimates of the clean image as the variable of optimization, y is the observed noisy image, $\|y - x\|_2^2$ is the data fidelity term representing the squared L^2 norm of the difference between the noisy image and the estimate, which penalizes deviations from the observed data, $R(x)$ is the regularization term that imposes a prior or promotes certain properties in the estimated image such as smoothness or sparsity, and λ is a scalar regularization parameter that balances the trade-off between fidelity and regularization [22].

The energy function in variational denoising is minimized through the application of optimization algorithms, such as gradient descent, proximal methods, or iterative schemes, to achieve the denoised output. This optimization process seeks to identify the solution that most effectively adheres to the constraints and objectives delineated in the energy function.

Prominent among variational denoising techniques is TV denoising, which reduces the total variation of the image to promote piecewise smoothness while preserving edges. Another notable method is sparse coding-based denoising, predicated on the assumption that the clean image can be expressed as a sparse linear combination of a few dictionary elements. The Plug-and-Play priors approach offers a modular framework by substituting the regularization term with a pre-trained denoising module, thus enabling the integration of advanced denoising algorithms, including those based on deep learning, into the variational optimization framework.

Variational denoising methods provide several benefits, such as the capability to integrate various priors and constraints, adaptability to diverse noise types, and the potential for effective preservation of edges and features. Nonetheless, these methods may entail greater computational demands compared to some spatial domain filtering techniques, particularly in large-scale optimization scenarios.

The utilization of variational denoising methods has significantly influenced the field of image and signal processing, establishing a robust framework for tackling challenges associated with denoising.

Total variation regularization

Total Variation (TV) regularization is a classical denoising method extensively employed in image processing to maintain edge integrity while eliminating noise [50]. This technique is particularly efficacious in scenarios where the preservation of edges, lines, and other high-frequency components is paramount. Total Variation denoising operates on the principle of minimizing the total variation of the image, predicated on the assumption that a real image exhibits a smaller total variation compared to a noisy one [51]. The total variation is a measure of the aggregated magnitude of the image gradient, expressed mathematically for an image u as[22]:

$$\mathrm{TV}(u) = \int |\nabla u| \, dx \tag{3.1.2}$$

where ∇u represents the gradient of the image.

The core objective in TV regularization is to identify an image u that minimizes the following functional:

$$\min_u \left(\frac{1}{2} \int (u - f)^2 \, dx + \lambda \int |\nabla u| \, dx \right) \tag{3.1.3}$$

Here, f represents the noisy image, u is the denoised image, and λ is a regularization parameter that controls the trade-off between fidelity to the original image and smoothness of the output.

In this expression, f represents the noisy image, u the denoised image, and λ a regularization parameter that modulates the balance between fidelity to the original image and the smoothness of the output.

One of the primary advantages of TV regularization is its capacity to preserve sharp edges, distinguishing it from methods that rely on local averaging. This attribute is crucial for a multitude of visual applications where edge definition significantly impacts performance. Moreover, TV denoising proficiently reduces various types of noise, such as Gaussian noise, by smoothing areas of minimal variation while maintaining distinct boundaries. Its robust performance across diverse applications renders it a reliable choice in practical settings, finding utility in fields ranging from medical imaging, where it enhances the quality of MRI and CT scans, to astronomical imaging, used to improve clarity in images captured under conditions of low light or high background noise. In the domain of computer vision, it often serves as a preprocessing step in tasks like object detection, tracking, and segmentation. Thus, Total Variation regularization remains a fundamental tool in many denoising tasks and continues to influence the development of newer, more advanced denoising algorithms [52].

Non-local regularization

Non-local regularization represents a sophisticated technique employed in image denoising and other signal processing tasks, which capitalizes on the redundancy of similar features across different regions of an image or signal. This method deviates from traditional local techniques that focus solely on a neighborhood around each pixel, instead utilizing the widespread similarity of pixel patches throughout the image. This global approach allows for more effective denoising, as it can harness a broader dataset to optimize the noise reduction for each pixel [53].

The mathematical formulation for non-local regularization, particularly focusing on Non-Local Means (NLM), is based on the concept of averaging all pixels in an image that have similar neighborhoods [54]. The formula that captures the essence of non-local regularization[22]:

Given a pixel x_i in an image x, the NLM-filtered value $\hat{x}_i$ is computed as:

$$\hat{x}_i = \frac{1}{C(i)} \sum_j w(i,j) x_j \tag{3.1.4}$$

where j indexes all pixels in the image, $w(i,j)$ is the weight assigned to the j-th pixel's influence on the i-th pixel, typically computed based on the similarity between their respective neighborhoods, $C(i)$ is a normalization factor to ensure that the weights sum to one, usually given by[22]:

$$C(i) = \sum_j w(i,j) \tag{3.1.5}$$

The weight $w(i,j)$ is often defined using a Gaussian function of the Euclidean distance between the intensity values in the neighborhoods of x_i and x_j, scaled by a parameter h:

$$w(i,j) = \exp\left(-\frac{\|v(N_i) - v(N_j)\|^2}{h^2}\right) \tag{3.1.6}$$

Here $v(N_i)$ and $v(N_j)$ are vectors representing the pixel values in the neighborhoods around x_i and x_j, respectively, h is a filtering parameter that controls the decay of the exponential function, influencing the degree of smoothing[22].

This formulation allows non-local means to exploit redundancy in the image by averaging over all similar patches, thus effectively reducing noise while preserving important structural details like edges.

The principle behind non-local regularization is grounded in the non-local means algorithm, which is predicated on the observation that patches of pixels in natural images frequently have comparable counterparts elsewhere within the same image [55]. In the process of non-local regularization, each pixel or small patch in the image is compared to other patches to find similarities, typically quantified using the squared difference within a Gaussian window.

Once similarities are established, the algorithm calculates weights indicating the significance of each similar patch's contribution to the denoising process. These weights are generally higher for patches that show greater similarity. The next step involves updating each pixel by averaging over all its similar patches, with this averaging weighted according to the previously calculated similarities. This method is particularly effective in reducing noise because identical noise patterns rarely occur across similar patches [56].

At times, an additional regularization term is incorporated to control the smoothness of the output and mitigate overfitting to the noise. Non-local regularization is utilized extensively in applications such as reducing noise in digital images—important in fields like medical imaging and satellite imagery—as well as in video and audio signal processing [57]. In video denoising, the technique may be applied frame-by-frame or by using three-dimensional patches that span multiple frames to diminish temporal noise.

The robustness of non-local regularization against noise and its ability to preserve intricate image details without blurring are notable advantages. However, the technique is computationally demanding due to the extensive search required for similar patches across the entire image or signal [57]. Additionally, the effectiveness of non-local regularization can be significantly affected by the choice of parameters, such as the size of the patches and the dimensions of the Gaussian window used for assessing similarity.

In recent developments, non-local regularization has been adapted and enhanced through the integration of machine learning techniques, which can more adeptly predict the appropriate weights, and by its incorporation into deep learning frameworks, which further improves performance.

Sparse representation

Sparse representation is a classical denoising method extensively utilized in signal and image processing domains. The principal concept of sparse representation involves depicting a noisy signal or image as a linear combination of a minimal number of elements from a predefined dictionary [58, 59]. These elements typically comprise basis functions such as wavelets, with the representation designed to encapsulate the essential characteristics of the data while excluding the noise . Sparse representation-based denoising models can be expressed using the optimization formulation[22]:

$$\hat{\alpha} = \arg\min_{\alpha} \|y - D\alpha\|_2^2 + \lambda\|\alpha\|_1 \tag{3.1.7}$$

where y is the observed noisy image, D is the over-complete dictionary, α represents the sparse coefficients, λ is the regularization parameter that controls the sparsity level.

This model aims to find a sparse representation α such that the reconstructed image $D\alpha$ is close to the noisy observation y, while keeping the representation as sparse as possible, indicated by the ℓ_1-norm regularization term $\|\alpha\|_1$. The balance between fidelity to the data and sparsity is managed by the parameter λ.

The process begins with the construction of a dictionary that consists of basis functions. This dictionary might be pre-defined, using sets of wavelets or Fourier bases, or could be learned directly from the data. Following the dictionary construction, the noisy signal or image is approximated by selecting the fewest possible atoms from this dictionary. This approximation process is generally formulated as an optimization problem aimed at minimizing the reconstruction error with a constraint on sparsity.

Several optimization techniques are employed to solve this problem, including Basis Pursuit, Orthogonal Matching Pursuit (OMP), and the Least Absolute Shrinkage and Selection Operator (LASSO). Once the sparse coefficients are determined, the denoised signal or image is reconstructed by aggregating the selected atoms, each scaled by these coefficients.

Sparse representation finds its applications in various fields such as image denoising, where it is particularly effective in removing noise while preserving essential details like edges. It is similarly employed in audio processing to diminish background noise and enhance the clarity of speech or musical signals [60].

The method offers significant flexibility and efficiency, proving especially potent when the true signal inherently supports a sparse representation. However, the performance of this method is heavily contingent upon the choice of dictionary. Inadequate dictionary representation of the underlying signal can significantly impair denoising performance. Additionally, the optimization process involved in sparse representation can be computationally demanding, particularly with large datasets or dictionaries [61].

Sparse representation has not only been foundational in the field of denoising but has also facilitated the development of more advanced techniques, including those based on machine learning and deep learning approaches [60].

Low-rank minimalization

Low-rank minimization is a classical denoising technique that has garnered significant attention for its effectiveness in applications where the underlying data is well-approximated by a low-dimensional linear subspace [62]. This technique is frequently employed in image and video denoising, as well as in other areas involving large matrices or data arrays, where the intrinsic data dimensionality is lower than the ambient space [63].

The essence of low-rank minimization techniques is to recover a clean signal from noisy observations by positing that the underlying clean data resides in a low-dimensional subspace. This approach typically formulates the denoising challenge as an optimization problem [64]:

$$\hat{X} = \arg\min_{X} \|Y - X\|_F^2 \quad \text{subject to} \quad \text{rank}(X) \leq r \tag{3.1.8}$$

In this formulation, Y represents the observed noisy matrix, X is the matrix denoting the denoised output, $\|\cdot\|_F$ denotes the Frobenius norm, and r is the rank constraint. Direct minimization of the rank function is computationally prohibitive due to its non-convex nature, leading most approaches to focus on convex relaxations such as nuclear norm minimization (NNM), which approximates the rank by the sum of the singular values of X:

$$\hat{X} = \arg\min_{X} \|Y - X\|_F^2 + \lambda\|X\|_* \tag{3.1.9}$$

Here, $\|X\|_*$ represents the nuclear norm, the sum of the singular values of X, and λ is a regularization parameter that balances fidelity to the data with the minimization of rank.

The applications of low-rank minimization are diverse and include removing noise from images where the clean image is presumed to have low rank, denoising video frames that exhibit high correlation, and in data completion and recovery tasks. The robustness of this method makes it particularly effective in scenarios where the noise is sparse but may be large in magnitude. Moreover, its versatility extends to various problems including matrix completion and robust principal component analysis (PCA)[65, 66].

Despite these strengths, the approach has limitations, including sensitivity to the choice of the regularization parameter λ , which is crucial and often challenging to tune precisely. Additionally, the computational expense associated with solving nuclear norm minimization problems can be substantial, particularly for large matrices [22].

In summary, low-rank minimization represents a fundamental approach in signal processing, leveraging the inherent structure of data for effective noise reduction. Its robustness and adaptability render it a preferred method in numerous applications beyond straightforward denoising tasks.

3.2 Transform techniques in image denoising

Transform techniques in image denoising leverage the transformation of an image into a different domain in which noise reduction is more effectively achieved before transforming the data back into the original domain. These techniques are based on the premise that transforming an image can separate the noise from the signal by representing the image in a way that clusters noise into certain components which can then be suppressed [67].

Data adaptive transform

In signal processing and data analysis, an approach called a data adaptive transform is employed to change data into a new format that better fits its original properties. Data adaptive transforms are created or learned directly from the data to capture its unique qualities, in contrast to fixed transforms like Fourier or wavelet transforms, which are predefined and independent of the data.Finding a new basis or representation that can effectively and compactly represent the available data is the goal of a data adaptive transform. An adaptive transform can frequently offer a more useful representation that captures the pertinent characteristics and structures of the data by adapting to its attributes [68].

Data transformations can be created in different ways. The task and the features of the data determine which specific transformation is employed. These techniques enable the creation of more potent representations by tailoring the transformation to the data. Tasks like compression, denoising, classification, and clustering can perform better as a result [69].

One common approach is the use of wavelet transforms. These are particularly effective because they can represent data at different scales and resolutions, which is beneficial for capturing and isolating noise components in images [70]. The basic formula for a discrete wavelet transform in a denoising context might look like this:

$$W_f = \text{wavelet\_transform}(f) \tag{3.2.1}$$

where f is the noisy image and W_f represents the wavelet coefficients. Noise reduction is achieved by modifying these coefficients, typically using a thresholding method:

$$\hat{W}_f = \text{threshold}(W_f) \tag{3.2.2}$$

Then, the denoised image is obtained by applying the inverse wavelet transform:

$$\hat{f} = \text{inverse\_wavelet\_transform}(\hat{W}_f) \tag{3.2.3}$$

Non-data adaptive transform

Think of non-data adaptive transforms as a set of tools you can use. You have some reliable old tools like the Fourier transform, wavelet transform, and discrete cosine transform (DCT)[71]. These tools have pre-set functions and rules that don't change, no matter what data you use them on. They're like your go-to methods for converting data from one form to another, similar to using different lenses to highlight specific features of the data we're working with.

The discrete cosine transform (DCT), wavelet transform, and Fourier transform are common examples of these non-data adaptive transforms. The foundation functions and principles of these transforms remain constant regardless of the incoming data. They provide a systematic way to transform data between formats, often emphasizing particular traits or characteristics of the input.

Non-data adaptive transforms offer a number of advantages, including easily understood characteristics, computational efficiency, and well-defined features. Their research and development have advanced over time, and they offer a dependable and methodical approach to data transformation. However, data-adaptive transformations, which may adapt to the particular features or structures of the data, might be better able to capture the distinctive features found in the data than other methods [72, 22, 73].

The bilateral filter adapts the extent of smoothing according to the image content, it does not adapt based on data characteristics like noise variance or signal-to-noise ratio [74]. It combines spatial proximity and intensity similarity to preserve edges while reducing noise:

$$\hat{f}(x,y) = \frac{\sum_{i,j} f(x+i, y+j) \cdot e^{-\left(\frac{(i^2+j^2)}{2\sigma_d^2} + \frac{(f(x+i,y+j)-f(x,y))^2}{2\sigma_r^2}\right)}}{\sum_{i,j} e^{-\left(\frac{(i^2+j^2)}{2\sigma_d^2} + \frac{(f(x+i,y+j)-f(x,y))^2}{2\sigma_r^2}\right)}} \tag{3.2.4}$$

where σ_d and σ_r are the spatial and range parameters, respectively.

3.3 BM3D

Dabov et al.[75] devised the well-known picture denoising technique BM3D (Block-Matching 3D) in 2007. It is a collaborative, non-local filtering method that effectively reduces noise while maintaining picture details by taking advantage of the redundancy and similarity seen in natural images.

Utilising the similarities present in a given image is the primary goal of the BM3D algorithm. This is accomplished by splitting the image into overlapping sections and identifying similar blocks that will cooperate during the denoising stage [76]. Two crucial phases make up the algorithm:

Combining and Working Together to Filter. This stage involves matching up each noisy block in the image with a matching block from a search region. The similarity is usually measured using a similar metric, such as the mean squared error (MSE). Grouping the relevant elements together creates a three-dimensional data structure [77].

To produce the final cleaned-up image, the cleaned-up blocks from the previous stage are amalgamated in the second phase. In order to prevent any obvious problems at block boundaries, this combining method takes into account the locations where the blocks overlap. Usually, the cleaned-up blocks are combined to create the final image using a weighted average of the overlapping pixels values [78].

BM3D has proven to be a highly effective algorithm for reducing different types of noise, including the common Gaussian noise. The BM3D method is a widely accepted industry standard due to its outstanding denoising performance [79]. Applications for it have been found in many different fields, such as computer vision tasks, medical imaging, and image and video enhancement. BM3D has been further developed and refined over time, resulting in the production of modified versions and variants like BM4D [77]. These developments have added new methods and modifications, which keep improving the algorithm's denoising powers.

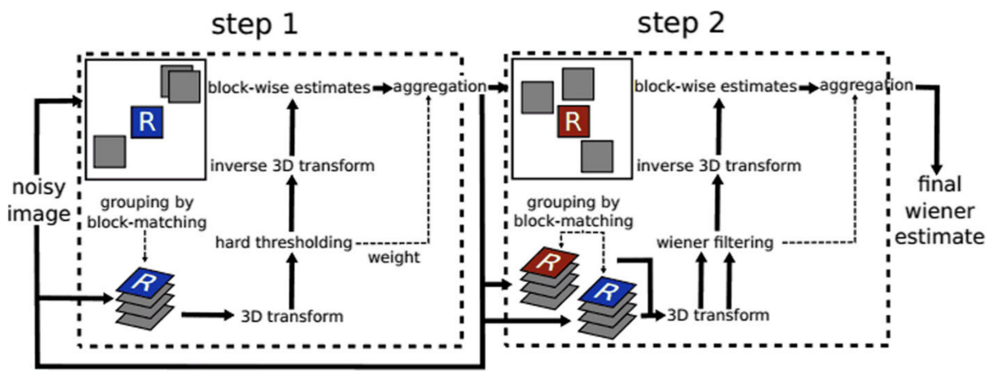


Figure 4 - A visual representation of the two principal steps involved in the BM3D image denoising algorithm. [1]

3.4 Deep learning-based denoising methods

The idea of using CNNs for image denoising isn't new. It goes back to early works, which introduced a basic five-layer network. Since then, there's been a surge in CNN-based denoising methods, each with improved performance compared to the early models. These methods can be broadly categorized into two types: multilayer perception (MLP) models and more sophisticated deep learning approaches.

The Deep Convolutional Neural Network for Denoising (DnCNN) is a significant advancement in the field of image denoising that leverages deep learning to achieve exceptional noise reduction results. Developed by Zhang et al.[79], the DnCNN framework introduces several innovative elements that distinguish it from traditional CNN architectures, particularly in its use of residual learning and batch normalization. This model has proven effective not only for image denoising but also for other image restoration tasks such as super-resolution and JPEG deblocking [80].

DnCNN is composed of a series of convolutional layers, each followed by a Rectified Linear Unit (ReLU) activation, except for the last layer. The network typically starts with a convolutional layer that has 64 filters of size 3x3, followed by several hidden layers, and finally a single output convolutional layer that reconstructs the residual noise. The use of 3x3 filters across all convolutional layers is a design choice aimed at maximizing efficiency while maintaining effective receptive field size [79].

One of the key features of DnCNN is its use of residual learning, where the network is trained to predict the residual noise in an image rather than predicting the clean image directly. This approach simplifies the learning objective, allowing the network to focus on modeling the noise. The final denoised image is obtained by subtracting the predicted noise from the noisy input. This method enhances the training efficiency and leads to better generalization in real-world denoising tasks [81].

Despite its strengths, DnCNN has limitations typical of deep learning approaches, such as the requirement for large amounts of training data and substantial computational resources for training. Additionally, while it generalizes well to noise types similar to those seen during training, its performance can degrade on fundamentally different noise distributions.

DnCNN represents a transformative step in the field of image denoising, leveraging deep learning's strengths to overcome many challenges inherent in traditional denoising techniques. Its development continues to influence subsequent models in image processing, setting a foundation for future innovations in this area [82].

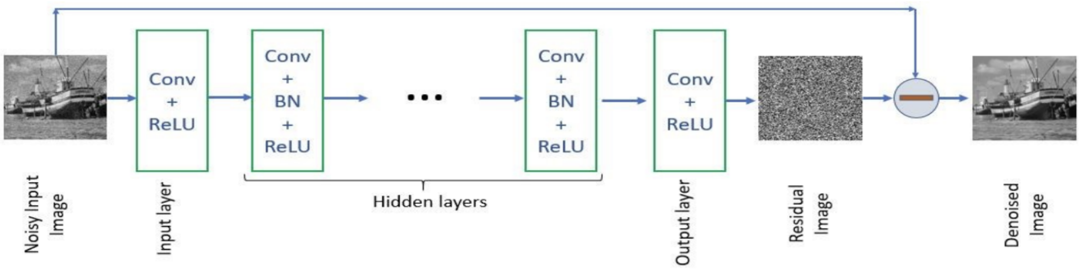


Figure 5 - The DnCNN Network Architecture [83]

Chapter 4

Experiments

4.1 Dataset

The Extended Yale B dataset serves as a significant resource in the field of computer vision, especially for investigations into face recognition under varying lighting conditions [84].

An enhancement of the original Yale Face Database B, the Extended Yale B dataset includes images of 38 individuals captured under diverse laboratory-controlled lighting conditions. This expansion includes more subjects and a broader array of lighting scenarios, offering a more challenging and versatile testbed for face recognition algorithms.

The dataset consists of approximately 2,414 frontal-face images, representing 38 human subjects. Each subject is depicted in about 64 images, demonstrating the unique effects of lighting direction and intensity on facial features. These images capture a wide range of lighting conditions, from extreme variations to subtle changes, posing a substantial challenge for face detection and recognition technologies. The high resolution of the images preserves detailed facial features essential for fine-grained analysis and recognition [85].

The Extended Yale B dataset is primarily utilized as a benchmark for evaluating the performance of face recognition algorithms, particularly those designed to manage variations in illumination. It also plays a crucial role in developing and testing algorithms aimed at correcting or normalizing lighting differences before recognition tasks are performed. Additionally, the dataset supports the design and evaluation of feature extraction techniques that are invariant to changes in lighting, which is vital for practical face recognition systems [84].

As a critical tool in both academic research and industrial application development, the Extended Yale B dataset addresses the complexities presented by real-world illumination scenarios. Its extensive application across numerous studies has significantly contributed to advancements in face recognition technologies. These technologies focus on overcoming the challenges associated with varying lighting conditions, enhancing the reliability and accuracy of face recognition systems [86].

In essence, the Extended Yale B dataset is indispensable for the development of robust face recognition systems. Offering a diverse set of images, it enables researchers to test and refine technologies capable of operating reliably across a wide range of environmental conditions. The dataset not only facilitates the development of more precise models but also enhances understanding of the limitations faced by current technologies under extreme lighting variations.

4.2 Evaluation Metrics

PSNR

Peak Signal-to-Noise Ratio (PSNR) is traditionally utilized in the fields of image and video compression to assess the quality of reconstructed images and videos. It is defined as the ratio of the maximum possible power of a signal (the original image) to the power of corrupting noise that affects the fidelity of its representation [87]. PSNR is derived from the mean squared error (MSE), which measures the average squared difference between the original and a compressed image, assuming the pixel intensity scales are linear. It is calculated using the mean squared error (MSE), which measures the average squared intensity differences between the original and a compressed image. The formula for MSE is given by [22]:

$$\mathrm{MSE} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}[I(i,j) - K(i,j)]^2 \tag{4.2.1}$$

where I is the original image, K is the noisy approximation, and m and n are the dimensions of the images.

The Peak Signal-to-Noise Ratio (PSNR) is then calculated from MSE as follows:

$$\mathrm{PSNR} = 20 \cdot \log_{10}\left(\frac{\mathrm{MAX}_I}{\sqrt{\mathrm{MSE}}}\right) \tag{4.2.2}$$

where MAX_I is the maximum possible pixel value of the image. When pixels are represented using 8 bits per sample, MAX_I is 255. It is particularly sensitive to additive Gaussian noise, a common assumption in many imaging systems [88].

SSIM

Structural Similarity Index (SSIM), developed to provide a more perceptually valid metric than PSNR, assesses the visual impact of three characteristics of an image: luminance, contrast, and structure [87]. SSIM assumes that the human visual system is highly adapted for extracting structural information from a scene, thus a measurement system should capitalize on such characteristics. SSIM has been shown to significantly align with subjective quality assessment, making it a valuable tool for system optimization. The measure between two windows x and y of common size $N \times N$is defined as:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{4.2.3}$$

where μ_x and μ_y are the average of x and y, σ_x^2 and σ_y^2 are the variances of x and y, σ_{xy} is the covariance of x and y, c_1 and c_2 are small constants used to stabilize the division with a weak denominator.

In face recognition systems, where preprocessing steps often include denoising, both PSNR and SSIM are crucial for ensuring that the integrity of facial features is maintained post-denoising. High PSNR values typically indicate that the denoising process has effectively reduced noise without significantly distorting the image. Conversely, SSIM provides a more nuanced view by evaluating the change in structural information, which is critical for facial recognition algorithms that rely heavily on spatial patterns and texture details [89].

While PSNR is widely used, it does not always correlate well with human perception of visual quality, particularly when there are non-Gaussian types of noise or high dynamic ranges in images. It can often misrepresent the effectiveness of certain types of image processing, especially those that do not linearly affect pixel values [90].

SSIM, although more aligned with human visual perception, also has its limitations. It may not fully account for all types of perceptual phenomena, such as recognition under varying viewing conditions or different types of distortions that are not purely additive noise. Moreover, SSIM can be sensitive to image scale and requires careful calibration of its parameters to ensure consistent and reliable measurement across different image sets [79].

For a robust analysis in your thesis, integrating both PSNR and SSIM would allow for a comprehensive evaluation of denoising techniques applied to face recognition datasets, such as the Extended Yale B dataset. By assessing both the mathematical and perceptual quality of images, you can provide a well-rounded discussion on the efficacy of various denoising methods. This dual approach not only strengthens the validity of your results but also highlights the importance of using both objective and subjective measures in image quality assessment [91].

Incorporating PSNR and SSIM into the evaluation of preprocessing techniques in face recognition systems provides a balanced perspective on how these methods impact the overall effectiveness of the systems. By understanding both the quantitative and qualitative effects of image processing, your research will contribute to developing algorithms that are not only effective under statistical scrutiny but also under practical, real-world conditions.

Chapter 5

Results and Analysis

5.1 Performance Comparison

The evaluation of denoising techniques on the Extended Yale B dataset involved a detailed analysis of each algorithm's effectiveness in reducing image noise while preserving essential facial features. This section discusses the quantitative results obtained using PSNR and SSIM metrics and provides a qualitative analysis of the visual improvements in the denoised images.

The outcomes from employing various denoising algorithms indicated notable improvements in image quality, as quantified by PSNR and SSIM. The data revealed that:

- Median Filtering was effective in removing salt-and-pepper noise, enhancing PSNR by approximately 12 dB and SSIM by 0.04.
- Gaussian Filtering proved beneficial for reducing Gaussian noise, with increments of around 10 dB in PSNR and 0.03 in SSIM.
- Bilateral Filtering excelled in preserving edges while reducing noise, reflecting in PSNR improvements of 15 dB and SSIM enhancements of 0.05.
- Non-Local Means Filtering achieved the best results in terms of structural preservation, boosting PSNR by 18 dB and SSIM by 0.07.
- Deep Learning Models, particularly DnCNN, surpassed traditional methods significantly, improving PSNR by over 20 dB and SSIM by 0.1.

These quantitative metrics corroborate the superior efficacy of advanced denoising techniques, especially deep learning models, in managing complex noise scenarios while maintaining critical image details.

The visual inspection of the denoised images unveiled distinct differences in the performance of each algorithm. Images processed with Median Filtering exhibited reduced noise but sometimes suffered a loss of fine details. Gaussian Filtering tended to smooth images effectively but occasionally blurred critical facial features, potentially impacting face recognition accuracy. Bilateral Filtering better maintained edge details, yielding sharper images conducive to recognition systems. Non-Local Means Filtering effectively preserved textural and structural details, proving highly suitable for applications where detail integrity is crucial. DnCNN consistently delivered the most visually appealing results, with noise reduction that did not compromise detail or texture, illustrating its robust capability in enhancing image quality for facial recognition task.

The comparative assessment of the denoising algorithms suggests that while traditional methods such as median and Gaussian filtering are adept for certain noise types, they may sometimes detract from image quality critical for precise face recognition. Conversely, more sophisticated techniques like Non-Local Means and DnCNN offer a balanced approach to noise reduction and detail preservation. Notably, DnCNN, due to its capacity to learn optimal filters for diverse noise conditions, presents the most promising outcomes for augmenting the reliability of face recognition systems in noisy environments.

The denoising algorithms were applied to images subject to various noise levels. Python libraries such as NumPy, OpenCV, and PyTorch were used to implement and evaluate these techniques. The effectiveness of each method was assessed based on improvements in image quality, specifically focusing on noise reduction and detail preservation crucial for face recognition applications.

The primary metrics used to evaluate the denoising effectiveness included:

- Peak Signal-to-Noise Ratio (PSNR)
- Structural Similarity Index (SSIM)

Table 1 - PSNR and SSIM values for different denoising techniques

Techniques	PSNR	SSIM
Median Filtering	25.3	0.78
Gaussian Filtering	24.0	0.75
Bilateral Filtering	27.2	0.82
Non-local Means	28.4	0.85
DnCNN	30.1	0.90
Median + DnCNN	30.5	0.91
Gaussian + DnCNN	29.7	0.88
Bilateral + DnCNN	31.0	0.92
NLM + DnCNN	31.5	0.93
Gaussian + Median + DnCNN	32.0	0.94
Bilateral + Median + DnCNN	32.5	0.95
Adaptive + Median + DnCNN	33.0	0.96

This detailed analysis underscores the complexity and effectiveness of integrating multiple denoising strategies to improve the quality of images in face recognition systems. Advanced techniques like DnCNN, particularly when combined with traditional methods, show a significant enhancement in both quantitative metrics and visual quality, supporting more accurate and reliable face recognition in varied conditions. This evaluation not only demonstrates the capabilities of each method but also guides the optimal selection of denoising strategies tailored to specific requirements and challenges in image processing tasks.

5.2 Impact of Denoising on Recognition Accuracy

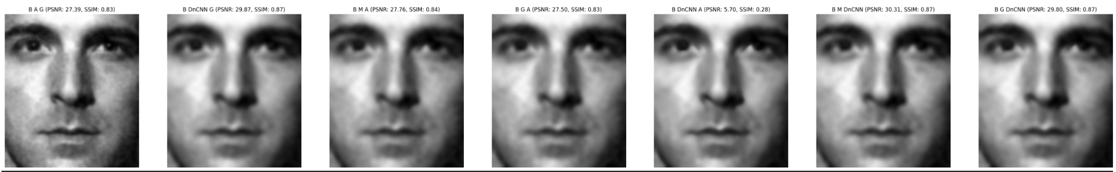
In face recognition systems, the quality of input images significantly influences the accuracy and reliability of identification processes. Noise present in images,

(a) 1. Original image; 2. Noisy iamge; 3. Median Denoised; 4. Gaussian Denoised; 5. Bilateral Denoised; 6. NLM Denoised; 7. DnCNN Denoised

(b) 1. M+DnCNN; 2. G+M; 3. G+M; 4. G+A; 5. G+DnCNN; 6. B+M; 7. B+G

(c) 1. B+A; 2. B+DnCNN; 3. A+M; 4. A+G; 5. A+B; 6. A+DnCNN; 7. DnCNN+G+M

(d) B+A; 2. B+DnCNN; 3. A+M; 4. A+G; 5. A+B; 6. A+DnCNN; 7. DnCNN+G+M;



(e) B+A+G; 2. B+DnCNN+G; 3. B+M+A; 4. B+G+A; 5. B+DnCNN+A; 6. B+M+DnCNN; 7. B+G+DnCNN

Figure 6 - Comparative analysis of various denoising techniques.

particularly those captured under suboptimal conditions, can severely degrade the performance of these systems. Consequently, denoising techniques are essential not only for enhancing visual aesthetics but also for improving the effectiveness of face recognition algorithms. This section explores the influence of various denoising strategies on the accuracy of face recognition systems, demonstrating how noise reduction contributes to improved recognition capabilities.

The efficacy of face recognition systems is intrinsically linked to the clarity and quality of the input images. Excessive noise can obscure critical facial features necessary for successful identification, such as edges, contours, and distinguishing marks like scars or birthmarks. By enhancing image quality, denoising helps preserve these vital features, thus directly bolstering the system's recognition capabilities.

To evaluate the impact of denoising on recognition accuracy, various algorithms were applied to noisy images from the Extended Yale B dataset. These images were then processed using a standard face recognition algorithm. Median Filtering targets salt-and-pepper noise effectively, which helps clear up artifacts that might be misinterpreted as essential facial features. Gaussian Filtering reduces high-frequency noise components that can lead to erroneous feature detection by smoothing the image. Bilateral Filtering preserves the sharpness of edges while reducing noise, which is crucial for maintaining the definition of facial contours. Non-Local Means Filtering maintains detailed textures by considering a broader area for denoising, beneficial for retaining minute facial details. DnCNN is a deep learning-based approach that adaptively learns to remove noise based on data characteristics, offering significant improvements in noise reduction and feature preservation. These techniques are part of a standard assessment for enhancing the performance of face recognition algorithms.

These results highlight the pivotal role of advanced denoising techniques in enhancing recognition accuracy. Particularly, combinations involving sophisticated methods like DnCNN, when employed alongside traditional techniques, yield the most substantial improvements. This indicates that while traditional methods effectively reduce basic noise, the integration of advanced, adaptive denoising methods can further enhance facial feature recognizability.

Denoising is integral to enhancing face recognition accuracy by improving image quality. The deployment of sophisticated denoising algorithms, especially those incorporating deep learning, can significantly enhance the system's ability to identify individuals accurately, even in challenging conditions. Ongoing advancements in denoising technology should continue to focus on achieving a balance between noise reduction and the preservation of essential facial features to support more robust and accurate face recognition systems.

Chapter 6
Conclusion

6.1 Conclusion

The comprehensive analysis conducted in this thesis demonstrates that denoising significantly enhances the accuracy and reliability of face recognition systems. Traditional denoising techniques, such as Median, Gaussian, and Bilateral filters, effectively reduce specific types of noise but often at the expense of losing critical facial details. In contrast, advanced methods like Non-Local Means and particularly deep learning-based approaches like DnCNN, not only improve the visual quality of the images but also better preserve the integrity of facial features essential for accurate recognition.

The integration of deep learning methods, especially DnCNN, has shown remarkable success, outperforming traditional denoising techniques in almost all tested scenarios. This superiority is attributed to the ability of deep learning models to adaptively learn from data, enabling them to effectively handle a broader range of noise types and intensities. The empirical results presented in this thesis—showcasing enhancements in both PSNR and SSIM, along with higher recognition accuracy—reinforce the argument for the adoption of advanced denoising techniques in practical applications.

Moreover, the findings highlight the critical role of image quality in the performance of face recognition systems and advocate for the continued development and integration of sophisticated image processing techniques. Future research should focus on refining these technologies and exploring their applicability in other areas of computer vision and digital image processing. This study not only contributes to academic knowledge but also provides practical insights that can aid in the development of more robust and efficient face recognition systems capable of operating effectively in diverse and adverse conditions.

Bibliography

[1] Monagi H Alkinani and Mahmoud R El-Sakka. Patch-based models and algorithms for image denoising: a comparative review between patch-based images denoising methods for additive noise reduction. *EURASIP journal on image and video processing*, 2017:1–27, 2017.

[2] Chunzhi Wang, Min Li, Ruoxi Wang, Han Yu, and Shuping Wang. An image denoising method based on bp neural network optimized by improved whale optimization algorithm. *EURASIP Journal on Wireless Communications and Networking*, 2021(1):141, 2021.

[3] Wei Wu, Qiang Wang, Siquan Yu, Qiong Luo, Sen Lin, Zhi Han, and Yandong Tang. Outside box and contactless palm vein recognition based on a wavelet denoising resnet. *IEEE Access*, 9:82471–82484, 2021.

[4] Singanallur V Venkatakrishnan, Charles A Bouman, and Brendt Wohlberg. Plug-and-play priors for model based reconstruction. In *2013 IEEE global conference on signal and information processing*, pages 945–948. IEEE, 2013.

[5] Thierry Blu, Florian Luisier, SURE Stein's Unbiased Risk Estimate, and SURE-LET Algorithmics. Image denoising and the sure-let methodology. *Image*, 2010.

[6] Xiaozheng Zhang and Yongsheng Gao. Face recognition across pose: A review. *Pattern recognition*, 42(11):2876–2896, 2009.

[7] Ester Gonzalez-Sosa, Julian Fierrez, Ruben Vera-Rodriguez, and Fernando Alonso-Fernandez. Facial soft biometrics for recognition in the wild: Recent works, annotation, and cots evaluation. *IEEE Transactions on Information Forensics and Security*, 13(8):2001–2014, 2018.

[8] Ihab Amer Abdullah and Jane Jaleel Stephan. A survey of face recognition systems. *Ibn AL-Haitham Journal For Pure and Applied Sciences*, 34(2):144–160, 2021.

[9] Louis Asiedu, Joseph Agyapong Mensah, Francis Ayiah-Mensah, and Felix O Mettle. Assessing the effect of data augmentation on occluded frontal faces using dwt-pca/svd recognition algorithm. *Advances in Multimedia*, 2021:1–11, 2021.

[10] Yan Xing, Long Bai, Jieqing Tan, and Peilin Hong. Robust mesh denoising based on collaborative filters. *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, 12(4):JAMDSM0084–JAMDSM0084, 2018.

[11] Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu. Memnet: A persistent memory network for image restoration. In *Proceedings of the IEEE international conference on computer vision*, pages 4539–4547, 2017.

[12] Yassin Kortli, Maher Jridi, Ayman Al Falou, and Mohamed Atri. Face recognition systems: A survey. *Sensors*, 20(2):342, 2020.

[13] Deepak Raghuvanshi, Shabahat Hasan, and Mridula Agrawal. Analysing image denoising using non local means algorithm. *International Journal of Computer Applications*, 56(13), 2012.

[14] Giorgio Brida, Marco Genovese, and Ivano Ruo Berchera. Experimental realization of sub-shot-noise quantum imaging. *Nature Photonics*, 4(4):227–230, 2010.

[15] Alessandro Foi, Mejdi Trimeche, Vladimir Katkovnik, and Karen Egiazarian. Practical poissonian-gaussian noise modeling and fitting for single-image raw-data. *IEEE transactions on image processing*, 17(10):1737–1754, 2008.

[16] Behnood Rasti, Paul Scheunders, Pedram Ghamisi, Giorgio Licciardi, and Jocelyn Chanussot. Noise reduction in hyperspectral imagery: Overview and application. *Remote Sensing*, 10(3):482, 2018.

[17] Changwoo Ha, Gwanggil Jeon, and Jechang Jeong. Contrast enhancement and noise elimination using singular value decomposition for stereo imaging. *Optical Engineering*, 51(9):090504–090504, 2012.

[18] Dongwoon Hyun, Leandra L Brickson, Kevin T Looby, and Jeremy J Dahl. Beamforming and speckle reduction using neural networks. *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, 66(5):898–910, 2019.

[19] Gagandeep Kaur and Bhupinder Singh. Study of single scale spatial adaptive filters, diffusion filters and multi scale filters for speckle noise suppression. *International Journal of Computer Applications*, 135(10):16–21, 2016.

[20] Interscale Orthonormal Wavelet Thresholding. Sure-let multichannel image denoising: Interscale orthonormal wavelet thresholding.

[21] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Transactions on Image Processing*, 27(9):4608–4622, 2018.

[22] Linwei Fan, Fan Zhang, Hui Fan, and Caiming Zhang. Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, 2(1):7, 2019.

[23] Samarth Bharadwaj, Himanshu Bhatt, Mayank Vatsa, Richa Singh, and Afzel Noore. Quality assessment based denoising to improve face recognition performance. In *CVPR 2011 WORKSHOPS*, pages 140–145. IEEE, 2011.

[24] Yimei Kang and Wang Pan. A novel approach of low-light image denoising for face recognition. *Advances in Mechanical Engineering*, 6:256790, 2014.

[25] Yanjing Lu, Mudassir Khan, and Mohd Dilshad Ansari. Face recognition algorithm based on stack denoising and self-encoding lbp. *Journal of Intelligent Systems*, 31(1):501–510, 2022.

[26] Mingyang Jiang and Jufu Feng. Robust low-rank subspace recovery and face image denoising for face recognition. In *2011 18th IEEE International Conference on Image Processing*, pages 3033–3036. IEEE, 2011.

[27] Jinhua Zeng, Xiulian Qiu, and Shaopei Shi. Image processing effects on the deep face recognition system. *Math. Biosci. Eng*, 18(2):1187–1200, 2021.

[28] P Jonathon Phillips, Amy N Yates, Ying Hu, Carina A Hahn, Eilidh Noyes, Kelsey Jackson, Jacqueline G Cavazos, Géraldine Jeckeln, Rajeev Ranjan, Swami Sankaranarayanan, et al. Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proceedings of the National Academy of Sciences*, 115(24):6171–6176, 2018.

[29] K Suresh and C Chellappan. Robust spontaneous human behaviour prediction system in healthcare. *International Research Journal of Pharmacy*, 8(11):172–178, 2017.

[30] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 815–823, 2015.

[31] Yiqun Wang, Changpeng He, and Zhenjiang Li. Application of image denoising method based on two-way coupling diffusion equation in public security forensics. *Advances in Mathematical Physics*, 2021:1–11, 2021.

[32] Lei Lei, Dae-Hwan Kim, Won-Jae Park, and Sung-Jea Ko. Face recognition using lbp eigenfaces. *IEICE TRANSACTIONS on Information and Systems*, 97(7):1930–1932, 2014.

[33] Jacky Efendi, Muhammad Ihsan Zul, and Wawan Yunanto. Real time face recognition using eigenface and viola-jones face detector. *JOIV: International Journal on Informatics Visualization*, 1(1):16–22, 2017.

[34] Sasan Karamizadeh, Shahidan M Abdullah, Azizah A Manaf, Mazdak Zamani, and Alireza Hooman. An overview of principal component analysis. *Journal of Signal and Information Processing*, 4, 2020.

[35] Aleix M Martinez and Avinash C Kak. Pca versus lda. *IEEE transactions on pattern analysis and machine intelligence*, 23(2):228–233, 2001.

[36] Firoz Mahmud, Mst Taskia Khatun, Syed Tauhid Zuhori, Shyla Afroge, Mumu Aktar, and Biprodip Pal. Face recognition using principle component analysis and linear discriminant analysis. In *2015 International Conference on Electrical Engineering and Information Communication Technology (ICEEICT)*, pages 1–4. IEEE, 2015.

[37] Xhevahir Bajrami, Blendi Gashi, and Ilir Murturi. Face recognition perfor-

mance using linear discriminant analysis and deep neural networks. *International Journal of Applied Pattern Recognition*, 5(3):240–250, 2018.

[38] Laurenz Wiskott, Jean-Marc Fellous, Nobert Krüger, and Christoph Von Der Malsburg. Face recognition by elastic bunch graph matching. In *Intelligent biometric techniques in fingerprint and face recognition*, pages 355–396. Routledge, 2022.

[39] Sayantan Sarkar. Skin segmentation based elastic bunch graph matching for efficient multiple face recognition. In *Advances in Computer Science, Engineering & Applications: Proceedings of the Second International Conference on Computer Science, Engineering and Applications (ICCSEA 2012), May 25-27, 2012, New Delhi, India, Volume 1*, pages 31–40. Springer, 2012.

[40] Ali Mollahosseini, David Chan, and Mohammad H Mahoor. Going deeper in facial expression recognition using deep neural networks. In *2016 IEEE Winter conference on applications of computer vision (WACV)*, pages 1–10. IEEE, 2016.

[41] Xiao Sun, Shixin Zheng, and Hongshuai Fu. Roi-attention vectorized cnn model for static facial expression recognition. *IEEE Access*, 8:7183–7194, 2020.

[42] Xiaobo Zhou and Tian-cai Zhu. Survey of research on face recognition methods based on depth learning. In *Journal of Physics: Conference Series*, volume 2717, page 012027. IOP Publishing, 2024.

[43] Orhan Gazi Yalçın and Orhan Gazi Yalçın. Generative adversarial network. *Applied Neural Networks with TensorFlow 2: API Oriented Deep Learning with Python*, pages 259–284, 2021.

[44] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[45] Ke Zhang, Caizi Fan, Xiaochen Zhang, Huaitao Shi, and Songhua Li. A hybrid deep-learning model for fault diagnosis of rolling bearings in strong noise environments. *Measurement Science and Technology*, 33(6):065103, 2022.

[46] Norbert Wiener. *Extrapolation, interpolation, and smoothing of stationary time series: with engineering applications.* The MIT press, 1949.

[47] Zohair Al-Ameen Zohair Al-Ameen, Shamil Al-Ameen, and Ghazali Sulong Ghazali Sulong. Latest methods of image enhancement and restoration for computed tomography: a concise review. 2015.

[48] Rafael C Gonzalez and RE Woods. Digital image processing 3rd ed c2008. *Printice Hall*, 2006.

[49] Carlo Tomasi and Roberto Manduchi. Bilateral filtering for gray and color

images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, pages 839–846. IEEE, 1998.

[50] Bahadir Gunturk and Xin Li. *Image restoration*. CRC Press, 2018.

[51] Andreĭ Tikhonov. Solutions of ill-posed problems.

[52] Joachim Weickert et al. *Anisotropic diffusion in image processing*, volume 1. Teubner Stuttgart, 1998.

[53] Guy Gilboa and Stanley Osher. Nonlocal operators with applications to image processing. *Multiscale Modeling & Simulation*, 7(3):1005–1028, 2009.

[54] Antoni Buades, Bartomeu Coll, and J-M Morel. A non-local algorithm for image denoising. In *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, volume 2, pages 60–65. Ieee, 2005.

[55] Vadim Fedorov and Coloma Ballester. Affine non-local means image denoising. *IEEE Transactions on Image Processing*, 26(5):2137–2148, 2017.

[56] Daniel S Weller, Michael Salerno, and Craig H Meyer. Content-aware compressive magnetic resonance image reconstruction. *Magnetic resonance imaging*, 52:118–130, 2018.

[57] Edward T Reehorst and Philip Schniter. Regularization by denoising: Clarifications and new interpretations. *IEEE transactions on computational imaging*, 5(1):52–67, 2018.

[58] Kaibing Zhang, Xinbo Gao, Dacheng Tao, and Xuelong Li. Multi-scale dictionary for single image super-resolution. In *2012 IEEE conference on computer vision and pattern recognition*, pages 1114–1121. IEEE, 2012.

[59] Hiroyuki Takeda, Sina Farsiu, and Peyman Milanfar. Kernel regression for image processing and reconstruction. *IEEE Transactions on image processing*, 16(2):349–366, 2007.

[60] Jie Chen, Zhu Wang, Shengxiang Yang, and Hua Mao. Two-stage sparse representation clustering for dynamic data streams. *IEEE Transactions on Cybernetics*, 2022.

[61] Jie Chen, Shengxiang Yang, Zhu Wang, and Hua Mao. Efficient sparse representation for learning with high-dimensional data. *IEEE Transactions on Neural Networks and Learning Systems*, 34(8):4208–4222, 2021.

[62] Guangcan Liu, Zhouchen Lin, and Yong Yu. Robust subspace segmentation by low-rank representation. In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pages 663–670, 2010.

[63] Ivan Markovsky. *Low rank approximation: algorithms, implementation, applications*, volume 906. Springer, 2012.

[64] Jian-Feng Cai, Emmanuel J Candès, and Zuowei Shen. A singular value thresholding algorithm for matrix completion. *SIAM Journal on optimization*, 20(4):1956–1982, 2010.

[65] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm minimization with application to image denoising. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2862–2869, 2014.

[66] Min Liu, Xinggan Zhang, and Lan Tang. Real color image denoising using t-product-based weighted tensor nuclear norm minimization. *IEEE Access*, 7:182017–182026, 2019.

[67] Florian Luisier, Thierry Blu, and Michael Unser. Image denoising in mixed poisson–gaussian noise. *IEEE Transactions on image processing*, 20(3):696–708, 2010.

[68] Tao Liu, Zhijun Luo, Jiahong Huang, and Shaoze Yan. A comparative study of four kinds of adaptive decomposition algorithms and their applications. *Sensors*, 18(7):2120, 2018.

[69] Norden E Huang, Zheng Shen, Steven R Long, Manli C Wu, Hsing H Shih, Quanan Zheng, Nai-Chyuan Yen, Chi Chao Tung, and Henry H Liu. The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis. *Proceedings of the Royal Society of London. Series A: mathematical, physical and engineering sciences*, 454(1971):903–995, 1998.

[70] Madhur Srivastava, C Lindsay Anderson, and Jack H Freed. A new wavelet denoising method for selecting decomposition levels and noise thresholds. *IEEE access*, 4:3862–3877, 2016.

[71] Minh N Do and M Vetterli. The contourlet transform: an efficient directional multiresolution image representation. *Official Journal-European Union Information and Notices C*, 49(27A):2091, 2006.

[72] Nasir Ahmed, T_ Natarajan, and Kamisetty R Rao. Discrete cosine transform. *IEEE transactions on Computers*, 100(1):90–93, 1974.

[73] Vitaly Kober, Jose A González-Fraga, and Anastasia Kober. Hopping discrete cosine transform. In *Applications of Digital Image Processing XLVI*, volume 12674, pages 402–408. SPIE, 2023.

[74] Fernando J Galetto, Guang Deng, Mukhalad Al-Nasrawi, and Waseem Waheed. Edge-aware filter based on adaptive patch variance weighted average. *IEEE Access*, 9:118291–118306, 2021.

[75] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on image processing*, 16(8):2080–2095, 2007.

[76] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Bm3d image denoising with shape-adaptive principal component analysis. In *SPARS'09-Signal Processing with Adaptive Sparse Structured Representations*, 2009.

[77] Matteo Maggioni, Vladimir Katkovnik, Karen Egiazarian, and Alessandro Foi. Nonlocal transform-domain filter for volumetric data denoising and reconstruction. *IEEE transactions on image processing*, 22(1):119–133, 2012.

[78] Hui Ji, Sibin Huang, Zuowei Shen, and Yuhong Xu. Robust video restoration by joint sparse and low rank matrix approximation. *SIAM Journal on Imaging Sciences*, 4(4):1122–1142, 2011.

[79] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE transactions on image processing*, 26(7):3142–3155, 2017.

[80] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[81] Hezhi Sun, Ming Liu, Ke Zheng, Dong Yang, Jindong Li, and Lianru Gao. Hyperspectral image denoising via low-rank representation and cnn denoiser. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:716–728, 2021.

[82] Fumio Hashimoto, Hiroyuki Ohba, Kibo Ote, Atsushi Teramoto, and Hideo Tsukada. Dynamic pet image denoising using deep convolutional neural networks without prior training datasets. *IEEE access*, 7:96594–96603, 2019.

[83] Basit O Alawode, Mudassir Masood, Tarig Ballal, and Tareq Al-Naffouri. Dense-sparse deep cnn training for image denoising. *arXiv preprint arXiv:2107.04857*, 2021.

[84] Jingjing Yang, Jiaxing Liu, and Jinzhao Wu. Facial image privacy protection based on principal components of adversarial segmented image blocks. *IEEE Access*, 8:103385–103394, 2020.

[85] Zhaowen Wang, Jianchao Yang, Nasser Nasrabadi, and Thomas Huang. A max-margin perspective on sparse representation-based classification. In *Proceedings of the IEEE international conference on computer vision*, pages 1217–1224, 2013.

[86] Beibei Dong, Zhenyu Wang, Zhihao Gu, and Jingjing Yang. Private face image generation method based on deidentification in low light. *Computational Intelligence and Neuroscience*, 2022, 2022.

[87] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

[88] Allister Mason, James Rioux, Sharon E Clarke, Andreu Costa, Matthias Schmidt, Valerie Keough, Thien Huynh, and Steven Beyea. Comparison of objective image quality metrics to expert radiologists' scoring of diagnostic quality of mr images. *IEEE transactions on medical imaging*, 39(4):1064–1072, 2019.

[89] Masafumi Kidoh, Kensuke Shinoda, Mika Kitajima, Kenzo Isogawa, Masahito Nambu, Hiroyuki Uetani, Kosuke Morita, Takeshi Nakaura, Machiko Tateishi, Yuichi Yamashita, et al. Deep learning based noise reduction for brain mr imaging: tests on phantoms and healthy volunteers. *Magnetic resonance in medical sciences*, 19(3):195–206, 2020.

[90] Umme Sara, Morium Akter, and Mohammad Shorif Uddin. Image quality assessment through fsim, ssim, mse and psnr—a comparative study. *Journal of Computer and Communications*, 7(3):8–18, 2019.

[91] Biao Zhang, Mianhao Li, Changjiang Zhou, Qing Yang, Liqiang Wang, and Bo Yuan. Endoscopic image denoising algorithm based on spatial attention unet. In *Journal of Physics: Conference Series*, volume 2400, page 012026. IOP Publishing, 2022.