

X-ray Image Denoising for Baggage Screening Using Learning Based Methods

A THESIS

Submitted by

VEMPULURU VENKATA POORNA SEKHAR

(M210262EC)

In partial fulfillment for the award of the degree of

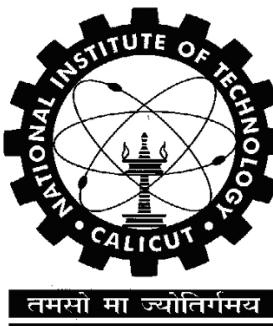
MASTER OF TECHNOLOGY

IN

**ELECTRONICS AND COMMUNICATION ENGINEERING
(Signal Processing)**

Under the guidance of

Dr.Deepthi P.P.



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
NATIONAL INSTITUTE OF TECHNOLOGY CALICUT
NIT CAMPUS PO, KOZHIKODE, KERALA, INDIA 673601**

July 2023

ACKNOWLEDGEMENTS

First and foremost, I would like to extend a deep sense of gratitude to my project advisor Dr. Deepthi P. P., Professor and Former Head of the Department of Electronics and Communication Engineering, and Dr. Renu M Rameshan, Lead Research Scientist, Vehant Technologies Private Limited, Noida, for their very valuable guidance, monitoring and motivation throughout the course of this project. It was their immense knowledge, stimulating suggestions, patience and encouragement that helped me coordinate this project in a decent manner.

I would also like to acknowledge with much appreciation the crucial role of the Project evaluation committee members: Dr. Sudhish N. George, Dr. Anup Aprem and Dr. Abhilash., for their valuable advice and critical assessments during the presentations.

I am grateful to Dr. Jaikumar M. G., Head of the Department of Electronics and Communication Engineering, and to the faculties of Electronics and Communication Department for all their help in completing the project. I also thank my family and my fellow classmates for their support throughout the work.

-Vempuluru Venkata Poorna Sekhar

DECLARATION

I hereby declare that this submission is entirely original work of mine, and that to the best of my knowledge and belief, it does not contain any text that has been accepted for the awarding of any other degree or diploma by the university or other institution of higher learning, with the exception of instances where due attribution has been made in the text.

Place: NIT Calicut

Date: July 11, 2023

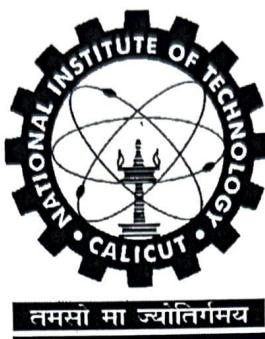
Signature: V.V Poorna Sekhar

Name: Vempuluru Venkata Poorna Sekhar

Reg.No: M210262EC

National Institute of Technology Calicut

Department of Electronics and Communication Engineering



CERTIFICATE

This is to certify that the Thesis entitled, "X-ray Image Denoising for Baggage Screening Using Learning Based Methods" submitted by Mr. Vempuluru Venkata Poorna Sekhar to the National Institute of Technology Calicut towards partial fulfillment of the requirements for the award of the Degree of Master of Technology in Electronics and Communication Engineering (Signal Processing) is a bonafide record of the work carried out by him under my supervision and guidance.

Guide:


Dr. Deepthi P. P.

Professor

ECED


Dr. Jaikumar M. G.

Head of the Department

ECED

Place: NIT Calicut

Date: July 11, 2023



TABLE OF CONTENTS

ABSTRACT	i
List of Abbreviations	ii
List of Important Symbols	iii
List of Figures	iv
List of Tables	v
1 INTRODUCTION	1
1.1 An Overview	1
1.2 Problem Definition	3
1.3 Motivation	3
1.4 Thesis Contribution	4
1.5 Thesis Organization	4
2 LITERATURE REVIEW	6
2.1 Traditional Methods	6
2.1.1 VST+BM3D	7
2.1.2 Non Local PCA	8
2.1.3 Limitations of Traditional Methods	8
2.2 Learning based Methods	9
2.2.1 Supervised Learning Method	10
2.2.2 Self Supervised Learning Method	11
2.2.3 GAN Based denoisers	12
2.3 X-ray imaging	12
2.4 Dataset	13

2.4.1	Fluorescence Microscopy dataset	13
2.4.2	Dual energy X-ray image dataset	13
2.5	Summary	14
3	METHODOLOGY FOLLOWED	15
3.1	Introduction	15
3.2	Blindspot Neural Network	15
3.3	Image Fusion Algorithms	16
3.3.1	Image Fusion using Local Spatial Information	17
3.3.2	Wavelet Based Fusion Algorithm	18
3.4	Summary	19
4	EXPERIMENTS AND RESULTS	20
4.1	Denoising Convolutional Neural Network (DnCNN)	20
4.2	Blindspot Neural Network:Denoising of Microscopy images	21
4.3	Experiments using pix2pix GAN	22
4.4	Blindspot Neural Network: Denoising of X-ray images	25
4.4.1	Evaluation Criteria	26
4.4.2	Results	26
4.4.3	Comparison Results	28
4.5	Image Fusion	28
4.6	Summary	29
5	BAGGAGE SCREENING WITH IMPROVED VISUALIZATION	32
5.1	Pseudo Coloring	32
5.2	Standard Deviation of single material patches	33
5.3	Summary	34
6	CONCLUSION	35
	PUBLICATIONS	36

““

ABSTRACT

The utilization of dual-energy X-ray detection technology in security inspection plays a crucial role in ensuring public safety and preventing crimes. However, the X-ray images generated in such security checks often suffer from substantial noise, degrading the image quality and hindering accurate judgments by security inspectors. Existing deep learning-based denoising methods have limitations, such as reliance on large training datasets and clean reference images, which are not readily available in security inspection scenarios. In this work, we addressed the denoising problem of X-ray images with a Poisson-Gaussian noise model, without requiring clean reference images for training.

To overcome these challenges, we employed the Blindspot neural network, which effectively removes noise from X-ray images. Additionally, we incorporated image fusion algorithms, leveraging image spatial information and wavelet-based fusion rules to enhance visualization. Experimental evaluations on a real X-ray image dataset demonstrated the effectiveness of our approach, achieving favorable BRISQUE scores across different baggage scenes. The denoised images obtained through our method, combined with image fusion, enable security inspectors to make more accurate judgments and enhance public safety in security inspection settings.

LIST OF ABBREVIATIONS

DL	Deep Learning
VST	Variance Stabilizing Transform
BM3D	Block Matching and 3D Filtering
NLPCA	Non Local Principal Component Analysis
PGPCA	Patch Gaussian PCA
DnCNN	Denoising Convolutional Neural Network
GAN	Generative Adversarial Network
BN	Batch Normalization
ReLU	Rectified Linear Unit
LowE	Low Energy Image
HighE	High Energy Image
BRISQUE	Blind/Referenceless Image Spatial Quality
BNN	Blindspot Neural Network
MSCN	Mean Subtracted Contrast Normalization
CTP	Combined Test Piece

LIST OF IMPORTANT SYMBOLS

x	Clean Image
y	Noisy Observation
v	Noise
$R(.)$	Residual Mapping Function
Θ	Parameters of the model
\hat{x}	Estimated output
μ_i	Mean of noisy pixel
$\hat{\sigma}_i^2$	Total Variance of noisy pixel
σ_i^2	Noise Variance

LIST OF FIGURES

2.1	DnCNN architecture	10
2.2	Normal working of neural network	11
2.3	Working of Blindspot neural network	11
3.1	Methodology followed.	15
4.1	Gaussian denoising Results for variance=0.1, Average PSNR=30.35, SSIM=0.93	20
4.2	Results of blind-spot network with Poisson-Gaussian noise model using FMD. Top row: 1) Noisy image, 2) Denoised image (from left to right). Bottom row: Corresponding zoom in versions.	21
4.3	Methodology followed to generate the noise model and denoise using GAN	22
4.4	Example of paired image	22
4.5	Left to right: Binary mask, Real image corresponding to binary mask, Generated image. Sorenson Distance = 0.0367	23
4.6	(a) Test image, (b) Generated noise model-background, (c) Cropped version of background	23
4.7	Examples of Noisy image samples after applying Gamma Distribution on binary masks	24
4.8	(a) Real Noisy Image, (b) Denoised Image	24
4.9	CTP image, Left to Right: Noisy image, Denoised version	27
4.10	BAG1, Left to Right: Noisy image, Denoised version	27
4.11	BAG2, Left to Right: Noisy image, Denoised version	27
4.12	Comparison between Traditional Methods and followed Deep Learning Method	28
4.13	Image Fusion using Local Spatial Information	30
4.14	Final Fused image after fusing 4.13a, 4.13b using wavelet transform.	31
5.1	Colored version of Figure 4.13e.	33

LIST OF TABLES

4.1	BRISQUE Scores of microscopy images before and after denoising	25
4.2	BRISQUE Scores of X-ray images before and after denoising	27
4.3	BRISQUE Scores of different baggage scenes before and after applying Denoising methods on X-ray image dataset and the best values are highlighted	28
5.1	Standard Deviation of Different Thicknesses in low energy CTP image	34
5.2	Standard Deviation of Different Thicknesses in high energy CTP image	34

CHAPTER 1

INTRODUCTION

1.1 An Overview

The application of X-ray detection technology is widespread, spanning across various domains such as medicine, industry, and security inspection. In the realm of security, X-ray technology is extensively utilized in locations like airports, customs checkpoints, and railway stations to combat criminal activities. Security personnel leverage the power of X-ray images as a means to determine the presence of prohibited items within packages. This enables them to make prompt and informed decisions, ensuring effective security measures are in place [1], [2]. However, the raw images obtained often contain noise, which can degrade image quality and make it difficult for inspectors to make accurate judgments. Therefore, denoising X-ray images is an important task. Image denoising technology reduces noise from images.

Noise is an undesirable by-product of an image capturing process that obstructs the desired information. Image noise may be caused by different intrinsic (i.e., sensor) and extrinsic (i.e., environment) conditions, which are often not possible to avoid in practical situations. As a result of noise, the original pixel values are sometimes replaced by random values, which can reduce the quality of the acquired image and result in information loss. Therefore, in low-level vision tasks and image processing, it is necessary to remove these noises from images. The process of removing such noises from images is known as image denoising. Image denoising is necessary for a variety of applications, including image restoration, image segmentation, and image classification, where recovering the original image content is essential for effective operation.

There are different types of noise that occur in an image. But Gaussian noise is commonly present in real-life images under well-lit conditions. Low-light images, however, suffer from Poisson-Gaussian noise, and it is very difficult to address this problem be-

cause Poisson noise is signal-dependent noise. One such example of low-light images is Fluorescence microscopy images. Fluorescence microscopy produces noisy images with weak signals affected by Gaussian and Poisson noise. Gaussian noise is evenly distributed, while Poisson noise occurs due to the uneven arrival of photons at a point. This noise combination hinders detailed visualization. Analyzing such data is challenging. In this case, obtaining clean training targets is difficult. Ground truth is often obtained by averaging multiple noisy images, but the noise persists. Averaging suppresses but does not eliminate noise.

Dual-energy X-ray imaging serves as a significant application scenario for addressing the challenges of Poisson-Gaussian noise in real-world images. In this context, X-ray imaging commonly combines Poisson (shot) noise and Gaussian (electronic) noise. X-ray emission and absorption are random processes that can be well modeled by the Poisson distribution, considering the proportional relationship between X-ray intensity and the probability of photons arriving at a point. Electronic noise arises from components, amplifiers, and digitization during the X-ray detector's readout process. Gaussian noise is often used to model electronic noise. This work aims to develop a denoising technique specifically designed for dual-energy X-ray imaging [3].

Furthermore, dual-energy security X-ray imaging plays a crucial role in threat detection and security screening. Capturing X-ray images at two different energy levels enables material discrimination and enhances the ability to detect various objects and substances. The purpose of the research is to overcome challenges such as overlapping materials, noise, and real-time processing requirements in dual-energy imaging systems. The utilization of image fusion algorithms and denoising techniques aims to improve visibility, discrimination, and image quality in fused images. These advancements contribute to the accurate identification of materials and objects, thereby enhancing threat detection capabilities and operational efficiency in critical security environments.

This work focuses on the study and application of deep learning methods in image denoising. Initial studies are conducted to perform denoising of microscopy images using blindspot neural network. The inclusion of microscopy images in this work is motivated by the fact that these images also exhibit the presence of Poisson-Gaussian noise. By

studying the denoising performance of the blindspot neural network on microscopy images, valuable insights can be gained regarding its effectiveness in handling the specific challenges posed by Poisson-Gaussian noise.

A detailed analysis of the performance of the blindspot network in denoising microscopy images provides a foundation for its potential applicability in the denoising of x-ray images where the noise is Poisson-Gaussian distributed, in general. Preprocessing techniques were applied to adapt the X-ray images for input into the blindspot network, and the denoising results obtained demonstrate its efficacy in mitigating Poisson-Gaussian noise in this context as well.

1.2 Problem Definition

This work aims at denoising of real-world images where the noise model is Poisson-Gaussian, using deep learning method. Further, a specific application scenario of baggage inspection through X-ray imaging is considered where the noise model follows Poisson-Gaussian distribution. Since the threat concealment strategies are highly sophisticated, the threat detection performance heavily depends on the performance of the denoising algorithm.

1.3 Motivation

The motivation for this project is to address the challenges posed by the Poisson-Gaussian noise model in security imaging. Traditional methods struggle to effectively handle the complex interaction between photon shot noise and electronic sensor noise. By developing a deep learning-based denoising method tailored to this noise model, the project aims to improve the quality and reliability of security imaging data. The evaluation of a specialized dataset contributes to advancements in denoising techniques, improving the quality and accuracy of security images for better decision-making in security-related tasks.

1.4 Thesis Contribution

The thesis contributions are outlined below:

- The thesis contributes by evaluating an existing deep learning method for image denoising on a security imaging dataset. The applied method demonstrates its effectiveness in reducing noise in security images, showcasing superior performance compared to traditional existing methods.
- Introduced a specific dataset consisting of security images, providing researchers and professionals in the field of security imaging with a specialized collection of images for denoising and subsequent analysis.
- In addition to denoising, the thesis explores the application of image fusion techniques on denoised security images. This contribution enhances the visualization capabilities of the images by integrating multiple sources of information, potentially improving the accuracy and interpretability of the results for security applications.
- Another contribution of the thesis is the application of pseudo-coloring on fused images. This approach assigns artificial colors to different intensities or features within the image, enhancing the visual representation and facilitating easier interpretation of threat item in the images.

1.5 Thesis Organization

The rest of this thesis is organised as follows. Chapter 2 focuses on the literature review, exploring traditional and learning-based denoising methods, including VST+BM3D, Non Local PCA, supervised learning methods, self-supervised learning methods, and GAN-based denoisers. The chapter also provides an overview of X-ray imaging and describes the datasets used for evaluation, including the fluorescence microscopy dataset and the dual-energy X-ray image dataset.

Chapter 3 presents the methodology followed in the research, introducing the Blindspot Neural Network as the deep learning-based denoising method. Additionally, image fu-

sion algorithms are discussed, including techniques based on local spatial information and wavelet-based fusion algorithms for dual-energy X-ray images.

Chapter 4 presents the experiments and results obtained from the denoising and fusion techniques. The Denoising Convolutional Neural Network (DnCNN), Blindspot Neural Network, and pix2pix GAN are evaluated on microscopy and X-ray images, and the results are compared using evaluation criteria. The chapter also includes a detailed analysis of the image fusion process and presents the outcomes.

Chapter 5 studies the applications of the developed techniques, specifically focusing on the pseudo-coloring of fused images. The concept and benefits of pseudo coloring are explained, and the HSI-based color transform is utilized to assign colors to grayscale levels in the fused images. The chapter also discusses the calculation of standard deviation to showcase the effectiveness of the denoising process.

Chapter 6 concludes the thesis by summarizing the key findings and contributions.

CHAPTER 2

LITERATURE REVIEW

In recent decades, a number of image denoising techniques have been developed using conventional techniques as well as learning-based methods. Many image processing and computer vision applications, including image denoising, have been revolutionized by deep learning using vast volumes of training data. Recently, it was demonstrated that deep learning-based denoisers give very promising results in reducing noise without introducing artifacts.

2.1 Traditional Methods

Many different methods have been developed to address denoising, with two primary challenges remaining: achieving the best possible removal of noise artifacts and preserving the important characteristics of an image, such as smooth features and edges. These methods are mainly divided into two categories namely, spatial domain filters and transform domain filters.

Spatial domain filters are effective in reducing additive noise. They can be classified as linear or nonlinear filters. The Gaussian filter, one of the earliest linear filters in the spatial domain, is widely used for denoising purposes. It operates by calculating a weighted average of the pixel's neighborhood. However, when applying Gaussian filtering, the edges in the image can become blurred or even disappear, leading to a decrease in contrast. The median filter is a non-linear filter in the spatial domain that effectively addresses denoising requirements while preserving edges more efficiently. It replaces the central pixel value with the median value of the pixels within a defined window. The choice of the window radius depends on the noise characteristics present in the image [5].

Transform domain filters have demonstrated remarkable performance in effectively removing additive noise. One of them is the wavelet transform. Wavelet transforms are

recognized as a crucial pre-processing tool in various domains such as image denoising, image compression, signal processing, pattern recognition, and computer vision applications. In the context of pre-processing medical and remote sensing images, wavelet transform-based methods have emerged as highly effective approaches [5].

Spatial domain filters are known for their ability to preserve edges effectively, but they often struggle to retain low-contrast details. On the other hand, transform domain filters excel in preserving low-contrast details while they may not be as successful in preserving edges. Consequently, a hybrid-domain approach combining both spatial and transform domains has been developed to address the preservation of both low-contrast details and edges. Some of them are Variance Stabilization Transform+BM3D and Non Local PCA methods.

2.1.1 VST+BM3D

The Variance-Stabilizing Transform (VST) is a preprocessing technique used to enhance the performance of image denoising algorithms. It aims to normalize the noise levels across different frequency bands or domains, making the noise characteristics more consistent and predictable for denoising algorithms. The first step is to transform the image into a different domain, such as the wavelet domain. After the transformation, the variance of the noise within each frequency subband or transformed domain is estimated. This step helps determine the noise characteristics within each frequency range accurately. Various methods can be employed to estimate the noise variance, such as statistical estimation techniques. Instead of directly dividing the values in each band by the estimated noise variance, the VST applies a specific mapping function that stabilizes the noise variance. By stabilizing the variance, the VST makes the noise more consistent and easier to handle for subsequent denoising algorithms [6].

BM3D stands for Block-Matching and 3D Filtering. It is a powerful image denoising algorithm that works by searching for similar patches in a noiseless image and then averaging the values of the corresponding patches in the noisy image. This effectively removes the noise while preserving the edges and details of the image. The first step is to divide the noisy image into overlapping patches. The size of the patches is typically

8×8 pixels. Next, a database of patches is created from a noiseless image. This database is created by dividing the noiseless image into overlapping patches in the same way as the noisy image. For each patch in the noisy image, BM3D searches for the most similar patches in the database. This is done by using a distance metric, such as the mean squared error. Once the most similar patches have been found, their values are averaged to create a denoised patch. The denoised patches are then stitched together to create a denoised image [7].

The normalized image that is obtained after VST is used as input to a BM3D denoising algorithm. The denoising algorithm will be able to remove the noise more effectively because the noise levels are now consistent across the image.

2.1.2 Non Local PCA

The non-local PCA denoising method involves several steps. Firstly, the noisy image is divided into patches of a specific size. These patches are collected to form a matrix, and then K-Means clustering is applied to group similar patches together. Each cluster represents a matrix of patches. Next, an iterative denoising process is performed for each cluster. Matrices U and V, which represent the estimated clean patches and their corresponding coefficients, respectively, are randomly initialized. These matrices are updated iteratively using specific equations until a stopping criterion is met. The updated matrices are used to generate denoised patches for each cluster. These denoised patches are then concatenated to form a matrix representing the collection of denoised patches. Finally, the denoised image is reconstructed by averaging the pixel estimates due to overlapping patches. This averaging process helps ensure smoothness and reduce artifacts in the final denoised image [8], [9].

2.1.3 Limitations of Traditional Methods

1. In the context of Non-local PCA and VST+BM3D, which are traditional methods for image denoising, it is important to consider their limitations. While these methods can be effective in reducing noise, they typically do not completely eliminate it.

They are designed to suppress the noise and enhance the visual quality of the image, but there may still be residual noise present after the denoising process.

2. One limitation of Non-local PCA and VST+BM3D is that they are based on assumptions about the noise distribution, such as the additive white Gaussian noise (AWGN) model. However, real-world noise in images, especially in certain domains like medical imaging or low-light photography, can deviate from the AWGN model. This can result in less-than-optimal denoising performance and the presence of residual noise artifacts.
3. Additionally, both Non-local PCA and VST+BM3D methods may introduce certain artifacts or distortions in the denoised image. These artifacts can be in the form of blurring, ringing effects, or loss of fine details and textures. The denoising process involves a trade-off between noise reduction and preserving image details, and sometimes these methods tend to smooth out the image excessively, leading to a loss of important visual information.
4. Furthermore, the performance of Non-local PCA and VST+BM3D can be dependent on parameter tuning. Selecting appropriate parameters and thresholds is crucial for achieving satisfactory denoising results. However, finding the optimal parameter settings can be challenging and time-consuming, especially when dealing with diverse types of images and noise characteristics.

Overall, while Non-local PCA and VST+BM3D are widely used and effective denoising methods, they have inherent limitations in terms of residual noise, artifacts, and parameter sensitivity.

2.2 Learning based Methods

Learning based methods try to learn the noise model from a collection of noisy images. Convolutional Neural Networks(CNNs) are recently used in many image denoising applications with successful performance. CNNs are a type of neural network that are commonly used for image processing tasks. CNNs have been shown to be effective at

removing noise from images by learning to identify the noise patterns. There are many supervised learning techniques for denoising where clean targets are required to train the neural network. But in unsupervised learning techniques, clean targets are not required to train the network; it learns by looking at the noisy samples.

2.2.1 Supervised Learning Method

Zhang et al. (2017) [10] proposed a deep convolutional neural network (CNN)-based denoising method called DnCNN. DnCNN stands for Denoising Convolutional Neural Network. DnCNN achieved better denoising performance than traditional image denoising methods. In this method, both noisy images and their clean targets are required to train the network. DnCNN works well in denoising Gaussian noise. Rather than directly outputting the denoised image \hat{x} , DnCNN is designed to predict the residual image \hat{v} , i.e., the difference between the noisy observation and the clean image. The input of DnCNN is a noisy observation $y = x + v$. In general, denoising models aim to learn a mapping function $f(y) = x$ to predict the clean image. By using residual learning formulation to train a residual mapping $R(y) \approx v$, clean image is retrieved as $x = y - R(y)$. The averaged mean squared error between the desired residual images and estimated ones from noisy input can be adopted as the loss function to learn the trainable parameters Θ in DnCNN.

$$l(\Theta) = \frac{1}{2N} \sum_{i=1}^N \|R(y_i; \Theta) - (y_i - x_i)\|^2 \quad (2.1)$$

The network architecture is a modified VGG network [11]. The network has 17 layers out of which first layer is Convolutional layer+Rectified Linear Unit(Conv+ReLU) next 15 Convolutional layer+ Batch Normalization + ReLU(Conv + BN + ReLU) layers and the last layer is Convolutional layer.

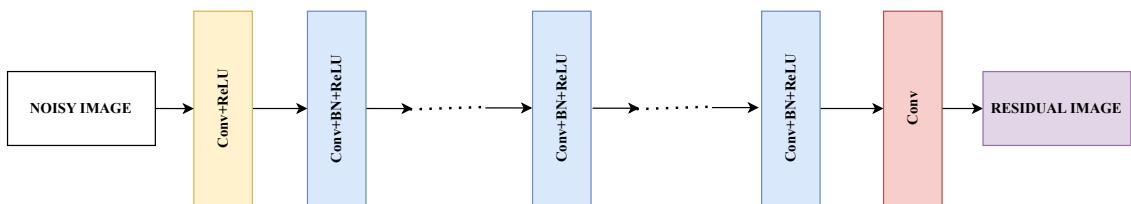


Figure 2.1: DnCNN architecture

2.2.2 Self Supervised Learning Method

In many cases, clean target images are not available. This limits the application of image-denoising methods that require them. To solve this problem, researchers have developed methods that do not require clean target images. One such method is Noise2Noise (N2N), which requires only independent pairs of noisy images. Another method, Noise2Void (N2V) [13], does not require any pairs of images, clean or noisy. Noise2Void [13] is a self-supervised denoising method that learns to predict the clean value of a pixel from its noisy neighbors.

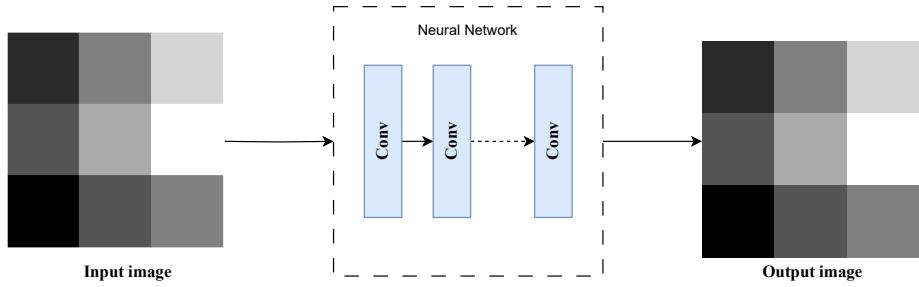


Figure 2.2: Normal working of neural network

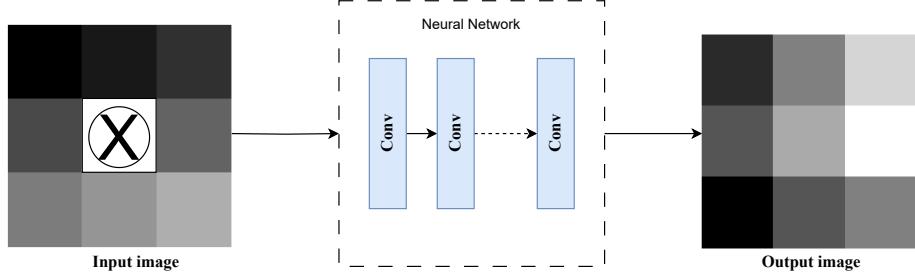


Figure 2.3: Working of Blindspot neural network

In self-supervised technique of denoising, the goal is to predict the values of a “clean” image $x = (x_1, \dots, x_n)$ given a “noisy” image $y = (y_1, \dots, y_n)$ by observing the neighbourhood of the pixel y_i . This is where the blindspot comes in, where clean pixel is predicted by looking at the neighbourhood of the noisy pixel but not the pixel itself. The normal network and the blind spot network are shown in Figures. 2.2, 2.3. Normal network takes same noisy image as input and label. The blind spot network works by not using the center pixel as input, but rather as the target. This encourages the network to

learn denoising by leveraging information from neighboring pixels, thereby restoring the true values of the blind spot area. The network architecture is a slightly modified version of the five-level U-Net architecture [14], as utilized by Lehtinen et al. [21]. Three additional 1×1 convolution layers are included at the end of network. Throughout the network architecture, all convolution layers utilize the leaky ReLU activation function, except for the final 1×1 convolution layer which employs a linear activation function [18]. Khademi et al. (2021) [16] used this concept and introduced a loss function which is appropriate for Poisson-Gaussian denoising.

2.2.3 GAN Based denoisers

Over the past few years, there has been significant research and development in the field of Generative Adversarial Networks (GANs) [12]. GAN models are designed to minimize a loss function that helps classify output images as either real or fake. By leveraging a training dataset, GANs aim to generate new data that exhibits similar statistical properties as the training data. Notably, recent advancements in GANs have demonstrated remarkable results, showcasing the capability of these models to learn and reproduce complex distributions effectively. In [25], SRGAN is used to generate noisy dataset and that dataset is used to train DnCNN for denoising. In [22], they generated a noisy dataset using pix2pix GAN [26] and trained another pix2pix GAN for denoising.

2.3 X-ray imaging

The principle behind x-ray imaging is the x-ray's capacity to penetrate objects. Depending on the substance, pixel values of the image changes. X-rays have a good penetration rate into wood and plastic. Metals are tougher because they are denser. The absorption of the specific energy of x-ray is detected by a line of sensors, allowing for the generation of a high-resolution image based on the level of x-ray absorption. However, single energy x-ray imaging has limitations when it comes to distinguishing between different types of materials or detecting specific substances, especially when those materials have similar X-ray attenuation properties. This limitation led to the development of dual-energy X-ray

imaging, which offers enhanced capabilities for material discrimination and threat detection. In dual energy x-ray imaging, two images are generated: a high energy image and a low-energy image. A high-energy image is formed when hard x-rays are absorbed by material. The attenuated X-rays result in reduced intensity on the detector. The remaining X-rays create brighter areas in the image, and a low-energy image is formed when soft X-rays are absorbed by material. The absorbed X-rays result in less exposure of the film or sensor, creating darker areas in the image.

X-ray imaging-based baggage inspection is a well established method for ensuring access control in high-security environments. But the complexity of both the contents of individual luggage and the methods used to conceal threats makes it difficult for screeners to detect low-density items and other types of threat items. The raw X-ray images obtained from inspection systems are often not clear enough to identify these items, which poses a significant challenge for security personnel. There is a rising demand for improved X-ray image processing algorithms to help in detecting potential threats and safeguarding the safety [19].

2.4 Dataset

2.4.1 Fluorescence Microscopy dataset

The available real Fluorescence Microscopy Images are of shape 512×512 . The Fluorescence Microscopy Denoising (FMD) benchmark is a collection of 12 datasets of images captured using three different types of microscopes: confocal, two-photon, and widefield. Each dataset in the FMD benchmark consists of 20 images of a sample, taken from 20 different viewpoints. Each image is 8-bit grayscale noisy, and there are 50 noisy images per view. 19 views are used for training and 20th view is withheld for testing [15].

2.4.2 Dual energy X-ray image dataset

The dual energy X-ray image dataset is obtained from a simulated baggage inspection environment, where bags were packed with random objects and subjected to X-ray scan-

ning using security equipment. The dual-energy X-ray images used in our study were provided by Vehant Technologies Private Limited, India. The grayscale images vary in size and are captured in 16-bit format. The dataset comprises a total of 1217 X-ray grayscale images, and they were divided into a training set consisting of 1085 images and a test set containing 132 images. This division allows us to train and evaluate our proposed method effectively on diverse X-ray scenes and configurations present in the dataset.

2.5 Summary

In this chapter, a comprehensive review of the existing literature related to denoising methods in security imaging is presented. Traditional methods such as VST+BM3D and Non Local PCA are discussed, along with learning-based approaches, including supervised learning methods, self-supervised learning methods, and GAN-based denoisers. The chapter also covers the relevant literature on X-ray imaging and describes the datasets used in the research.

CHAPTER 3

METHODOLOGY FOLLOWED

3.1 Introduction



Figure 3.1: Methodology followed.

Figure 4.2 gives the block diagram representation of different stages in the X-ray imaging based baggage screening denoising method followed in this work. The method consists of two steps: First, the image is denoised using a blindspot neural network, and then image enhancement algorithms are applied to the denoised images to improve their visibility.

3.2 Blindspot Neural Network

Blindspot neural network follows the self-supervised learning technique. In self-supervised technique of denoising, the goal is to predict the values of a “clean” image $x = (x_1, \dots, x_n)$ given a “noisy” image $y = (y_1, \dots, y_n)$ by observing the neighbourhood of the pixel y_i . The noise model considered in this work is Poisson-Gaussian model. In the case of Poisson-Gaussian noise, a noisy observation is created by first adding Poisson noise to a clean observation, and then adding Gaussian noise that is independent of the clean observation. Further since the mean of Poisson distribution is high, it is approximated as Gaussian with equal mean and variance [16]. Khademi et al. (2021) [16], proposed a loss function which is appropriate to address Poisson-Gaussian model. This loss function is used to train the blindspot network.

This network is trained in such a way that it takes inputs as noisy images and estimates

the mean μ_i and total variance $\hat{\sigma}_i^2$ at each noisy pixel by observing it's neighbourhood. The loss function used to train this network is as follows:

$$\mathcal{L} = \sum_i \left(\frac{(y_i - \mu_i)^2}{\hat{\sigma}_i^2} + \log(\hat{\sigma}_i^2) \right) \quad (3.1)$$

The mean μ_i and total variance $\hat{\sigma}_i^2$ are estimates of the noisy pixel y_i . The clean pixel value is estimated using noisy pixel y_i , Poisson-Gaussian noise parameters a and b as shown below:

$$\hat{x}_i = \frac{y_i \sigma_i^2 + (a\mu_i + b)\mu_i}{(a\mu_i + b) + \sigma_i^2} \quad (3.2)$$

where a and b are Poisson-Gaussian noise parameters, $\sigma_i^2 = \hat{\sigma}_i^2 - (a\mu_i + b)$ and y_i is noisy pixel value [16]. These noise parameters a and b are estimated using Nelder-Mead optimization [17] with objective function as

$$a, b = \arg \min_{a,b} \sum_i \left(\frac{(y_i - \mu_i)^2}{ax_i + b} + \log(ax_i + b) \right) \quad (3.3)$$

where x_i is the clean pixel value. Since the network does not have access to clean data μ_i is used instead of x_i .

3.3 Image Fusion Algorithms

Our X-ray image dataset contains dual energy images (high and low energy images). Each image is 16-bit grayscale image. After denoising image fusion algorithms are applied to improve the enhancement.

High-energy X-ray images are better at detecting dense materials, such as metals. This means that high-energy X-ray images will typically appear brighter in areas where there is a lot of metal, and darker in areas where there is less metal. Low-energy X-ray images are better at detecting organic materials, such as plastic and clothing. This means that low-energy X-ray images will typically appear brighter in areas where there is a lot of organic material, and darker in areas where there is less organic material. By combining

high-energy and low-energy X-ray images, it is possible to create a composite image that has improved contrast and material discrimination. This is because the high-energy X-ray image will highlight the dense materials, while the low-energy X-ray image will highlight the organic materials. This can be helpful for detecting concealed or hidden objects, as well as for identifying different materials in a complex object.

3.3.1 Image Fusion using Local Spatial Information

The algorithm works by first creating a difference image, which is the difference between the low-energy and high-energy images. The next step is to categorise each pixel in the image as either a background pixel or a detail pixel using this difference image. Background pixels are pixels that do not contain any objects of interest, while detail pixels are pixels that contain objects of interest. Following classification of the pixels, the background image is produced by averaging the intensity values of the low-energy and high-energy images for the background pixels. The intensity value of the detail pixels for the formation of detail image is then taken from either the low-energy image or the high-energy image, depending on which image has the higher contrast. The fused image is then formed by adding the background image and the detail image [19].

Algorithm 1 Image Fusion Algorithm using Image Spatial Information

- 1: **Input:** Low energy image (I_{low}), High energy Image (I_{high})
- 2: **Output:** Fused Image
- 3: **Step 1:** Computing the Difference image
- 4: Diff = $|I_{\text{low}} - I_{\text{high}}|$
- 5: **Step 2:** Categorize Pixels
- 6: **if** Intensity(Diff) > Threshold **then**
- 7: Pixel in difference image is classified as a detail pixel
- 8: **else**
- 9: Pixel in difference image is classified as a background pixel
- 10: **end if**
- 11: **Step 3:** Background Image
- 12: Background Image =
$$\begin{cases} \frac{I_{\text{low}} + I_{\text{high}}}{2}, & \text{for background pixels} \\ 0, & \text{for detail pixels} \end{cases}$$
- 13: R = Apply 3×3 AveragingFilter($\frac{I_{\text{high}} + I_{\text{low}}}{2}$)
- 14: **Step 4:** Detail Image
- 15: Detail Image =
$$\begin{cases} I_{\text{low}}, & \text{if } |R - I_{\text{high}}| < |R - I_{\text{low}}| \text{ for detail pixels} \\ I_{\text{high}}, & \text{if } |R - I_{\text{high}}| \geq |R - I_{\text{low}}| \text{ for detail pixels} \\ 0, & \text{for background pixels} \end{cases}$$
- 16: **Step 5:** Fused Grayscale Image
- 17: Fused Image = Background Image + Detail Image

3.3.2 Wavelet Based Fusion Algorithm

To fuse multiple images using wavelet based algorithm, the discrete wavelet transform (DWT) is used to decompose each image into a set of wavelet coefficients. Then fusion rule is applied to combine the wavelet coefficients from the different images. Finally, inverse DWT (IDWT) is used to reconstruct the fused image from the fused wavelet coefficients. Algorithm goes this way:

1. Wavelet Decomposition: Perform the discrete wavelet transform (DWT) separately on the I_{low} and I_{high} images. Select a wavelet family, such as the Daubechies wavelet

transform (e.g., db4), and choose the desired number of scales for the decomposition.

2. Low-Pass Filtering: Apply a low-pass filter to the approximation coefficients of I_{low} and I_{high} . This step aims to create a smoother representation of the scene by averaging the respective approximation coefficients from both images.
3. Detail Coefficients Combination: Combine the detail coefficients from I_{low} and I_{high} to obtain the detail coefficients for the fused image. This step adds details that are specific to I_{low} or I_{high} and enhances the visibility of details present in both images. The relevant detail coefficients from I_{low} and I_{high} are added to generate the fused detail coefficients at each decomposition level.
4. Inverse Discrete Wavelet Transform: Perform the inverse discrete wavelet transform (IDWT) on the fused approximation coefficients and detail coefficients obtained in Steps 2 and 3, respectively. This reconstruction step uses the IDWT to generate the fused image from the wavelet coefficients [19].

3.4 Summary

This chapter introduces the methodology followed in this research to achieve X-ray image denoising based baggage screening with improved performance. The Blindspot Neural Network, a deep learning-based denoising method, is explained in detail. The chapter also discusses different image fusion algorithms employed for enhancing the denoised security images. Specifically, the fusion methods using local spatial information and a wavelet-based fusion algorithm for dual-energy X-ray images are described.

CHAPTER 4

EXPERIMENTS AND RESULTS

4.1 Denoising Convolutional Neural Network (DnCNN)

This work started experimenting with DnCNN [10] for denoising purposes. DnCNN was trained on Berkeley Segmentation Dataset images in supervised manner. The objective of this experiment was to observe the performance of DnCNN on Poisson-Gaussian noise model. DnCNN method was designed to act as an effective Gaussian denoiser and this work observed that such a denoiser is not able to perform well when Poisson-Gaussian noisy images are given as input. Because, the output of the DnCNN is residual noise, as Gaussian noise is additive the denoised image is obtained by subtracting that residual noise output from input noisy image. But in case of Poisson-Gaussian noisy image, the noise is not additive that is why, when residual noise is subtracted from noisy image with Poisson-Gaussian noise, the performance was bad. Results and PSNR values of Gaussian denoising using DnCNN are shown in Figure 4.1.

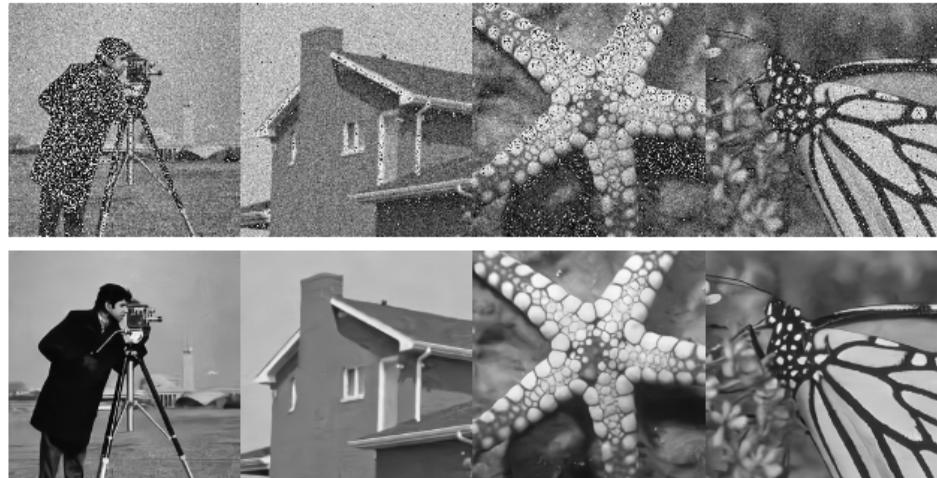


Figure 4.1: Gaussian denoising Results for variance=0.1, Average PSNR=30.35, SSIM=0.93

4.2 Blindspot Neural Network:Denoising of Microscopy images

To address the Poisson-Gaussian denoising issue, Khademi et al. [16] used this blindspot network for denoising microscopy images. Experiments were conducted using this network where a loss function was introduced which is appropriate for this noise. It was observed that this network works good for real microscopy images where standard noise model is Poisson-Gaussian. The network architecture follows modified U-net architecture, trained for 300 epochs with a learning rate of 0.0003, each epoch consisting of 50 batches of 128×128 crops from random images from the training set and batch size of 4.

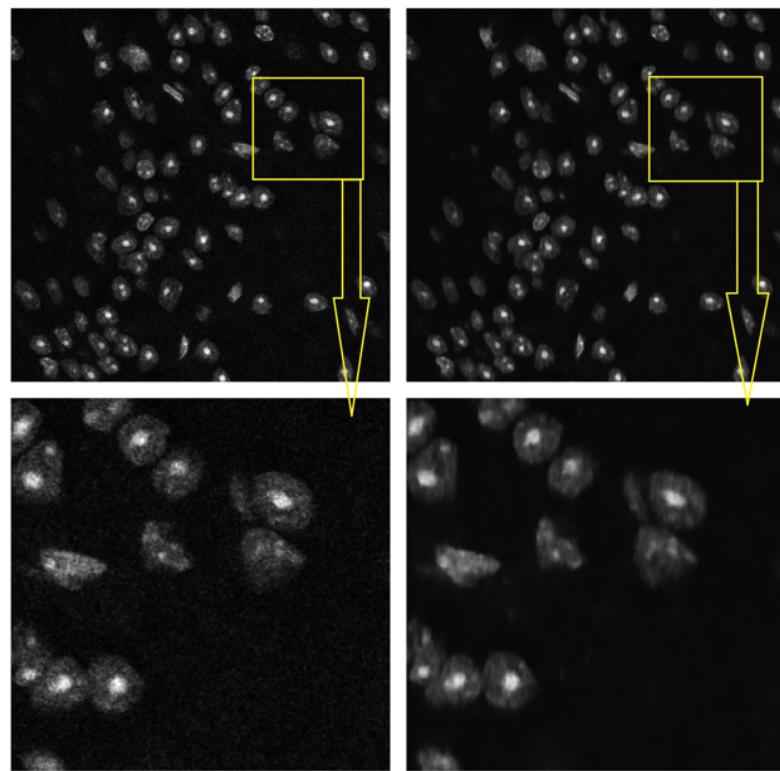


Figure 4.2: Results of blind-spot network with Poisson-Gaussian noise model using FMD. Top row: 1) Noisy image, 2) Denoised image (from left to right). Bottom row: Corresponding zoom in versions.

4.3 Experiments using pix2pix GAN

The microscopy dataset that is available was very small in size. So, Poisson-Gaussian noise model was generated for fluorescence images using pix2pix GAN [22]. To generate the noise model, first the GAN is trained with paired image data where each image pair contains real microscopy image(taken from the already existing microscopy dataset [15]) and it's binary mask shown in Figure 4.4.

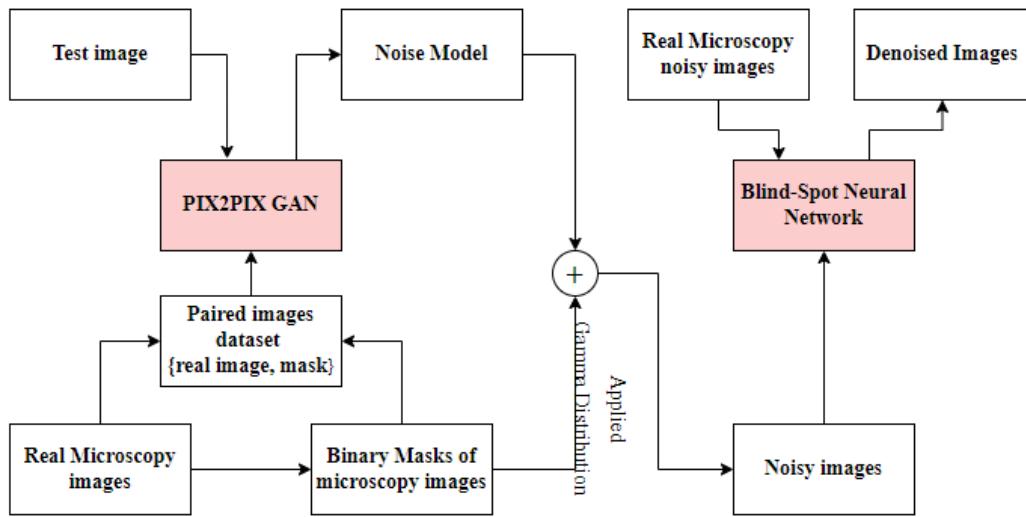


Figure 4.3: Methodology followed to generate the noise model and denoise using GAN

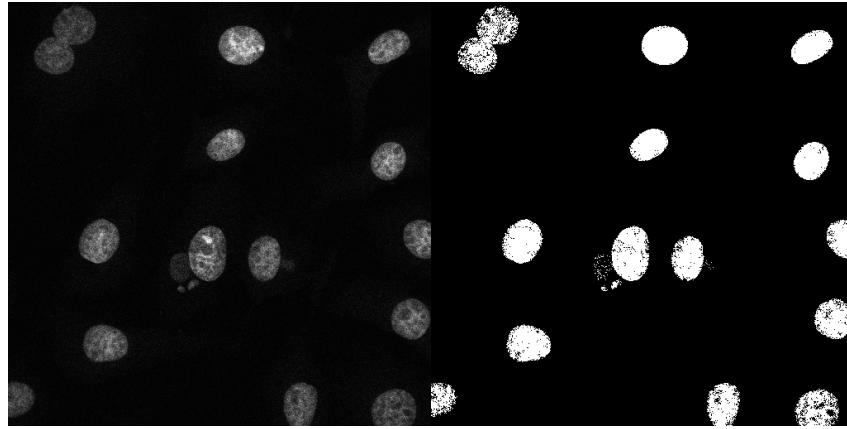


Figure 4.4: Example of paired image

Here, pix2pix GAN learns mapping from binary masks to real images. During training GAN generates noisy images and the generated noisy images are shown in Figure 4.5.

To check how similar generated and real images are, Sorensen distance method [24] is used to measure the variation in their distributions. It takes values from 0 to 1. If the two distributions are dissimilar, then a value of 0 means there is no difference between those two distributions, and a value of 1 means the distributions do not share any similarity.

$$S_D(A, B) = \frac{\sum_{i=1}^n |A_i - B_i|}{\sum_{i=1}^n |A_i + B_i|} \quad (4.1)$$

Where A and B represent two probability density functions of the real image and generated image.

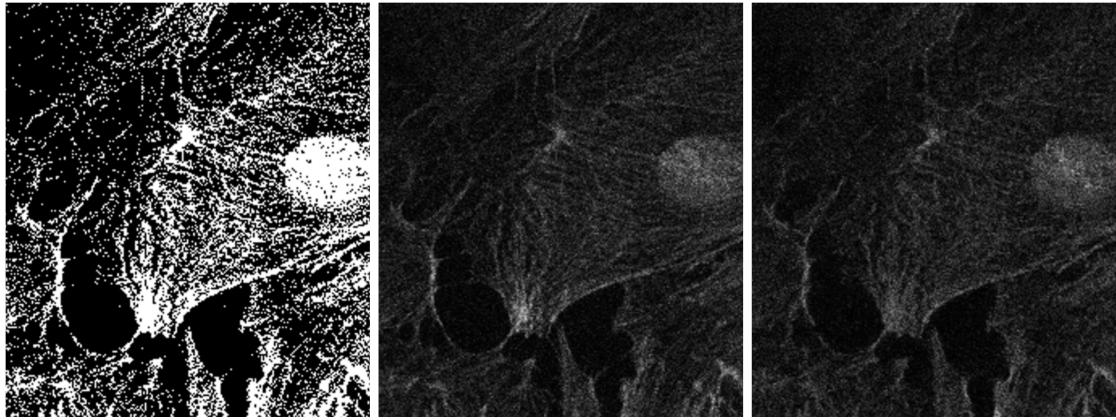


Figure 4.5: Left to right: Binary mask, Real image corresponding to binary mask, Generated image.
Sorensen Distance = 0.0367

The background of the microscopy images is black. So during testing, a nearly all-black mask is fed to a well-trained pix2pix GAN, to generate noise on that mask. So, a realistic synthetic background(noise) is formed. Images are shown in Figure 4.6.

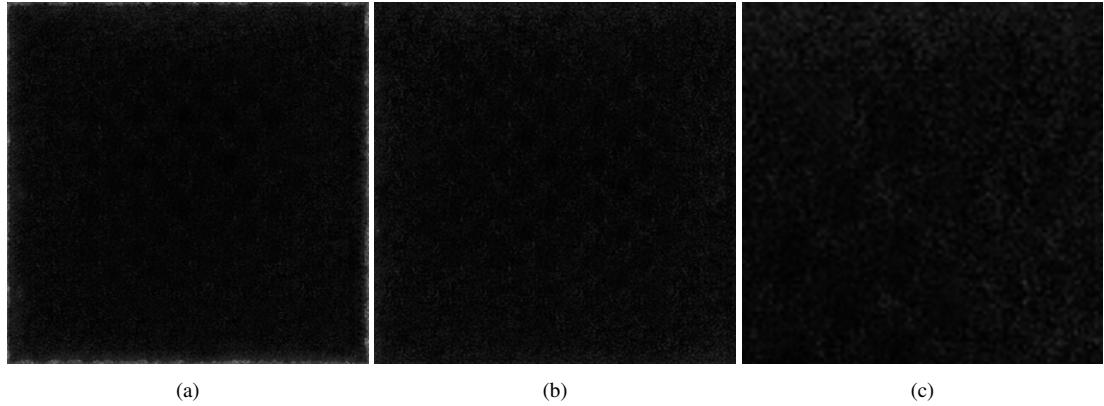


Figure 4.6: (a) Test image, (b) Generated noise model-background, (c) Cropped version of background

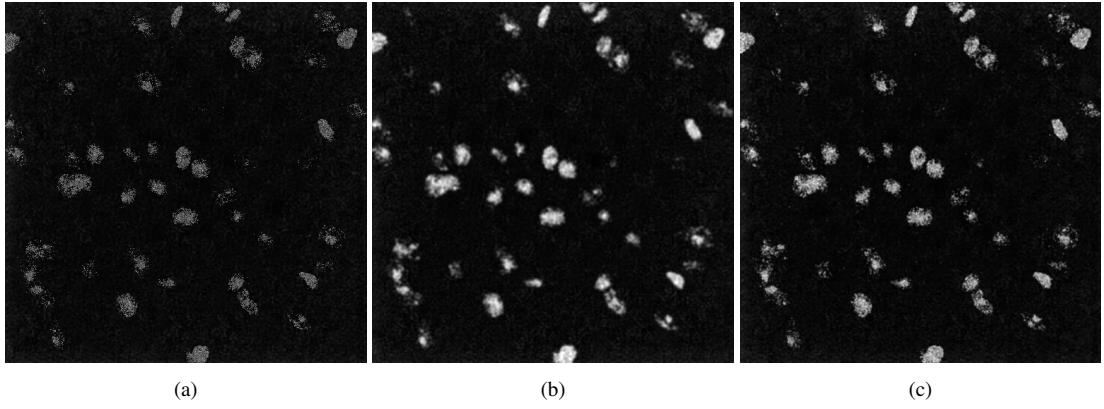


Figure 4.7: Examples of Noisy image samples after applying Gamma Distribution on binary masks

Now, the foreground is generated by filling the signal regions of binary masks(that are created to train the GAN) randomly with samples from Gamma Distribution. Assuming that the noise model that is generated is additive, a synthetic realistic noisy image is simulated by adding a foreground image and background image. Like this, a dataset is generated to train the blindspot neural network and tested this trained network with already available real microscopy images. Results are shown in Figure 4.8.

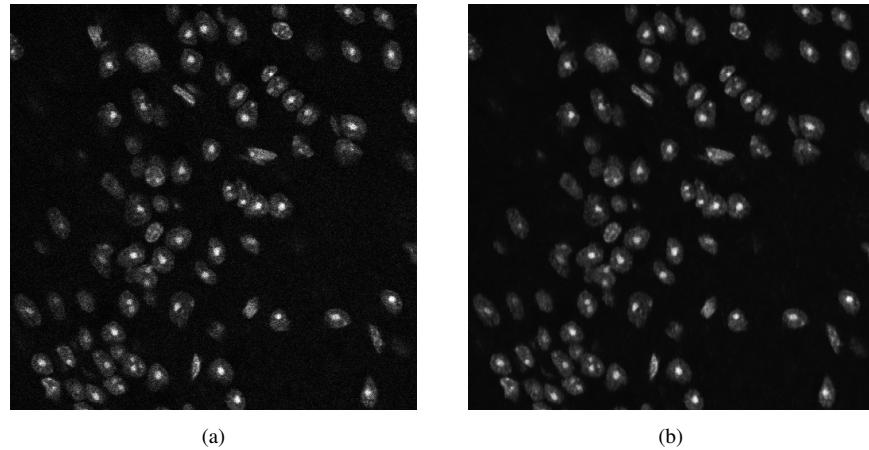


Figure 4.8: (a) Real Noisy Image, (b) Denoised Image

The objective of this experiment was to observe the performance of the blindspot neural network which is trained on synthetic images and tested with real images and it was observed that it is giving closer BRISQUE scores when compared with a network trained on real noisy images. This could be because of assuming the generated Poisson-Gaussian noise model as additive noise which is not true. Creating the noisy dataset is adopted from [22]. For pix2pix GAN, a total of 1080 mask-image pairs of size 512×512

were used for training, and the number of training epochs was 1000. BRISQUE scores are shown in Table 4.1. BRISQUE stands for Blind/Referenceless Image Spatial Quality Evaluator, which is an algorithm used for No-Reference Image Quality Assessment. BRISQUE score is a referenceless metric, meaning it does not require a reference image for comparison. It evaluates image quality solely based on the statistical characteristics of the image itself. It is typically a single numerical value, with lower scores indicating better image quality. Higher BRISQUE scores imply that the image has more perceptible distortions or artifacts, and is considered to have lower visual quality [20].

Table 4.1: BRISQUE Scores of microscopy images before and after denoising

Method	Before denoising	After denoising
Blindspot Neural Network	41.55	31.65
GAN + Blindspot Neural Network	41.55	33.59

Creating a dataset in case of security x-ray images is not possible using above mentioned GAN-based method. Because x-ray images have a white background and most of the noise is Gaussian and the Poisson noise only present in the darker regions of the image. So, the blindspot neural network was chosen to use and the network was trained on real x-ray images and the performance was evaluated.

4.4 Blindspot Neural Network: Denoising of X-ray images

In the case of X-ray images, dual-energy X-ray images (high and low energy images) are used. Each image is 16-bit grayscale image and the shape of the images is 848×704 . But the network can be trained only on 8-bit grayscale images of shape 512×512 . So, the original image is mirror padded to 1024×1024 and that image is split into four 512×512 images. As the network was previously trained on microscopy images, assuming that the network learned the noise model, the network directly tested with those split 512×512 x-ray images. Then it was observed that the denoised images got blurred and visibility became bad. This is because most of the pixels in microscopy images are black and X-ray images have white background. So, the Blind-spot network was re-trained with those split 16 bit 512×512 images and tested. During testing, the four parts of the mirror padded image is tested separately, then merged those denoised parts to form 1024×1024

and cropped out the mirror padded part to bring the denoised image back to its original shape as the test image.

4.4.1 Evaluation Criteria

Due to the unavailability of clean and noise-free X-ray images, traditional evaluation metrics like peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) [23] cannot be utilized. To address this challenge, we opted for a non-reference image quality assessment method called Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) to objectively evaluate the performance of our method. BRISQUE leverages scene statistics of locally normalized luminance coefficients to quantify potential quality losses caused by distortions, enabling the evaluation of image quality. A lower BRISQUE value indicates higher image quality, providing a reliable metric for assessing the effectiveness of our approach.

BRISQUE algorithm involves normalizing the pixel intensities using Mean Subtracted Contrast Normalization (MSCN), MSCN images are multiplied by shifted versions of themselves in four orientations (horizontal, vertical, off diagonal, and on diagonal). This captures the neighborhood relationships between pixels. Features are extracted at 2 scales - the original image scale, and at a reduced resolution (low pass filtered and downsampled by a factor of 2). From the normalized and pairwise product images, a feature vector of size 36×1 (18 at each scale) is computed. Now, these feature vector is fed to a pre-trained support vector machine regressor to calculate the final quality score of the image [20].

4.4.2 Results

The results of denoising are shown in Figures 4.9-4.11. In some cases visually it will be difficult to find the differences between noisy and denoised images, as the images used are grayscale images. So, qualitative results are shown in Table 4.2. From Table 4.2, it is observed that after denoising the quality of the image has been improved as denoised images have low BRISQUE scores.



Figure 4.9: CTP image, Left to Right: Noisy image, Denoised version

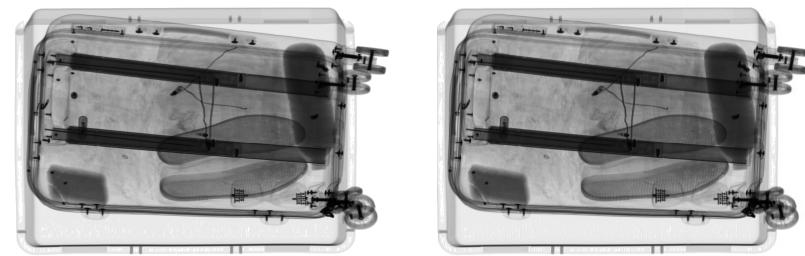


Figure 4.10: BAG1, Left to Right: Noisy image, Denoised version

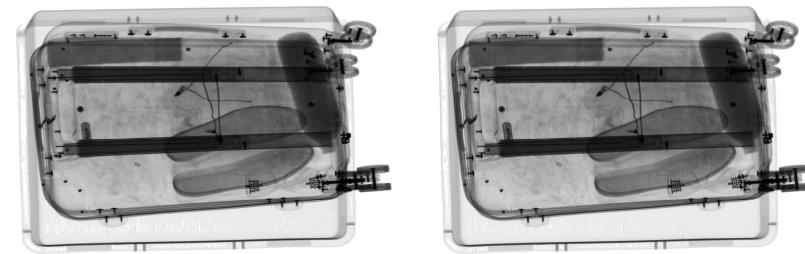


Figure 4.11: BAG2, Left to Right: Noisy image, Denoised version

Table 4.2: BRISQUE Scores of X-ray images before and after denoising

Image	Before denoising	After denoising
Combined Test Piece	43.89	39.44
BAG 1	36.38	32.54
BAG 2	36.94	33.29

4.4.3 Comparison Results

For comparing the deep learning method with traditional approaches like VST+BM3D and non-local PCA for image denoising, the BRISQUE score metric was used. The results consistently favored the deep learning method, indicating that it outperforms VST+BM3D and non-local PCA in terms of image quality assessment. The deep learning method's ability to capture and learn complex features from the data allows it to achieve higher perceptual quality, resulting in lower BRISQUE scores compared to the traditional methods. Results are shown in Table 4.3.

Table 4.3: BRISQUE Scores of different baggage scenes before and after applying Denoising methods on X-ray image dataset and the best values are highlighted

Baggage Scene	Before Denoising	Denoising methods			
		VST+BM3D	NLPCA	PGPCA	BNN(DL)
BAG3	36.5	41.29	46.99	36.35	33.64
BAG4	36.8	38.02	47.28	36.09	33.65
CTP	43.89	45.9	47.67	43.99	39.44

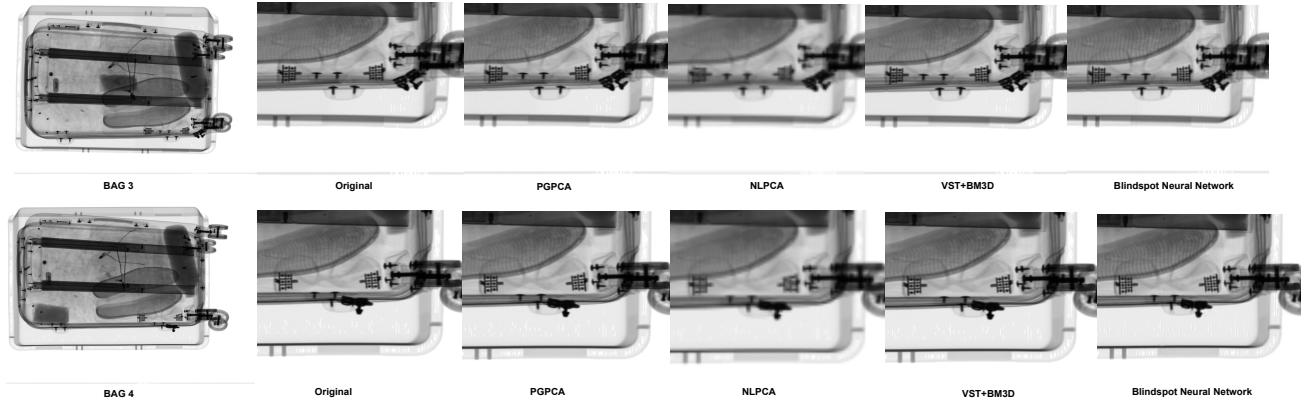


Figure 4.12: Comparison between Traditional Methods and followed Deep Learning Method

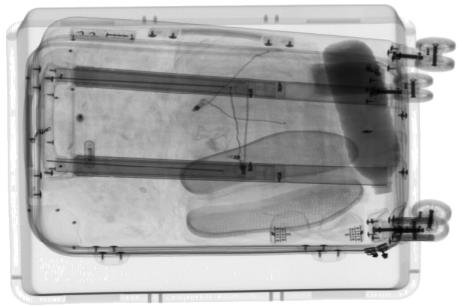
4.5 Image Fusion

After denoising, high and low-energy X-ray images, image fusion algorithms are applied for better visualization. The goal of image fusion in this context is to combine the infor-

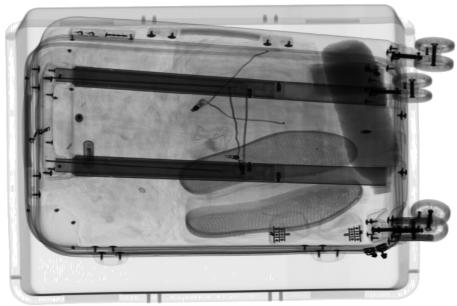
mation from two X-ray images, one with low energy and one with high energy, to create a single image that is easier for screeners to interpret. This image fusion method makes it possible to see details that are hidden or not clearly visible in the original dual-energy images. Results of image fusion using Image Spatial Information are shown in Figure 4.13. Wavelet-based image fusion is shown in Figure 4.14. Details have been improved with the wavelet-based fusion method but visually it does not look good because shadows are being formed at the edges of the objects.

4.6 Summary

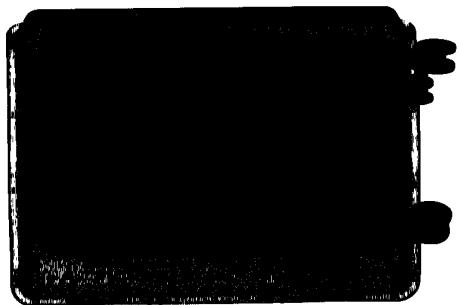
This chapter presents the experimental setup and the results obtained during the research. It includes the evaluation of the Denoising Convolutional Neural Network (DnCNN) and the Blindspot Neural Network for denoising microscopy and X-ray images. Additionally, experiments using the pix2pix GAN for denoising are discussed. The evaluation criteria, results, and comparison with traditional methods are presented. The chapter also covers the image fusion process and its outcomes.



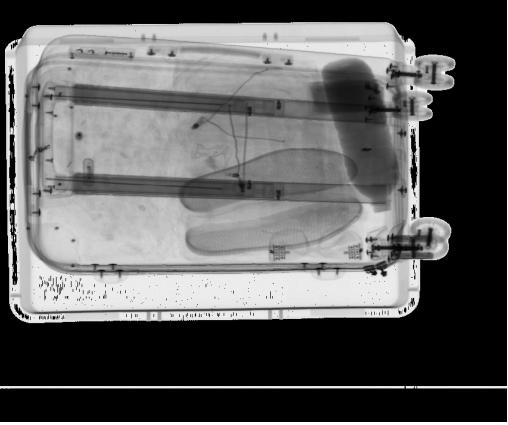
(a) High Energy Image



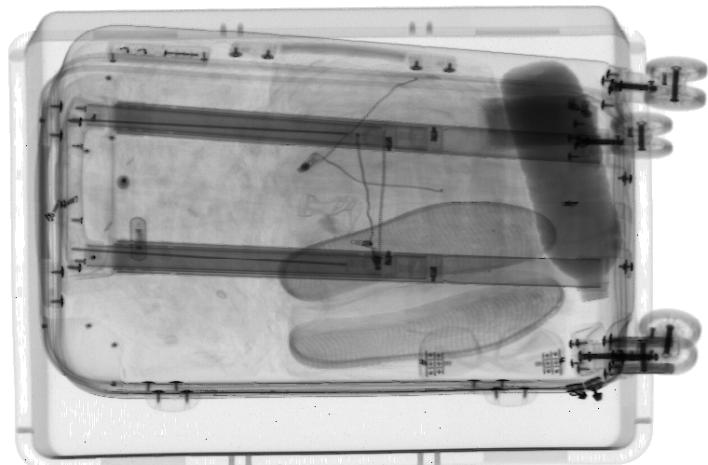
(b) Low Energy Image



(c) Background Image



(d) Detail Image



(e) Final Fused Image

Figure 4.13: Image Fusion using Local Spatial Information

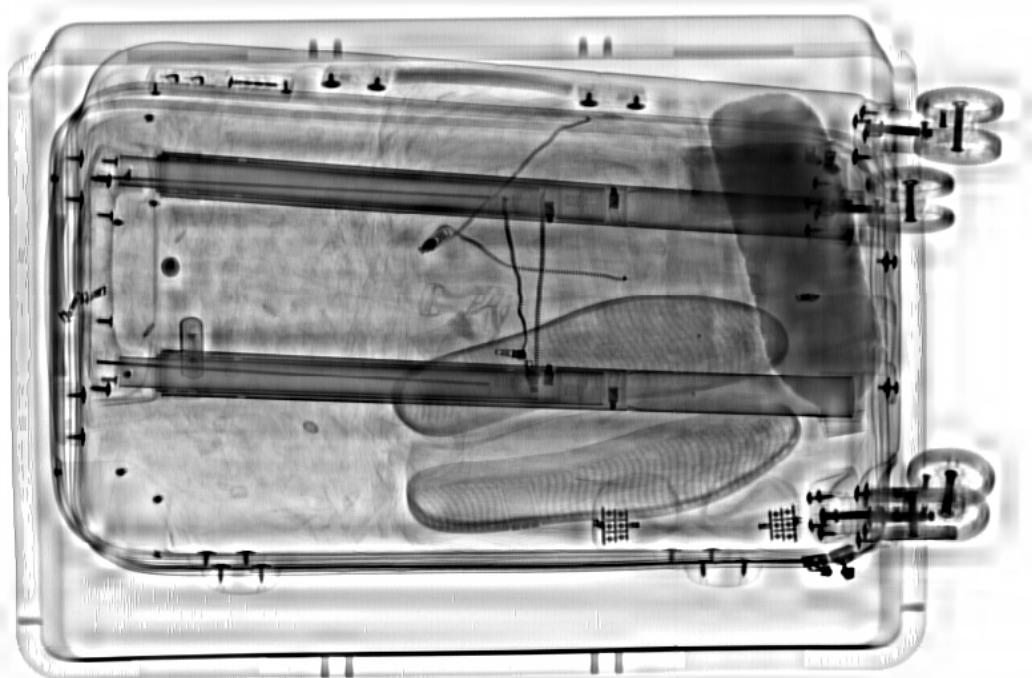


Figure 4.14: Final Fused image after fusing 4.13a, 4.13b using wavelet transform.

CHAPTER 5

BAGGAGE SCREENING WITH IMPROVED VISUALIZATION

5.1 Pseudo Coloring

The use of color serves two primary purposes: Firstly, incorporating color can significantly enhance object discrimination, as humans have the ability to perceive a broader spectrum of colors compared to a limited range of gray shades. By utilizing color, images become more vibrant and engaging, thereby improving the attention span of viewers. To achieve this effect, the technique of Pseudo Coloring is employed to add color to fused images. Pseudo-color processing involves assigning specific colors to different gray levels in a black-and-white image.

HSI-based color transform is used to color the image. The color assigned to each gray level is as follows:

- As there are three channels Hue(H), Saturation(S), and Intensity(I) each channel is assigned separately and finally stacked those three channels to form a single color image.
- **H-channel:** A fused grayscale image that is generated using an image fusion algorithm.
- **S-channel:** Constant value of 0.8, this means a grayscale image is created where each pixel value is 0.8.
- **I-channel:** Applied Gamma Correction on fused image.
- Stack above three channels and that will be the final color image [19].

The image used in coloring is Figure 4.13e and the final colored image is shown in Figure 5.1. The high-density material has a low degree of transparency and consequently higher

pixel intensities. In Figure 4.13e the bottom right corner of the bag contains eyewear. Due to its high pixel values, it is barely visible, but after coloring, the visibility of eyewear is improved which can be seen in Figure 5.1.

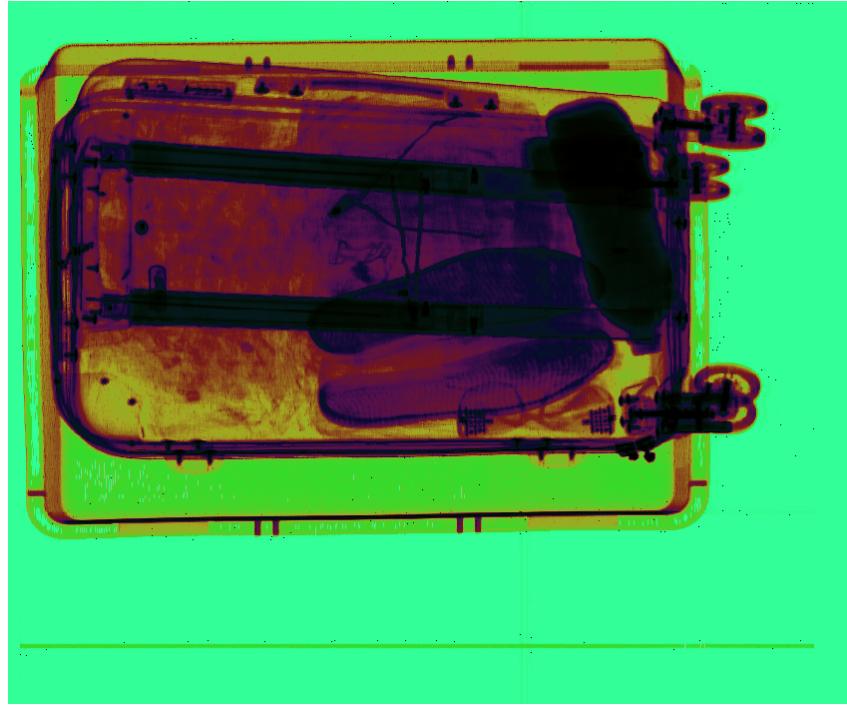


Figure 5.1: Colored version of Figure 4.13e.

5.2 Standard Deviation of single material patches

Combined Test Piece(CTP) images that are shown in Figure 4.9 contains different materials of different thickness. Standard Deviation is calculated to show that denoising is being done. So, a 30×30 patch is selected on those materials, and the calculated standard deviation of each patch before and after denoising. It is observed that the standard deviation has been reduced after denoising, which means the noise has been suppressed. Standard Deviation values are shown in Tables 5.1, 5.2.

Table 5.1: Standard Deviation of Different Thicknesses in low energy CTP image

Patch	Before denoising	After denoising
Thickness 1	8376.84	7788.38
Thickness 2	515.08	432.26
Thickness 3	670.33	651
Thickness 4	2497.88	2434.88
Thickness 5	18817.25	18677.20

Table 5.2: Standard Deviation of Different Thicknesses in high energy CTP image

Patch	Before denoising	After denoising
Thickness 1	4854.46	4737.36
Thickness 2	507.62	469.56
Thickness 3	697.56	639.48
Thickness 4	1007.23	997.62
Thickness 5	18374.22	18225.59

5.3 Summary

In summary, Chapter 5 demonstrates the practical applications of the developed techniques in security imaging. Pseudo-coloring enhances the visual interpretation of fused images, making them more lively and aiding in object discrimination. Additionally, the evaluation of standard deviation validates the effectiveness of the denoising process. These applications contribute to improved decision-making and analysis in security-related tasks by enhancing image quality, visibility, and noise reduction.

CHAPTER 6

CONCLUSION

In this thesis, we have addressed the challenges posed by the Poisson-Gaussian noise model in security imaging and proposed a deep learning-based denoising method tailored to this specific noise model. Our experimental results have shown that the deep learning method outperforms traditional denoising methods, such as VST+BM3D and Non-Local PCA, in terms of denoising performance, as evaluated by the BRISQUE score. This indicates the effectiveness of deep learning techniques in handling the complex interplay between photon shot noise and electronic sensor noise in security imaging.

Furthermore, we have applied image fusion techniques to the denoised dual-energy X-ray image dataset for better visualization. The fusion of denoised images enhances the clarity and interpretability of the security imaging data, aiding in the identification of important details and anomalies. Additionally, we have introduced pseudo-coloring to the fused images, leveraging color to improve object discrimination and engagement in security screening tasks.

Our research contributes to the field of security imaging by providing a comprehensive approach to address the challenges of Poisson-Gaussian noise and improve the quality and reliability of security imaging data. The developed deep learning-based denoising method, along with image fusion and pseudo-coloring techniques, offers practical solutions for enhancing the visual interpretation and analysis of security images. This has the potential to enhance decision-making processes, improve security-related tasks, and aid in threat detection.

PUBLICATIONS

- Vempuluru Venkata Poorna Sekhar, Deepthi P.P, Renu M Rameshan. "*X-ray Image Denoising for Baggage Screening Using Learning Based Methods*", to be submitted in a conference, Image Processing: Algorithms and Systems XXII.

REFERENCES

- [1] Riffo, Vladimir & Flores, Sebastian & Mery, Domingo. (2017). Threat Objects Detection in X-ray Images Using an Active Vision Approach. *Journal of Nondestructive Evaluation*. 36. 10.1007/s10921-017-0419-3.
- [2] Wangzhe, Du & Shen, Hongyao & Jianzhong, Fu & Zhang, Ge & He, Quan. (2019). Approaches for Accuracy Improvement of the X-ray Image Defect Detection of Automobile Casting Aluminum Parts Based on Deep Learning. *NDT & E International*. 107. 102144. 10.1016/j.ndteint.2019.102144.
- [3] Lee, Sangyoon, Min Seok Lee, and Moon Gi Kang. 2018. "Poisson–Gaussian Noise Analysis and Estimation for Low-Dose X-ray Images in the NSCT Domain" *Sensors* 18, no. 4: 1019. <https://doi.org/10.3390/s18041019>
- [4] D. Liu, J. Liu, P. Yuan and F. Yu, "A Lightweight Denoising Method Based on Noise2Void for X-ray Pseudo-color Images in X-ray Security Inspection," 2022 4th International Conference on Industrial Artificial Intelligence (IAI), Shenyang, China, 2022, pp. 1-6, doi: 10.1109/IAI55780.2022.9976566.
- [5] Girdher, Amina and Goyal, Bhawna and Dogra, Ayush and Dhindsa, Anaahat and Agrawal, Sunil, Image Denoising: Issues and Challenges (September 2, 2019). Proceedings of International Conference on Advancements in Computing & Management (ICACM) 2019.
- [6] M. Makitalo and A. Foi, "Optimal Inversion of the Generalized Anscombe Transformation for Poisson-Gaussian Noise," in *IEEE Transactions on Image Processing*, vol. 22, no. 1, pp. 91-103, Jan. 2013, doi: 10.1109/TIP.2012.2202675.
- [7] K. Dabov, A. Foi, V. Katkovnik and K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering," in *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 2080-2095, Aug. 2007, doi: 10.1109/TIP.2007.901238.

- [8] J. Salmon, C. -A. Deledalle, R. Willett and Z. Harmany, "Poisson noise reduction with non-local PCA," 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, 2012, pp. 1109-1112, doi: 10.1109/ICASSP.2012.6288081.
- [9] Deledalle, Charles & LTCI, CNRS & Salmon, Joseph & Dalalyan, Arnak. (2011). Image denoising with patch based PCA: local versus global. 10. 10.5244/C.25.25.
- [10] K. Zhang, W. Zuo, Y. Chen, D. Meng and L. Zhang, "Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising," in IEEE Transactions on Image Processing, vol. 26, no. 7, pp. 3142-3155, July 2017, doi: 10.1109/TIP.2017.2662206.
- [11] Simonyan, Karen & Zisserman, Andrew. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv 1409.1556.
- [12] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., Generative Adversarial Networks, Jun. 2014.
- [13] Krull, Alexander & Buchholz, Tim-Oliver & Jug, Florian. (2019). Noise2Void - Learning Denoising From Single Noisy Images. 2124-2132. 10.1109/CVPR.2019.00223.
- [14] Ronneberger, O., Fischer, P., Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab, N., Hornegger, J., Wells, W., Frangi, A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science(), vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28
- [15] Yide Zhang, Yinhao Zhu, Evan Nichols, Qingfei Wang, Siyuan Zhang, Cody Smith, and Scott Howard. A Poisson-Gaussian Denoising Dataset with Real Fluorescence Microscopy Images. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 11710–11718, 2019.
- [16] Khademi, Wesley & Rao, Sonia & Minnerath, Clare & Hagen, Guy & Ventura, Jonathan. (2021). Self-Supervised Poisson-Gaussian Denoising. 2130-2138. 10.1109/WACV48630.2021.00218.

- [17] John A Nelder and Roger Mead. A simplex method for function minimization. *The computer journal*, 7(4):308–313, 1965.
- [18] Samuli Laine, Tero Karras, Jaakko Lehtinen, and Timo Aila. High-quality self-supervised deep image denoising. In *Advances in Neural Information Processing Systems*, pages 6968–6978, 2019.
- [19] Zheng, Yue. “X-Ray Image Processing and Visualization for Remote Assistance of Airport Luggage Screeners.” (2004).
- [20] A. Mittal, A. K. Moorthy and A. C. Bovik, "No-Reference Image Quality Assessment in the Spatial Domain," in *IEEE Transactions on Image Processing*, vol. 21, no. 12, pp. 4695-4708, Dec. 2012, doi: 10.1109/TIP.2012.2214050.
- [21] Lehtinen, Jaakko & Munkberg, Jacob & Hasselgren, Jon & Laine, Samuli & Karras, Tero & Aittala, Miika & Aila, Timo. (2018). Noise2Noise: Learning Image Restoration without Clean Data.
- [22] L. Zhong, G. Liu and G. Yang, "Blind Denoising of Fluorescence Microscopy Images Using GAN-Based Global Noise Modeling," 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 2021, pp. 863-867, doi: 10.1109/ISBI48211.2021.9434150.
- [23] Setiadi, De Rosal Ignatius Moses. “PSNR vs SSIM: imperceptibility quality assessment for image steganography.” *Multimedia Tools and Applications* 80 (2020): 8423 - 8444.
- [24] Cha, Sung-Hyuk. (2007). Comprehensive Survey on Distance/Similarity Measures Between Probability Density Functions. *Int. J. Math. Model. Meth. Appl. Sci.*. 1.
- [25] Tran, Linh & Nguyen, Son & Arai, Masayuki. (2021). GAN-Based Noise Model for Denoising Real Images. 10.1007/978-3-030-69538-5_34.
- [26] P. Isola, J.-Y. Zhu, T. Zhou, and A. A. Efros, "Image-to-image translation with conditional adversarial networks," in *CVPR*, 2017, pp. 1125-1134.