IMPROVING ULTRASOUND KIDENY STONE DETECTION USING DEEP LEARNING

Bekalu Gedifew Molla

A Thesis Submitted to the Centre of Biomedical Engineering, Addis Ababa Institute of Technology, Addis Ababa University in Partial Fulfilment of the Requirements for the Degree of Master of Science in Biomedical Engineering (Bioinstrumentation & Imaging)

August 2023

Addis Ababa, Ethiopia

ADDIS ABABA UNIVERSITY

ADDIS ABABA INSTITUTE OF TECHNOLOGY

CENTER OF BIOMEDICAL ENGINEERING

IMPROVING ULTRASOUND KIDENY STONE DETECTION USING DEEP LEARNING

By

Bekalu Gedifew Molla

Advisors:

Dr. Dawit Assefa Haile (PhD)

August 2023

Addis Ababa, Ethiopia

DECLARATION

I, the undersigned, declare that this thesis is my own and original work and has not been presented for a degree in any other university, and that all sources of materials used for the thesis have been duly acknowledged, following the scientific guidelines of the institute.

Student's Name: Bekalu Gedifew Molla

Signature _____

This thesis has been submitted for examination with my approval as an advisor.

Advisor's Name: Dr. Dawit Assefa Haile (PhD)

Signature _____

August 2023

CERTIFICATE OF EXAMINATION

Addis Ababa University
School of Graduate Studies

This is to certify that the thesis prepared by Bekalu Gedifew Molla entitled "IMPROVING ULTRASOUND KIDENY STONE DETECTION USING DEEP LEARNING" submitted in partial fulfilment of the requirements for the Degree of Master of Science in Biomedical Engineering (Bioinstrumentation and Imaging) complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the examining committee

_____	_____	_____
Name of External Examiner	Signature	Date
_____	_____	_____
Name of Internal Examiner	Signature	Date
_____	_____	_____
Name of Advisor	Signature	Date
_____	_____	_____
Chair Person	Signature	Date

_____

Chief of Department or Graduate Program Coordinator

ACKNOWLEDGMENT

Firstly, I would like to express my gratitude to Institution Review Board Director of St. Paul Millennium Medical College (SPMMC), Professor Tariku Chanyalew. Also to the medical director of Torhailoch Comprehensive Specialized Hospital for their unlimited support to collect all the necessary data for this study. I am sincerely grateful to my scientific advisor, Dr Dawit Assefa Haile for his valuable advice on this research subject. As a thoughtful instructor, he was guiding me through all the way, until I reach to the end of this Study. I am also sincerely grateful to the librarians at SPMMC who helped me with search for literature and articles on my study subject. Last but not the least; I would like to thank my family for supporting me spiritually to work hard throughout from the beginning to the end of this study.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBRIVIATIONS

2D-GDWT – 2 Dimensional - Gaussian Directional Wavelet Transform

CKD - Chronic Kidney Disease

CLAHE - Contrast Limited Adaptive Histogram Equalization

CNN - Convolution Neural Network

CT - Computed Tomography

IEF - Image Enhancement Factor

EME - Effective Measure of Enhancement

ENLSS - Equivalent Number of Looks for Speckle Suppression

ESRD - End Stage Renal Disease

FN - False Negative

FP - False Positive

GFR - Glomerular Filtration Rate

GHE - Global Histogram Equalization

HCC - Hepatic Cellular Carcinoma

HE - Histogram Equalization

KUB - Kidney Urethral Bladder

MSD - Mean Square Difference

MSE - Mean Square Error

mSv - Milli Sievert

NME - Noise Mean Error

PACS - Picture Archiving and Communication System

PSNR - Peak Signal to Noise Ratio

RMSHE - Recursive Mean-Separate Histogram Equalization

SAR - Synthetic Aperture Radar

SSIM - Structural Similarity Index Matrix

TN - True Negative

TP - True Positive

US - Ultrasound

WT - Wavelet Transform

ABSTRACT

Background: Currently a CT scan is preferred over Ultrasound images for Kidney stone diagnosis. The major problem regarding diagnosis of these stones under CT imaging modality is that once these patients are diagnosed as positive; there is a high chance for the stone to be formed again in the patient's lifetime after removal. As a result, use of the CT modality repeatedly exposes the patient for unwanted radiation exposure.

Purpose of the Research: In ultrasound imaging, additive and multiplicative noises are taken as disadvantage for its imaging output. However, its ability to form real time imaging makes it preferable in many diagnosis procedures. The current study is aimed at developing an effective kidney stone detection scheme using a Convolutional Neural Network (CNN) by incorporating useful image pre-processing tools applied on Ultrasound images.

Methods: The approach implemented in the proposed kidney stone detection scheme mainly involves two stages. The first stage employed using useful pre-processing steps applied on the Ultrasound images, which include image filtering, contrast enhancement and 2-D Directional Wavelet Transforms. The second stage is employed using multiclass semantic segmentation CNN models, which include Deep Lab V3 , U-net and Seg-net models. In order to detect multiclass regions of ultrasound kidney stone image.The performance of the models was evaluated using useful quantitative matrices.

Results and Conclusion: Results have shown that the Deep Lab V3 CNN model had greater performance than U-net and Seg-net CNN models tested in this study. The model was able to maintain a global accuracy and mean accuracy of 95.1% and 80.9% respectively showing its great promises in improving the detection of kidney stones based on Ultrasound images. Compared to performances reported in the literature by previous scholars who have developed different method of kidney stone detection algorithms, the proposed method has offered commendable results in terms of global accuracy and mean accuracy.

1 INTRODUCTION

1.1 Background

Kidney stones are hard deposits of minerals and acid salts that stick together in the urine. This stones may vary based on their type and size. If these stones are not managed properly on time, they will affect the normal function of the Kidney and result in Chronic kidney diseases (CKD) [1]. CKD is a recognized complication of kidney stones as a result of nephron calcinosis which can lead to progressive loss of Glomerular Filtration Rate (GFR) and end stage renal diseases (ESRD). The international guideline defines CKD as a decrease in kidney function as shown by GFR less than 60ml/min/1.73m^2 [2].

The existence of solid particles in the urinary system, the so-called urinary calculi, is considered a major problem in many nations. The prevalence of urinary calculi is estimated to be 1-5% worldwide and 2-13% in developed countries.It is on a rising trend in tropical and sub-Saharan African countries due to growing socio-economic status as manifested by westernized diet and lifestyle [3] [4] [5]. When kidney stone exit the renal pelvis and move into the remainder of the urinary collecting system, the condition is called Urolithiasis. Recent studies have shown that the prevalence of Urolithiasis has been increasing in developed as well as developing countries. According to an institutional based cross section retrospective study conducted from July 2016 to December 2017 in Addis Ababa, Ethiopia, 247 cases (around 30%) were affected by Urolithiasis from a total of 824 patients considered in the study which were all linked to urology cases. Of the 247 cases, 202 were further investigated and 186 (92%) had upper urinary tract stones of which 96 (51.6%) were with renal stones. More than two-thirds (164, 81.2%) of the patients had complications and 148 (73.3 %) had hydro nephritis [6]. Another study conducted between January 1st and December 31st in 2019 at St. Paul Millennium Medical College (SPMMC) has reported that from a total of 620 urologic admissions, 139 (22.4%) had chronic kidney diseases using cut-off value of 60ml/min/1.73m^2 on a GFR scale. Stone formation generally affects all ages, sexes, and races [7].

Ultrasound imaging is one useful modality in medical imaging. It is beneficial in examining of different internal organs including heart, blood vessels, liver, gallbladder, spleen, pancreas, kidneys, uterus, different glands, and many more.

From all these applications, most often times ultrasound is widely used for diagnosis of kidney stones as a primary choice due to its ability to form quick, safe and easy procedure [8]. Other imaging modalities such as CT Kidney-Ureter-Bladder (CT-KUB) and X-ray KUB are also nowadays available. The three imaging options namely ultrasound, CT and X-ray, have sensitivity and specificity of 69% & 87%, 94% & 100% and 52% & 87% for detection of kidney stones less than 3mm, respectively [9]. Although CT imaging is preferred above all these modalities for better diagnosis of kidney stones with better sensitivity and specificity, it is also known for its drawbacks in that a standard CT scan expose patients to an effective dose of ionizing radiation of ~10 mSv. Such a radiation dose might be insignificant. However, malignancies have been recorded at 100 mSv, The US National Academy of Sciences defined low doses of radiation as those up to ~100 mSv [10]. While the dose is additive over a person's lifetime [11]. If patients do not apply Meta-phylaxis, the relapsing rate of secondary stone formations is estimated to be 10–23% in one year, 50% in 5–10 years, and 75% in 20 years of the lifetime of the patient [12]. The US National Academy of Sciences defined low doses of radiation as those up to ~100 mSv [10]. X-ray also exposes patients to radiation dose but minimally compared to CT. However, X-ray comes with poor sensitivity and specificity. On the other hand, Ultrasound uses non-ionizing radiation. Moreover, its real time imaging option and cost effectiveness makes it preferable. Its low contrast, presence of speckle noise and occlusion are considered some of the disadvantages of the ultrasound.

Noises and other artifacts tend to hide the visibility of some parts on the ultrasound images and reduce the ability of observers to resolve the actual information. Such drawbacks make ultrasound based kidney stone detection a challenging task. Studies have shown that the sensitivity and specificity of ultrasound in kidney stone detection are rather low [9]. This might call for the development of image enhancement schemes to reduce the false negative/false positive detection rate. Now-a-days, with the success of deep learning and different image enhancement techniques, the detection performance of the medical imaging modalities is considerably increasing. In the meantime, the key challenges associated with deep learning in ultrasound imaging including reliability, generalizability, and bias have continued [13].

The main intent of the current thesis work is to use different image enhancement techniques and deep learning models to increase the robustness of kidney stone detection. Different image enhancement techniques are used in this study, aimed to reduce noises and improve the visual effect in ultrasound kidney stone imaging. These include use of different adaptive filters; transform domain filters and useful contrast enhancement techniques. All these techniques were selected based on different quality matrices. The effectiveness of the stone detection is tested by using different Multi class semantic segmentation deep learning models.

1.2 Statement of the Problem

Imaging modalities such as CT come with good sensitivity and specificity. Nevertheless, due to its inability to form real time imaging a repetitive scan is necessary for diagnosing of kidney stones. These phenomena leads a patient to radiation exposure. Besides, its inaccessibility, unaffordability makes it uncommon modality to use most often times. Therefore, using an imaging modality with real time imaging feature has advantage to diagnose kidney stones. In this regard, Ultrasound is commonly used imaging modality with no issue of radiation exposure. In addition, availability of Ultrasound in a clinical setting is often much higher than CT. However, detecting kidney stones accurately based on Ultrasound has stayed challenging due to its low contrast, presence of speckle noise and occlusion in tissue boundaries. This calls for the development of different digital image processing options to create an automated platform in kidney stone detection based on Ultrasound imaging with better sensitivity and specificity.

1.3 Significance of the Research

This thesis work mainly aims to improve ultrasound kidney stone detection by using appropriate image processing techniques. Accurate kidney stone detection later enables physicians during therapeutic intervention.

1.4 Research Questions

1. What are the current challenges in Ultrasound based kidney stone detection?

2. How can we deal with the limitations in the current kidney stone detection practices using deep learning models?

3. How can we enhance kidney stone detection capabilities of Ultrasound based on image processing?

1.5 Objectives

1.5.1 General objective

The general objective of the current thesis work is to detect kidney stones in Ultrasound images using improved image enhancement techniques.

1.5.2 Specific objectives

- Develop better image processing techniques to improve the quality of Ultrasound images.
- Develop a deep learning model that can be used for accurate detection of kidney stones seen on enhanced Ultrasound images.
- Evaluate the performance of the each deep learning model.

1.6 Scope of the Research

This study develops and tests a deep learning scheme that could be used in accurate detection of kidney stones making use of ultrasound images acquired locally from two hospitals found in Addis Ababa, Ethiopia. Ultrasound images captured from cases with confirmed kidney stone/s and those control healthy subjects have been used to train, validate and test the developed deep learning scheme. No effort has been exerted to test the scheme in actual clinical setting. Only data/Ultrasound images captured with acceptable image quality have been used to develop the stone detection algorithm, which in effect reduced the number of data available for training, validating and testing the algorithm, and this was considered as one of the limitations in the current study.

2 LITRETURE REVIEW

2.1 Ultrasound Imaging – Oveview

Images in an Ultrasound are produced when sound energy from the transducer is propagated through different tissues which have different acoustic impedance [14]. Ultrasound images come with different kinds of noises and artifacts hence might provide incomplete representation of the examined tissues. Therefore, examination procedures have to be facilitated by both technical and physical measures in this kind of imaging modality. For better ultrasound image quality, proper type of transducer and proper amount of energy have to be selected. These transducers can emanate proper type of beam pattern that are constructive with corresponding appropriate energy to scan the necessary area of interest. Good image quality is dependent on wavelength, attenuation effect of the tissue under consideration, velocity of propagation, wavelength of the beam, scattering effect of the tissue, and acoustic impedance of the tissue. These are also important parameters for determining the sensitivity for the area to be examined [14].

Another study emphasised on the issue of speckle noise when using Ultrasound scanning [15]. In low contrast images, with the presence of speckle noise, the image resolution is significantly reduced. Hence, it is very important to reduce speckle noise for better and accurate interpretation of image data. The wavelength of the beam that propagates through the media determines the axial and lateral resolution of the image to be formed. Hence, when one has to exploit information embedded in an Ultrasound image, different kinds of image pre-processing and enhancement techniques might have to be considered [15].

Reflection and shadowing are two clinically valuable features for Ultrasound imaging of the kidney. They indicate presence of cyst and calcification (stone/s). The reflection and shadowing phenomena are important to discover while interpreting Ultrasound images. Generally, physicians have to check the following golden rules in Ultrasonography interpretations [14].

1. One should never make interpretation based on a single image and a positive real diagnosis has to be cross-checked from different scanning planes by creating cross sectional images.
2. If the feature (cyst, stones) does not display, it does not mean it is not there.

2.2 Noise in Ultrasound Imaging

Many previous studies discussed issue of noise in ultrasound imaging. Both additive noises (Gaussian and Salt and Pepper noise) and multiplicative noise (Speckle noise) are present on ultrasound images. In that regard, image enhancement techniques are mandatory where their performance is measured in terms different matrices such as peak signal-to-noise ratio (PSNR), mean square error (MSE) and root mean square error (RMSE) values [16] [17] [18] [19].

Understanding and identifying the nature of Ultrasound imaging like attenuation effect and low contrast property are very important and the resulting noises and artefacts should be dealt with accordingly. This might require identification and categorizing the noise type [20]. In one of the previous studies, it was reported that presence of speckle noise in Ultrasound images affects edges and fine details and decreases image visibility making continuous structures appear discontinuous [21]. Hence, removal of these noises using image pre-processing is essential for better visualization of different regions of interest [14] [22].

2.3 Image De-noising Techniques

There are many previous researches done comparing different kinds of noise reduction and image enhancement techniques for Ultrasound images. One of the studies recommends using spatial linear filtering to remove noise from most anatomical details. However, such spatial filters often have too much blurring effect [23]. Another study made use of and compared spatial high pass filter with shock, spatial low pass filter and contrast limited adaptive histogram equalization (CLAHE) by using quantitative matrices including MSE and PSNR values to enhance Ultrasound images of the liver and classify them into benign, malignant and normal cases. The study showed that CLAHE outperformed the rest of the filters. Another study utilized a soft-thresholding de-noising method using wavelet decomposition method [24]. In another study, a wavelet filter was compared with several other standard speckle filters which are widely used for speckle noise suppression and concluded that the wavelet-based approach is preferred for speckle removal [25] [26]. Another method also exploited wavelet filters with different thresholding values applied on Ultrasound images and observed that the method was able to effectively supress noise by preserving edge details [27]. Another study has also used this filters and confirmed that the approach can enhance the

image quality by preserving image information and suppressing different noises [28]. A different approach using the wavelet filters recommended to perform directional filtering which is sensitive to directions when the aim is to detect useful image features and edges [29]. Another study also used directional filtering using wavelets relying on angular and scale selectivity (resolving power) in a wavelet [30]. These studies clearly showed the effectiveness of the wavelet approach when dealing with noises and other artefacts in Ultrasound images.

Other researchers used other enhancement methods for Ultrasound images. For example, Yang et al. used histogram matching for enhancing Ultrasound images. Their experimental results showed that their method leave speckles unchanged but enhanced tissue boundaries [31]. Li et al. al study has explained on adaptive image enhancement method for speckle detection based on pre-defined adaptive spatial filters including Kuan, Lee, Frost, Adaptive Homomorphic, Wiener and Anisotropic Diffusion filters, which are widely used in different fields particularly in areas of geo-physics to de-noise Synthetic Aperture Radar (SAR) images [32]. In their studies, Khan M. N. et al. have also made an experimental evaluation of these filters (Mean, Median, Kuan, Lee, Frost, Adaptive Homomorphic, Wiener and Anisotropic Diffusion filters) for removing speckle noise in Ultrsound images. Based on their study report, Frost and adaptive Homomorphic filters worked well for the Ultrasound images degraded with high level of speckle noise [33]. Another study by P. Singh and R. Shree used anisotropic diffusion filter with homomorphic scheme in db2-type wavelet transform to reduce noise without disturbing the significant parts of the SAR images. Their results showed the selected filters showed better edge preservation and noise suppression capability [34]. A study by Jaybhay J. et al. compared different noise reduction techniques used for Ultrasound images including Spatial filters, Frequency domain filters, Adaptive filters, and Deep learning models. The study asserted that choice of noise reduction methods depends on the specific application. According to the study, spatial filters are preferred for applications where image resolution is critical while frequency domain filters are good choice for applications where speckle noise reduction is more important and adaptive filters can be good choice for applications with non-uniform noise levels [35].

A study by Hiremath P. S. et al. compared Wiener filter with Lee and Frost and their results showed that Wiener filter was able to de-speckle better than Lee and Frost filters. However, the study has suggested the effectiveness of this filter is very much dependent on the noise strength and power spectrum [23] [36]. Another study by Mageshkumar G. et al. analyzed two adaptive filters, namely Frost and Kuan filters. The study emphasized that both filters were suitable for removing speckle noise from SAR images preserving useful edge information and there is insignificance difference between the two. These filters have a feature designed to retain the edge and maintain characteristics feature of the original image [37]. Similar review by Pradeep S. et al. 2021 on Adaptive filters have showed, a frost filter is better for an image corrupted by noise of speckle nature [38].

Another study by Ponmani E. et al. on de-speckling of SAR images revealed that adaptive filters used for de-speckling perform better than other transform based methods. The researchers also stated that deep learning method can achieve even much superior results with high quality image outputs specifically when applied on SAR images [39]. A study by P. Perona et al. showed that anisotropic diffusion filter applied on Ultrasound images offered best qualitative performance over other filters while similar outcomes were observed when the same filter is applied on SAR images [40]. There are also other researchers who utilized an anisotropic diffusion filter based on the so called doubly degenerate diffusion equation and achieved a good balance between speckle noise reduction and edge preservation. The methods had one major drawback in that they need high computational time to generate outputs [40] [41] [42].

As can be seen from the studies reviewed so far, it is evident that there are different filters that are available for use in noise and artefact reduction particularly for images generated in areas of Ultrasonography and SAR imaging. However, the performance of the filters is dependent on their specific usage. In the case of image de-speckling, the cost of blurring and lower speckle suppression capacity must be seriously considered. In that regard, many studies still recommend to develop new methods of de-noising based on a hybrid of different techniques that are available in the literature [15] [42] [43].

2.4 Image Contrast Enhacment Techniques

Other than noise filtering, most image pre-processing procedures involve some kind of image contrast enhancement. Image contrast enhancement aims to improve image quality and non-conformity. It is considered as one of the most important problems in medical imaging as adjusting the contrast often helps to identify details in the image. In that regard, different contrast enhancement techniques have been proposed in the literature.

One classical method used for image contrast adjustment is histogram equalization [44] [45]. The global version on histogram equalization, however, may not be always ideal. A study by Abdullah-Al-Wadud M. et al. asserted this problem of global histogram equalization (GHE) in that the method fails to capture changes in local image features [46]. A dynamic/adaptive approach proposed by the authors showed a good performance in enhancing images retaining important image details. Their method allows control over the extent of enhancement using adjusting parameter. One way to circumvent the issue of global histogram equalization is to make it adaptive, the so called local histogram equalization (LHE) [47]. Another method proposed a dynamic method based on histogram specification that generates the specified histogram dynamically from the input image [48]. Recent studies focused on improvement of image histogram equalization based contrast enhancement by separating the input image histogram into two parts based on the input image mean. After separation, each part is equalized independently to overcome the brightness preservation problem [49] [50] [51].

2.5 Deep Neural Networks

Artificial intelligence (AI) is being increasingly used in many areas of image processing and received a wide level of interest by many researchers. One subset of AI that has shown great promises in image processing is Deep Learning that has been used in various applications including segmentation, classification, object detection and other computer vision applications. One popular deep learning model is Convolutional Neural Network (CNN). The CNN based approach has shown promising strength especially in medical image segmentation and classification tasks [52]. The convolutional layer, which is the backbone of the deep learning scheme, applies a series of small filters to the input image transforming

the raw data into a unique map. A batch-normalization layer, a pooling layer and different types of activation functions (linear or non-linear including step, sigmoid, ReLU, tanh, softmask, and others) are considered important building blocks of CNN [53].

Deep learning has enjoyed different applications over the last decade applied on medical images including image enhancement, segmentation, registration, feature extraction and object recognition [54]. There are also a number of publications by different researchers that came out at different times specific to use of deep learning in detection of kidney abnormalities, kidney stone being one of them. A work by Black et al. utilized a deep CNN with computer vision algorithm applied on 63 digital images and was able to detect kidney stone compositions: calcium oxalate monohydrate (COM), uric acid (UA), magnesium ammonium phosphate hexahydrate (MAPH/struvite), calcium hydrogen phosphate dihydrate (brushite), and cystine stones. The study was able to obtain 83.2% overall accuracy for detecting four types of stone compositions [55]. Another study by Martin et al. utilized CNN applied on 465 high-resolution CT images and was able to detect ureteral stones (those kidney stones that specifically affect the ureter) with area under the curve (AUC) of 0.9971, a sensitivity of 100% and an average of 2.68 false-positives per patient on a test set of 88 scans [55]. Another study by Aziyus et al. was able to detect urinary stones with 99.59% test accuracy by using 2430 Micro-CT scans [57].

Xiang et al. has done experimental study using CNN model using the ResNet50 deep learning architecture. This study used 80 microscopy images taken from urine sediments of patients and was able to recognize 74% of Urolithiasis [58]. A study by Kazemi et al. used the Ensemble Learning model of CNN to be able to predict kidney stones using 42 biological features acquired from 936 patients and was able to detect kidney stones with an accuracy of 97.1% [59]. A study by Chaitanya et al. reported a 78% accuracy in detecting Chronic Kidney Disease (CKD) based on MRI kidney images using machine learning classifiers and Gravitational Search Algorithm [60].

In another study, Suresh et al. reported an accuracy of 92.57% kidney stone detection rate based on Ultrasound imaging utilizing image processing techniques

comprising of a pre-processing stage followed by segmentation and morphological analysis [61]. A study by Kobayashi et al. made use of 1017 KUB X-ray images and CNN model to achieve a 0.752 F-score in detecting urinary tract stones [62]. Liu et al. also used KUB X-ray images (355 images in total) and recorded 97.7% accuracy in detecting Urolithiasis [63].

The above literatures clearly show that there have been a number of efforts made by different researchers to be able to automatically detect kidney stones using various methods utilizing AI. Most of them relied on information captured from patients using a given imaging modality (CT, Ultrasound, X-ray, MRI, Microscopy) while few used biological features derived from patients. The current thesis work is more concerned with those methods that made use of information that could be obtained by imaging patients non-invasively using a given modality. For the reasons made in the first chapter of this thesis, emphasis was given to Ultrasound images. The fact that the procedure does not involve any ionizing radiation and its availability in many clinics made Ultrasound a preferred imaging modality for detecting stones. Yet, there is a still a gap to be filled in using Ultrasound for automatic kidney stone detection as one has to deal with the noises and other artifacts that are embedded in it. From the literature review we could see that few researches carried out by the different scholars using Ultrasound images showed good promises. However, there are stillrooms for improvement. It was the main intent of the current study to exploit the capabilities of deep learning in accurately detecting kidney stones based on Ultrasound images involving useful pre-processing steps. Enough number of Ultrasound images were obtained to train, validate and test the proposed kidney stone detection scheme.

Over all Methodology Diagram

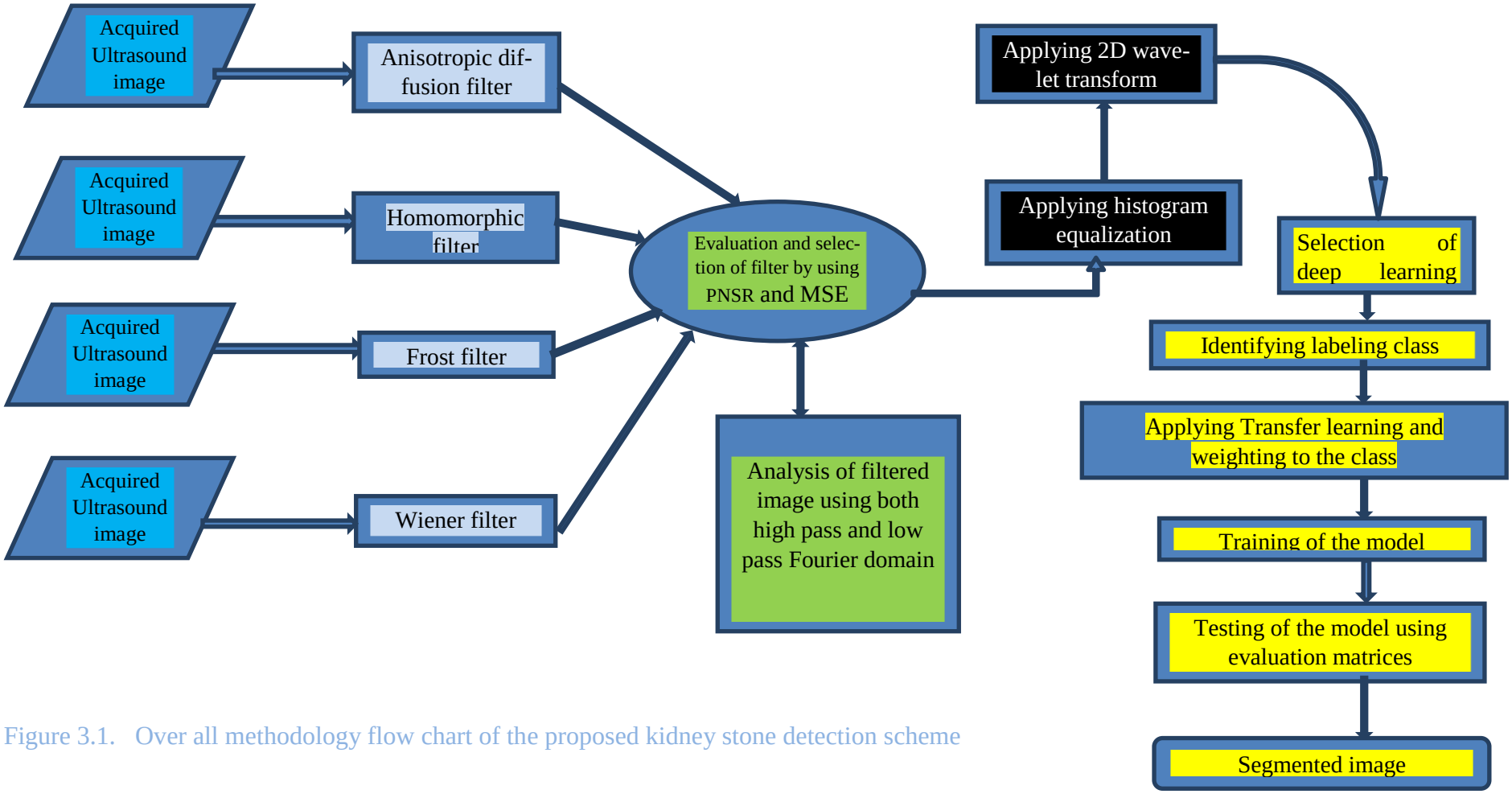

Figure 3.1. Over all methodology flow chart of the proposed kidney stone detection scheme

3 MATERIALS AND METHODS

This study mainly aims at improving Ultrasound based kidney stone detection performance using improved image enhancement techniques. The Ultrasound image data were collected locally from St. Paul's Hospital Millennium Medical College and Torhailoch comprehensive specialized hospital, Addis Ababa, Ethiopia. These institutions were selected, due to availability of the relevant Urology department and their high flow of patients. While carrying out this study, the data has been granted by the Ethical Committee of the Hospital administration unit. Previous case studies were also reviewed from the institutions library on the subject matter.

3.1 Study Design

In this study, an experimental design approach was implemented. The overall flow diagram of the methodology used in the current study to detect kidney stone from Ultrasound images is shown in Figure 3.8 at the end of this chapter. The first section of focuses on noise removal from the Ultrasound images using different kinds of adaptive filters. In addition, contrast enhancement techniques and directional wavelet are used to enhance the images further. Finally, a deep learning model is used for region of interest (ROI) detection. Detection of kidney stones consists of five major steps:

i. Filtering of additive and speckle noises using adaptive filters.
ii. Enhancing the contrast of the image.
iii. Image labelling.
iv. Applying semantic segmentation model to detect required ROIs.

3.2 Data Collection and Procedure

Data was collected in co-cooperation with a radiologist physician. Questionnaires were prepared while carrying out the data collection. The questionnaires consisted of information on significance of the current study, technical issues for Ultrasound kidney stone imaging, current challenges regarding Ultrasound kidney stone imaging and current practice of stone detection. The data is composed of Ultrasound images acquired from patients with confirmed kidney stones as well as those controls with no sign of kidney stone as confirmed by the expert physician. A non-

probabilistic sampling method was used to collect data. Sampling was done based on pre-determined knowledge about the targeted group population so that the collected Ultrasound image are representative of a larger group. Only primary data has been used in this study and data was not available from secondary data sources.

3.3 Study Datasets

A total of 250 Ultrasound images were acquired for use in training, validating and testing of the proposed automatic kidney stone detection scheme. The images were collected from 250 subjects out of which 200 were confirmed to have kidney stones while the rest 50 were controls. Among the subjects, 140 had left kidney images, 70 had right kidney images and 40 were images taken from kidney ureteral track. All the images were in transversal view. The images were collected from six different (Philips (EPIQ7, Affinity70), Siemens (AcusonX150, X300PEUI), GE Healthcare (LOGIQE 10), TOSHIBA_MEC (TUSA 500), Samsung (HS40, SonoAceR5), and SonoscapeS2) Ultrasound machines that were available at St. Paul millennium medical college Hospital and Torhailoche comprehensive specialized Hospital and were all in DICOM format. The images came with varying resolutions with pixel size ranging from 768 x 576 to 1280 x 1024. The images were further augmented (cropped, resized and flipped) to replicate them in order to increase the data size and assure robustness of the developed kidney stone detection model. During data collection, the quality of the Ultrasound images was checked and confirmed by expert radiologists. Ultrasound images acquired from those with no indication of kidney stones were used as control variables in this study.

3.4 Study Period

The time of data collection was from Feb 2022 to Oct 2022.

3.5 Data Inclusion and Exclusion Criteria

3.5.1 Inclusion criteria

- Images with acceptable image quality or further processing as confirmed by the radiologists.

3.5.2 Exclusion criteria of the study

- Images that have unacceptable degree of image artefacts as confirmed by the radiologists.

3.6 Tools Used in This Study

Matlab version R2021a was used to numerically implement the proposed stone detection algorithm. Different Matlab toolboxes were also used which have helped for executing different image processing tasks. These were Data acquisition version 4.3, 2D directional Wavelet toolbox version 5.6, Transfer learning version1.0, Computer vision toolbox version10.0, Deep learning toolbox version 14.2 and statistical and machine learning toolbox.

3.7 Data Processing and Analysis

The following steps were used to develop the proposed stone detection scheme:

- Data set preparation.
- Selection of image de-noising filter.
- Selection of image contrast enhancement technique.
- Application of 2D directional Wavelet transform to the filtered and enhanced images.
- Training the datasets using different CNN models.
- Selection of CNN models.
- Final image segmentation.
- Performance evaluation using useful quantitative matrices.

3.8 Preparation of Images for Training

3.8.1 Image Labelling

Images were labelled into four different classes: Kidney stone/s, Acoustic Shadow, Kidney Morphology and Outer region (the rest of the image background). The CNN models were trained accordingly. There are two ways to perform image labelling in CNN:

1. Manual labelling: This is the process of manually assigning labels to each pixel in an image. This can be a time-consuming and tedious process but it is the most accurate way to label images.

2. Automatic labelling: This is the process of using a machine-learning algorithm to automatically assign labels to pixels in an image. This is a faster and more efficient way to label images, but it is not as accurate as manual labelling.

In the current study, image labelling was done manually. Images were labelled using image labeller toolbox in Matlab. For the sake of image labelling accuracy, pixel labeller was used in order to label precisely pixel intensity values. Images were labelled to the four different classes with a unique labelling Id. The CNN model finally segments the four distinct objects in a given Ultrasound image and comparison would be made against the available ground truth. Figure 3.1 depicts the labelling information per each of the four classes. The x-axis is the image number while the y-axis shows the percentage of pixels labelled per each of the four classes.

Figure 3.2. Labelling information for each respective class.

3.8.2 Semantic Segmentation - Training using Deep Learning Models

Currently, there are many models which are used for sematic segmentation applications and some of the most popular once include Fully convolutional network, Seg-Net, U-net and Deep labV3. Semantic segmentation is usually applied for pixel level segmentation of images. In the current study, these pixels could belong to any of the four distinct classes.

Most often, semantic segmentation is carried out using CNN, which is useful in wide areas of applications. While using a CNN layer of network in semantic segmentation, a big challenge is the resolution of the input image. Input image with poor resolution becomes difficult to define its feature map when it is convolved in the down sampling network. Before inserting the images into the deep learning models, the characteristics of each model should be analysed including the size of the input image, the number of classes and the image channel arrangement. In addition, the images need to be cropped, primarily to discard unnecessary image information. When training the CNN model, we need to optimize the training loss and the validation loss and minimize under and over fitting issues. Under fitting occurs when the model does not fit the training data model and thus produces a large error and this happens when the validation accuracy is greater than the training accuracy. This may indicate a defect in the model. While in the case of overfitting, the model cannot generalize on new data and in this case, the training accuracy is higher than the validation accuracy.

Data partitioning was performed on the available 250 Ultrasound images with a 70:10:20 ratio for training, validation and testing, respectively. The training data size was further increased to 700 using augmentation techniques during pre-processing as discussed later. The available datasets were then trained using a single GPU with variable training cycles (epochs), batch sizes and the corresponding number of iterations. The entire deep learning algorithm was implemented on a Matlab environment using available toolboxes using a transfer-learning approach. The toolbox is easy to use and allows a layer-by-layer assessment of the used architectures/models. The input layers, the convolutional layers and the output layers are arranged accordingly depending on the deep learning model one wants to implement. The number of convolution filters in the output layers depend on the pooling layer and stride used in the model. In the convolutional layer, spatial information is recovered using up sampling through deconvolution. Two important considerations while selecting a given deep learning model are:

- Its convolutional network layers' parameters, and
- Transforming architectures based on the input image sizes.

3.8.3 Deep Learning Models used in This Study

There are many deep learning models that have been proposed in the literature previously by different researchers. In the current study, three of the popular ones were considered, namely U-Net, Seg-Net and DeepLabV3, and their applicability was tested in automatic detection of kidney stones found on Ultrasound scans of patients.

✓ **U- Net**

U-Net is a CNN architecture that is designed for image segmentation tasks. It was first introduced for biomedical image segmentation. It mainly uses convolutional networks, which are architecturally structured as U. The U-Net architecture is composed of two main parts, a contracting path and an expanding path. The contracting path is responsible for extracting features from the input image, while the expanding path is responsible for reconstructing the image with the segmented objects. The contracting path consists of a series of convolutional layers, each followed by a max-pooling layer. This down sampling process allows the network to learn features at different scales. Whereas, the expanding path consists of a series of convolutional layers, each followed by an up-sampling layer. This up sampling process allows the network to reconstruct the image with the segmented objects. The up and down sampling processes are pictorially depicted in Figure 3.2.

The U-Net architecture has been shown to be very effective for image segmentation tasks. It has been used for a variety of applications, including medical image segmentation, satellite image analysis, and object detection in autonomous vehicles. Here are some of the advantages of U-Net:

- It is able to segment objects at different scales;
- It is relatively easy to train and deploy, and
- It has been shown to be effective on a variety of different datasets.

However, there are some disadvantages of using U-Net deep learning model:

- It can be computationally expensive to train;
- It can be sensitive to the choice of hyper parameters, and
- It may not be as accurate as some other segmentation models in some applications.

Overall, the U-Net architecture is a powerful and versatile CNN architecture for image segmentation tasks. It is a good choice for applications where accuracy and

ease of use are important. U-Net has been used to segment medical images generated on different imaging modalities including CT, MRI and PET. This has been used to improve the diagnosis and treatment of diseases. Besides, U-Net has been used in satellite image analysis to segment and identify different objects, such as buildings, roads, and vegetation. It has been used to monitor changes in the environment and to plan for future development. Currently this network is also used in autonomous object detection in order to detect objects in the environment, such as cars, pedestrians, and cyclists.

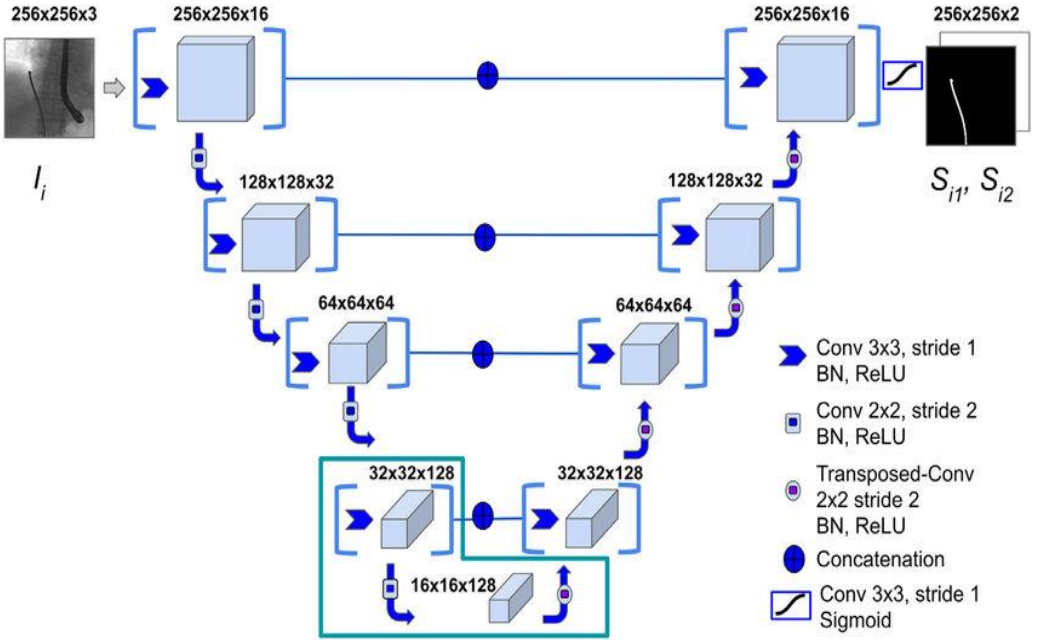

Figure 3.3. Up sampling and down sampling in CNN [64].

✓ **Seg-Net**

Seg-Net is another deep learning model based on CNN. First introduced in the year 2015, this architecture is also specifically designed for image segmentation tasks. The Seg-Net architecture is also composed of two main parts, encoder and decoder. The encoder is responsible for extracting features from the input image while the decoder is responsible for reconstruction of the image with the segmented objects. The encoder consists of a series of convolutional layers, each followed by a max pooling layer. This down sampling process allows the network to learn features at different scales. The decoder consists of a series of convolutional layers, each followed by an up-sampling layer. This up sampling process allows the network to reconstruct the image with the segmented objects.

The Seg-Net architecture is similar to the U-Net architecture, but there are a few key differences. First, the Seg-Net encoder does not use any fully connected layer.

This makes the Seg-Net architecture more efficient and easier to train. Second, the Seg-Net decoder uses a different up sampling method than the U-Net decoder. The Seg-Net decoder uses a transposed convolution layer, which allows the network to learn more accurate spatial information.

The Seg-Net architecture has been shown to be very effective for image segmentation tasks. Like the U-Net model, Seg-Net has also been used for a variety of applications, including road scene understanding, medical image segmentation, and satellite image analysis. The advantages and disadvantages of the Seg-Net model are very similar to that of the U-Net model.

✓ **DeepLabV3**

The DeepLabV3 model is a semantic segmentation model that uses a number of different modules to improve its performance. These modules include:

Atrous convolution: this is a form of convolution that permits the model to see larger region (receptive fields) in the input that produces features without growing the number of parameters. This is crucial for semantic segmentation.

Atrous Spatial Pyramid Pooling (ASPP) Module: this is a module that helps the model to seize context from specific scales. It does this through the use of a series of parallel convolutions. DeepLabV3 makes use of the ASPP module to enhance its overall performance on a selection of datasets (see Figure 3.3).

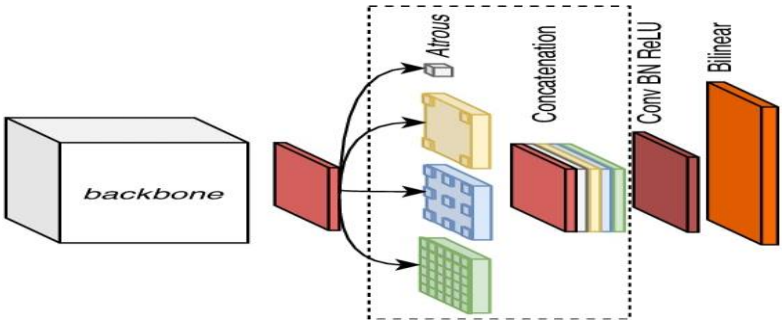

Figure 3.4. Diagram of the Atrous Spatial Pyramid Pooling module [65].

Image-level Features Module: is another module to improve the performance of the DeepLabV3 model. These features are extracted from the entire image and help the model to better understand the context of the objects and to segment them more accurately.

The DeepLabV3 model takes an input image and passes it through a series of convolutional layers. Then the convolutional layers extract features from the image, such as edges, shapes and colors. The ASPP module combines these features

from different scales to create a more comprehensive representation of the image. The image-level features are then added to this representation to help the model understand the context of the objects in the image. Finally, the model uses a classifier to segment the objects in the image. Figure 3.4 presents the DeepLabV3 model architecture.

The DeepLabv3 model has been shown to be very effective for semantic segmentation. Some of the advantages of DeepLabV3 include:

- It is able to capture context from different scales, which allows it to segment objects more accurately;
- It uses image-level features to help it understand the context of the objects in the image, which further improves its accuracy;
- It has been shown to be effective on a variety of datasets.

One known disadvantage of the DeepLavV3 model is that it requires a large amount of data to train.

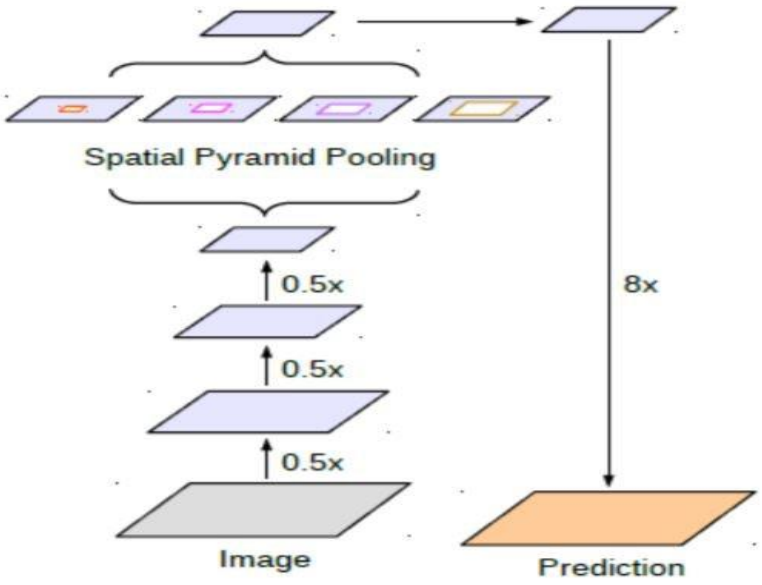

Figure 3.5. DeepLabV3 architectural diagram [66].

The receptive field is increased without increasing the size of the convolution kernel or decreasing the resolution of the images from the backbone network. The ASPP integrates the pointwise convolution mechanism to reduce computation time. Pointwise convolution is a form of convolution using a kernel of size 1x1 iterating through the channels of the image for each pixel of the image. This kernel has a depth equal to the number of channels in the input image. It can be useful

for applying multiple filters with a reduced number of operations and to change the dimensionality of the input.

The pointwise convolution allows to obtain a dilated convolution (or "atrous convolution") by adding a dilation rate r, which defines the spacing between the cells of the kernel (see also Figure 3.5). This allows to increase the size of the kernel, and thus the receiving field, without having to process too many pixels. This results in considerable speed gains without loss of accuracy.

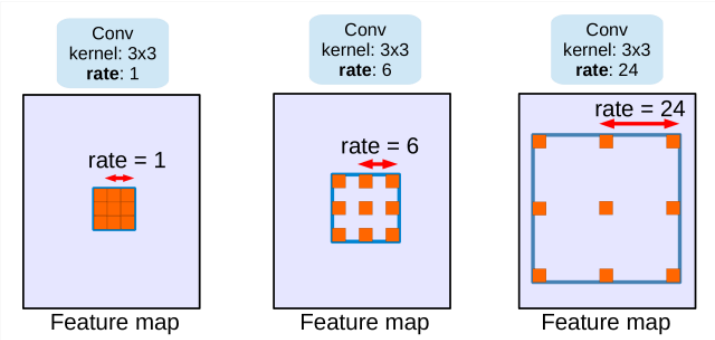

Figure 3.6. Explanatory diagram of the relationship between the expansion rate r and the feature map [66].

Formally, the dilated convolution is written as:

$$Y[i] = \sum_{k=1}^{K} x[i + r * k]w[k] \qquad \text{----------} \qquad \textit{Equation 3.1.}$$

where r is the expansion rate and K is the size of the kernel applied to the input. If r is equal to 1, then we obtain a simple convolution. Figure 3.6 compares a standard and dilated convolutions.

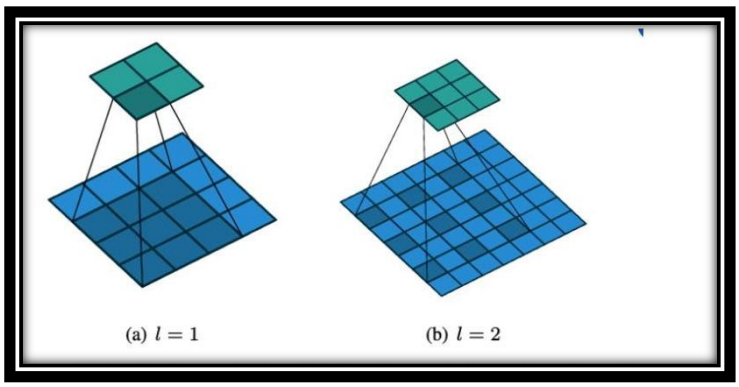

Figure 3.7. Comparison of a standard convolution and a dilated convolution [67].

3.9 Pre-processing

All 250 Ultrasound kidney images (which were originally in DICOM format) were manually cropped in order to remove unnecessary data from the images. Then the images were all resized to size 224x224. All images were then converted

to .jpeg format, so that it fits all deep learning models considered in the current study. Three data augmentation techniques were used to enhance the training data size namely image resize, flip and rotate. All images were then subjected to useful image pre-processing steps.

3.9.1 Image Filtering

Usually ultrasound images are prone to different kind of noises: additive and multiplicative noises. Speckle noise is a typical multiplicative noise that is present in Ultrasound images that results from scattering effect. The additive component arises mostly due to sensor signal tracing capability.

Different filters have been proposed in the literature to deal with the noises that appear in Ultrasound images. Four popular filters have been tested in the current study namely, Wiener, Frost, Homomorphic and Anisotropic diffusion filters, to deal with the noises. These filters are widely used for De-speckling. These filters are believed to maintain fine image details while removing noises. The performance of each filter is evaluated based on PNSR and MSE. Accordingly, the filter with highest PSNR and lowest MSE is selected. The MSE is computed as the average error of the pixels in the entire image. The MSE value per pixel is calculated as:

$$MSE = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}[Q'(i,j)-Q(i,j)]^2}{M*N} \quad \text{----------} \quad \textit{Equation 3.2.}$$

where i and j are pixel positions, Q and Q' are original and enhanced images respectively and (M, N) is the size of the input image.

The PSNR is computed (in decibel, dB) as:

$$\text{PSNR} = 10 * log_{10}\frac{[L-1]^2}{MSE} \quad \text{----------} \quad \textit{Equation 3.3.}$$

where *L-1* is the maximum possible grey value the image could assume. Higher value of PSNR of restored image shows better quality.

Some also use another matrix called Image Enhancement Factor (IEF) to measure filter performance. The IEF is a quantitative measure defined as the ratio of the mean square error before filtering and the mean square error after filtering. Higher IEF denotes not only higher noise reduction but also greater enhancement of the image. The IEF is computed as:

$$IEF = \frac{\sum_{m,n}(Q'(i,j)-X(i,j))^2}{\sum_{m,n}(Q(i,j)-Q'(i,j))^2}$$ ---------- *Equation 3.4.*

where Q′ is the noisy image, X denotes the original noise free image and Q represents the de-noised image.

3.9.2 Image Enhancement

The purpose of image enhancement is primarily to improve the quality of the image. A good image enhancement technique enhances the image retaining its natural look. The applied technique should not enhance the noise level of the input image. This image quality improvement can be measured through visual inspection or using other quantitative matrices. However, there is no comprehensive measure used to show the purpose and effectiveness of an enhancement algorithm. In the current study, two popular measures of enhancement are considered: Effective Measure of Enhancement (EME) and PSNR.

PSNR is the ratio between maximum possible power of the input signal and the power of corrupting noise that affects the quality of the image under consideration. A higher PSNR indicates reconstruction of higher quality image [67]. The PSNR used to quantify the effectiveness of the image enhancement technique is given by:

$$PSNR = 10log_{10}[\frac{[L-1]^2}{\frac{1}{M*N}\sum_i\sum_j|x(i,j)-y(i,j)|^2}]$$ ---------- Equation 3.5.

where *x(i, j)* is the intensity of the input image and *y(i, j)* is the intensity of the processed (enhanced) image.

EME is one of a complex measures of contrast suitable to evaluate contrast enhancement particularly for images that have attributes of non-periodicity, random texture and non-uniform background. EME is considered as an important contrast enhancement evaluation criteria for this kinds of images. From Weber-Fechner Law, contrast ratio for a given image is defined as:

$$C_R = \frac{I_{max}(k,l)}{I_{min}(k,l)}$$ ---------- Equation 3.6.

where I_{max} and I_{min} are the maximum and minimum intensities. Let A be an image with size *M*N*. A is split in to *r *c* blocks and let each block be represented as $B_{k,l}(i,j)$ where *i* and *j* are rows and columns inside each block. The contrast ratio is measured for each block [68]. Then the EME can be computed as:

$$EME_{r,c} = \frac{1}{r*c}\sum_{l=1}^{r}\sum_{k=1}^{c} 20\log(C_R(k,l))$$ ---------- Equation 3.7.

For good segmentation of images using deep learning, image enhancement plays a great role by raising the image's quality. The Ultrasound images considered in the current study show high local intensity variations and this could hide important information in the images. Hence, we need to utilize methods which could be used to normalize the intensity variations. In that regard, different normalization techniques have been developed in the literature. Four of these methods have been tested in the current study: Histogram Equalization (HE), Contrast Limited Adaptive Histogram Equalization (CLAHE), Recursive Mean Separate Histogram Equalization (RMSHE) and a combination of CLAHE and RMSHE. The performance of the enhancement methods is evaluated using PNSR and EME as defined above.

- ✓ **Histogram Equalization (HE)**

HE is an image processing technique that distribute the histogram of an image evenly using a transform function. This technique is commonly employed for image enhancement because of its simplicity and comparatively better performance. The operation of HE is performed by remapping the grey levels of the image based on the probability distribution of the input grey levels. It flattens and stretches the dynamic range of the input image's histogram. Ideally, the output (equalized) image histogram should have equal number of pixels for each discrete grey value assuming a uniform distribution [45].

- ✓ **Contrast Limited Adaptive Histogram Equalization (CLAHE)**

CLAHE operates by splitting the original input image into several sub-images without overlapping (blocks of histogram). These blocks of input image histogram is limited to the value of the improvement per each pixel and then equalization is performed, where the contrast enhancement of an image is applied on small data regions called tiles rather than the entire image. The resulting neighbouring tiles are then stitched back seamlessly using bilinear interpolation. Therefore, the contrast in the homogeneous region can be limited so that noise amplification can be avoided. At the end, it makes the image to appear sharpened and more cohesive. The idea of dividing the image into small blocks increases the contrast in each block while maintaining the contrast between the sub images [69] [70].

✓ **Recursive Mean-Separate Histogram Equalization (RMSHE)**

Like CLAHE, RMSHE works by recursively partitioning the input image histogram into smaller and smaller pieces (squares) and equalize the histogram of each square. The histogram is decomposed up to a recursion level r, generating 2^r sub-histograms. The individual sub histograms are then equalized. This increases the contrast between individual blocks while maintaining the brightness of the image. RMSHE is a good preservation technique when the value of 'r' is large because the output mean converges to the input mean. When 'r' is large, the output histogram will approach the input histogram, which will eventually becomes the exact copy of the input image and there will be no more enhancement [72] [70].

✓ **A Combination of CLAHE and RMSHE**

A combination of CLAHE and RMSHE has also been proposed in the literature before to further improve the contrast and brightness of images [68]. CLAHE can be used to initially equalize the histogram of the image whereas RMSHE performs scalable brightness preservation. This can help to improve the contrast and brightness of the image while preserving the details of the image. The method has shown promises in a variety of applications including image editing, medical imaging, and computer vision. Some potential disadvantages of the hybrid CLAHE-RMSHE method include its higher computational cost and the fact that it may not be suitable for all types of images.

3.9.3 Image Transformation

The 2-D Directional Wavelet Transform (DWT) has been used in a number of studies, in a wide variety of problems that have to do with analysing images [72]. It can be used for the detection of specific features, such as a hierarchical structure, edges, filaments, contours, boundaries between areas of different luminosity, and the like. Gaussian DWT (GDWT) is a wavelet transform that combines the advantages of Gaussian Wavelet Transform and directional wavelet transform [73]. The Gaussian wavelet transform is a wavelet transform that uses the Gaussian mother wavelet as the wavelet function. This makes it ideal for images with smooth edges. The GDWT is used to effectively enhance edges in images [74]. It is also used for effective image noise reduction while preserving image edges and image compression to render the image without losing many details. It also has the property of modifying edge discontinuities.

The 1D Gaussian wavelet (which is basically the derivative of the Gaussian function) is given by:

$$\varphi(x) = xe^{\frac{-x^2}{2\sigma_x{}^2}}$$ ---------- Equation 3.8.

For image processing application, we need wavelets that are 2D. In practice the usual choice for 2D wavelets is to use a product of two 1D functions:

$$\varphi(x,y) = \varphi(x)\varphi(y)$$ ---------- Equation 3.9.

The Fourier transform of the 2D Gaussian wavelet is given by:

$$\hat{\varphi} = -\omega_x\omega_y{}^p e^{\frac{[(\sigma_x\omega_x)^2+(\sigma_y\omega_y)^2]}{2}}$$ ---------- Equation 3.10.

Where $\hat{\varphi}$ is the Fourier transform the 2 D gauss function with respect to x and y here the scaling and angulation are the two fundamental operation we can do to a function by keeping the position translation fixed. In practice, a black and white image will be represented by a bounded non-negative function $0 \leq S(x) \leq M$ the operations that will be applied to an image s are obtained by combining three elementary transformations of the plane, namely, translations, dilations and rotations. These transformations are represented by the following unitary operator in the space $L^2(R^2, d^2x)$ of signals:

$$(\Omega(a,\theta,b)s)(x) = s_{a,\theta,b}(x) = a^{-1}w(a^{-1}r_\theta(x-b))$$ ------- Equation 3.11.

or, equivalently, in the space of Fourier transforms

$$\left(\widehat{\Omega}(a,\theta,b)\hat{s}\right)(k) = \hat{s}_{a,\theta,b}(k) = ae^{-ib.k}s(ar_{-\theta}(k))$$ ------- Equation 3.12.

In these relations the reproducing kernel k, which measures the correlation length in each variable(a,θ,b). b ∈ R2 is the displacement parameter which is used to translate the wavelet, in the time/space axis so that, the wavelet can be centred for any value of x. a > 0 the dilation parameter used for the scaling which has the effect of squeezing and stretching the whole wavelet., θ the rotation angle acts on x = (x, y) as

$$r_\theta(x) = (xcos\theta - ysin\theta, xsin\theta + ycos\theta)$$

Where, $0 \leq \theta \leq 2\pi$

If the function s is rotation invariant, we simply omit the index θ:

$$s_{a,b}(x) = a^{-1}s(a^{-1}(x-b))$$ ------- Equation 3.13.

Similarly, to Fourier transformations, we can use a convolution to calculate the transform. Defining the transform of a function, s(x), as $[W_{(\varphi)}s](a, \theta, b)$, evaluating the integral gives the transform.

$$[W_{(\varphi)}s](a, \theta, b) = \int_{-\infty}^{\infty} s(x, y) \frac{1}{\sqrt{a}} \varphi * \frac{(x-b)}{a} \frac{(y-c)}{a} dxdy \quad \text{--------- Equation 3.14.}$$

Here $\varphi *$ means the complex conjugate of the wavelet function, since the wavelet and the function are in general complex functions. Notice that the transform, $[W_{(\varphi)}s](a, \theta, b)$, is a function of three variables.

The gauss function in Equation 3.10 is a so called a "mother" wavelet. This mother is then scaled and shifted to get all the other possible variations of the transform, $[W_{(\varphi)}s](a, \theta, b)]$.From Plancherel's theorem we can see that

$$\int_{-\infty}^{\infty} f(t)\bar{h}(t)dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega)\, \hat{h}(\omega)dw \quad \text{---------- Equation 3.15.}$$

where $\hat{f}$ is the Fourier transform of f, $\hat{h}$ is the Fourier transform of h, and $\bar{h}$ is the complex conjugate of h . This is the general form of the theorem, and f and h can be complex valued. Here, in $h^{\wedge} * (\omega)$ means the complex conjugate of the Fourier transform of h. Where the hat describes the fourier transform.

$$\mathcal{f}(\varphi(\frac{t-b}{a})) = e^{iu\omega} \mathcal{f}\left(\frac{t}{a}\right) = e^{iu\omega}\, \varphi^{\wedge}(\frac{\omega}{a}) e^{-iu\omega} \quad \text{---------- Equation 3.16.}$$

$$W_{f(a,b)} = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\omega) \frac{1}{\sqrt{a}} \varphi^{\wedge}(\frac{\omega}{a}) e^{-iu\omega} \quad \text{---------- Equation 3.17.}$$

Note: By varying type of wavelet used, we can also vary the output of transformation.

3.10 Deep Learning Model Evaluation

The accuracy of labelling determines the segmentation accuracy of the deep learning, meaning that the models create a feature map from the labelled images. Therefore, while labelling indicating the right classes is mandatory. Once the images are fully labelled, training follows. The trained model will then be validated and tested on a different data set and the performance should be quantitatively measured. In this regard, there are different quantitative matrices used in the literature that are used to check model performance. The once that were used in the current research are global Accuracy, mean Accuracy, Mean IOU, weighted IOU, and mean BF score values.

Intersection Over Union (IOU) - is the most common method used in deep learning for comparing different models. IOU, also known as Jaccard Index, is the ratio of correctly classified pixels to the total number of pixels that are assigned to that class by the ground truth and the predictor giving rise to true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN) as also indicated in Figure 3.7. The IOU can be expressed as:

$$IoU = \frac{TP}{TP + FP + FN} \quad \text{----------} \quad \textit{Equation 3.18.}$$

- ✓ Global Accuracy: is the ratio of correctly classified pixels to total pixels, regardless of class.
- ✓ Mean Accuracy: is the ratio of correctly classified pixels in each class to total pixels, averaged over all classes.
- ✓ Mean IOU: is the average intersection over union (IOU) of all classes.
- ✓ Weighted IOU: is the average IOU of all classes, weighted by the number of pixels in the class.
- ✓ Mean BF Score: is the boundary F1 score for each class, averaged over all images.

3.11 Confusion Matrix

In the current study, a normalized confusion matrix was used specified as a table with rows and columns, which are equal to the number of classes used in training a specific model. Each table element (*i,j*) is the count of pixels known to actual class *i* and predicted class *j* as also indicated in Table 3.2.

Table 3.2. The relation between actual and predictive values.

		Predicted value	
Actual value		Positive	Negative
	Positive	True Positive	False Negative
	Negative	False Positive	True Negative

The confusion matrix dictates the calculation of model sensitivity, specificity and over all accuracy as defined below:

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \quad \text{----------} \quad \textit{Equation 3.19.}$$

$$Sensitivity = \frac{TP}{TP+FN} \quad \text{----------} \quad \textit{Equation 3.20.}$$

$$Specificity = \frac{TN}{TN+FP}$$ ---------- *Equation 3.21*.

In our context, true positives indicates the number of cases who have kidney stone/s and the test correctly identifies them as having the stone whereas the false negatives values indicate the number of cases who have kidney stone/s but the test incorrectly identifies them as not having. False positive and true negatives are similarly defined.

4 RESULTS AND DICSUSSION

4.1 Noise Filtering

The four noise reduction filters namely Wiener, Frost, Homomorphic and Anisotropic filters were tested for their efficacy in dealing with the noises that come with the Ultrasound images. All filters were implemented on a Matlab environment (Matlab, version 2021a). Figure 4.1 presents an Ultrasound test image and results obtained after the application of the four filters.

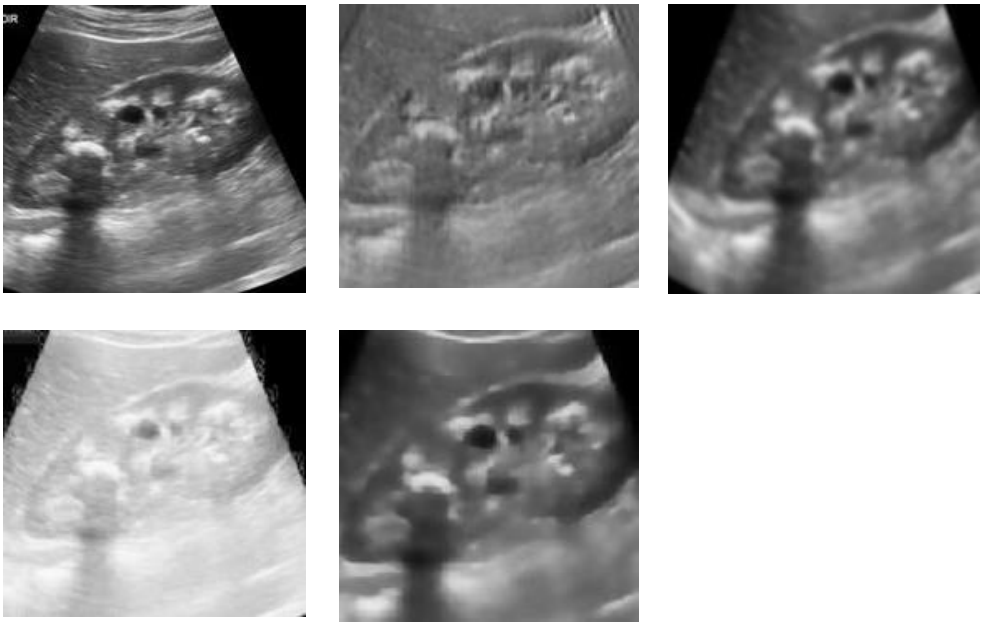

Figure 4.1. Top: Original image (left), Weiner filtered (middle), and Frost filtered (right); Bottom: Homomorphic filtered (left), Anisotropic filtered (middle).

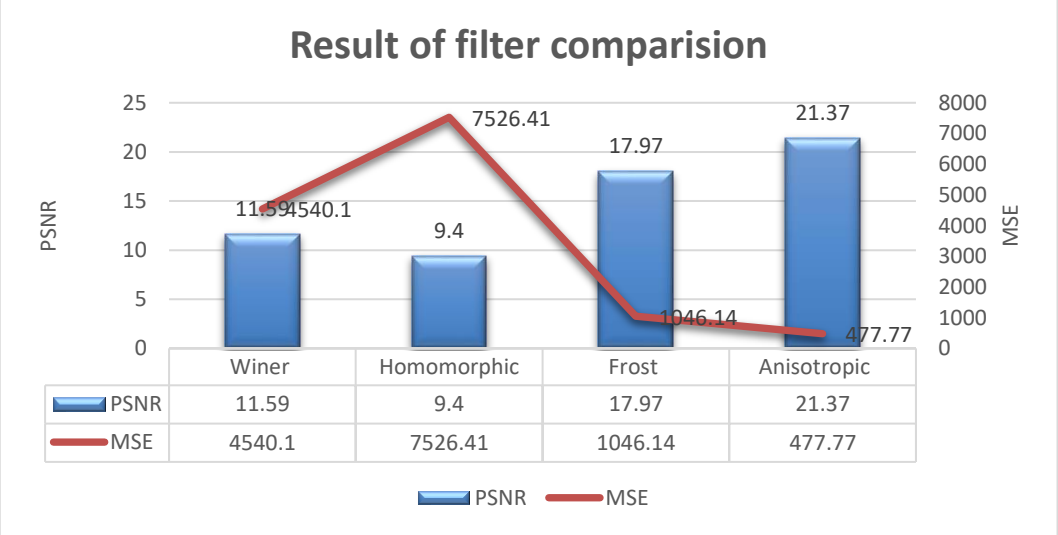

Figure 4.2. Comparison of the four filters using PSNR and MSE values.

As explained in the methodology section in the previous chapter, the performance of these filters was quantitatively evaluated using PSNR and MSE. Figure 4.2 presents the computed PSNR and MSE values for each of these filters for a typical Ultrasound image considered in the current study. Based on the results, Anisotropic filter offered the maximum PSNR and minimum MSE and is therefore has the best filter performance. Similar results were obtained on all 250 Ultrasound images acquired from the available dataset. Accordingly, all images had to pass through Anisotropic filter before further processing.

4.2 Contrast Enhancement

All Ultrasound images filtered with Anisotropic filter were further enhanced using contrast enhancement techniques described in the previous chapter (HE, CLAHE, RMSHE and combination of CLAHE and RMSHE). Figure 4.3 presents the results obtained after the application of these four enhancement techniques. The performance of each enhancement technique was evaluated using the PSNR and EME values computed using the method discussed in the previous chapter. Figure 4.4 shows the comparison between the contrast enhancement techniques applied on a representative filtered image based on these two performance matrices. As can be seen on the results, in terms of both PSNR and EME values, the combined CLAHE-RMSHE method outperformed the rest three enhancement techniques. It is also clear from Figure 4.4 that the resulting histogram after the application of the hybrid CLAHE-RMSHE approach assumed a much bigger dynamic range with a distribution which is close to uniform than the histograms derived after the application of the rest three enhancement methods. Similar results were obtained for all Ultrasound images considered in the current study. Accordingly, all images once filtered using the Anisotropic filter, had to pass through the combined CLAHE-RMSHE enhancement method before further processing. It was these filtered and enhanced images that were put into the 2D Gauss directional wavelet transform for further processing before they were used as inputs to the deep learning models employed in the current study to be able to perform a final segmentation of the kidney stones seen on the Ultrasound images. Note that the entire stone detection algorithm was implemented on a Matlab environment.

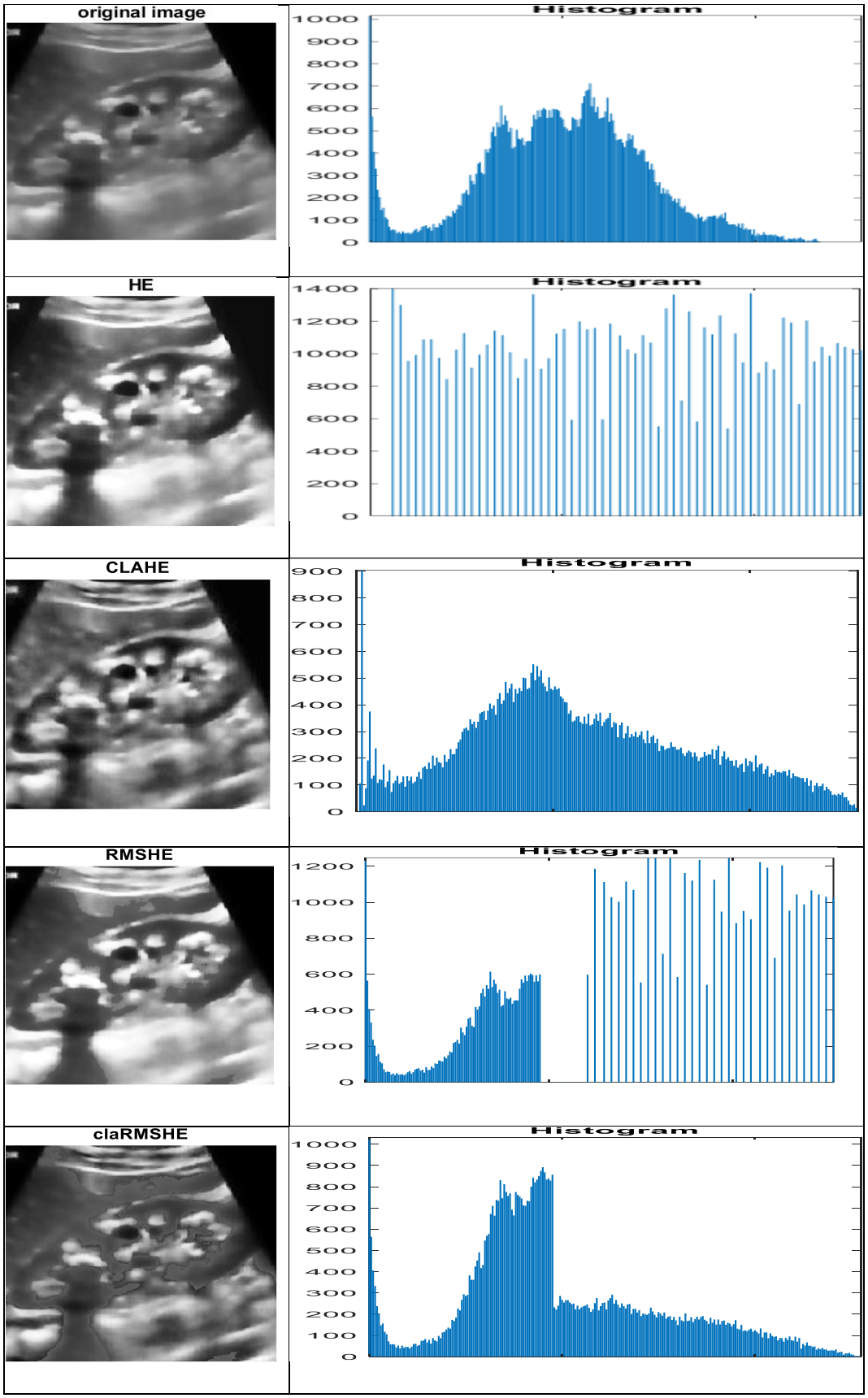

Figure 4.3. Original filtered image and its histogram (top), HE applied (second), CLAHE applied (third), RMSHE applied (fourth), combination of CLAHE and RMSHE applied (bottom).

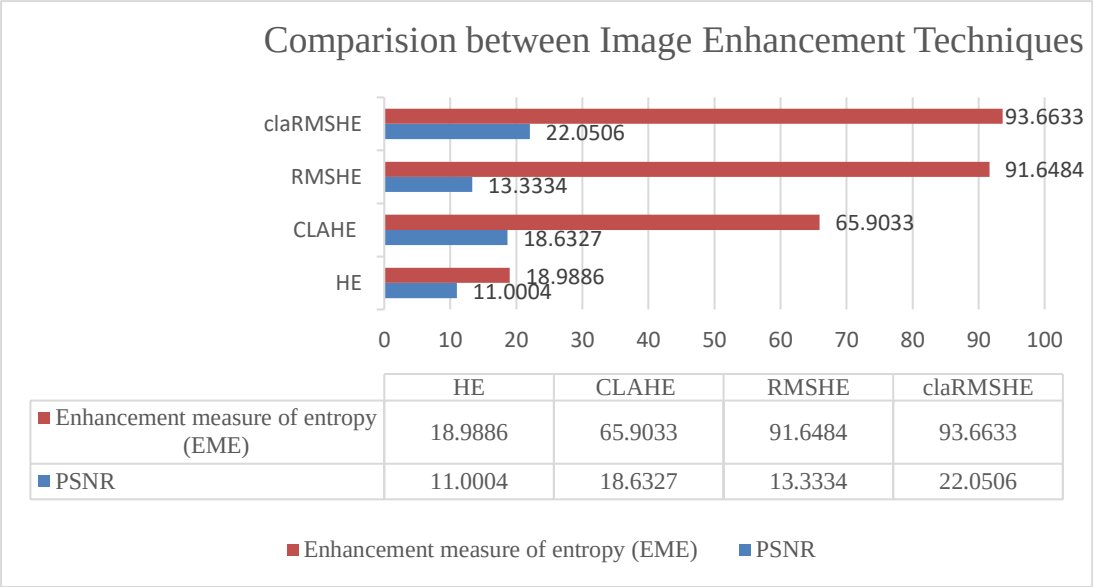

Figure 4.4. Comparison between Image enhancement techniques.

4.3 2D Directional Wavelet Transform

As explained in the previous chapter, the 2D GDWT works by decomposing an image into a set of wavelet coefficients. These coefficients represent the image at different scales and orientations. Then the resulting image was feed to deep learning neural network through labelling different region of interest. Figure 4.5 demonstrates the effect of applying the 2D GDWT on a typical Ultrasound image. The 2D GDWT is generally complex valued and assumed modulus (amplitude) and angle (phase). Figure 4.5 displays the modulus, angle, real and imaginary components of the 2D GDWT of a pre-processed Ultrasound image at different scales and certain orientation ($\theta = \frac{7\pi}{4}$). Same analysis was performed on all 250 Ultrasound images. The best visual information was obtained in the real component of the 2D GDWT at scaling value of 2 and orientation at $\frac{7\pi}{4}$. The images that are filtered using Anisotropic filter, enhanced by the hybrid CLAHE-RMSHE technique and 2D GDWT transformed with the respective parameters were inputted to the deep learning schemes for training. As explained in the previous chapter, three known deep learning sematic segmentation schemes were tried in the current study for their efficacy in effectively segment kidney stones observed on Ultrasound images.

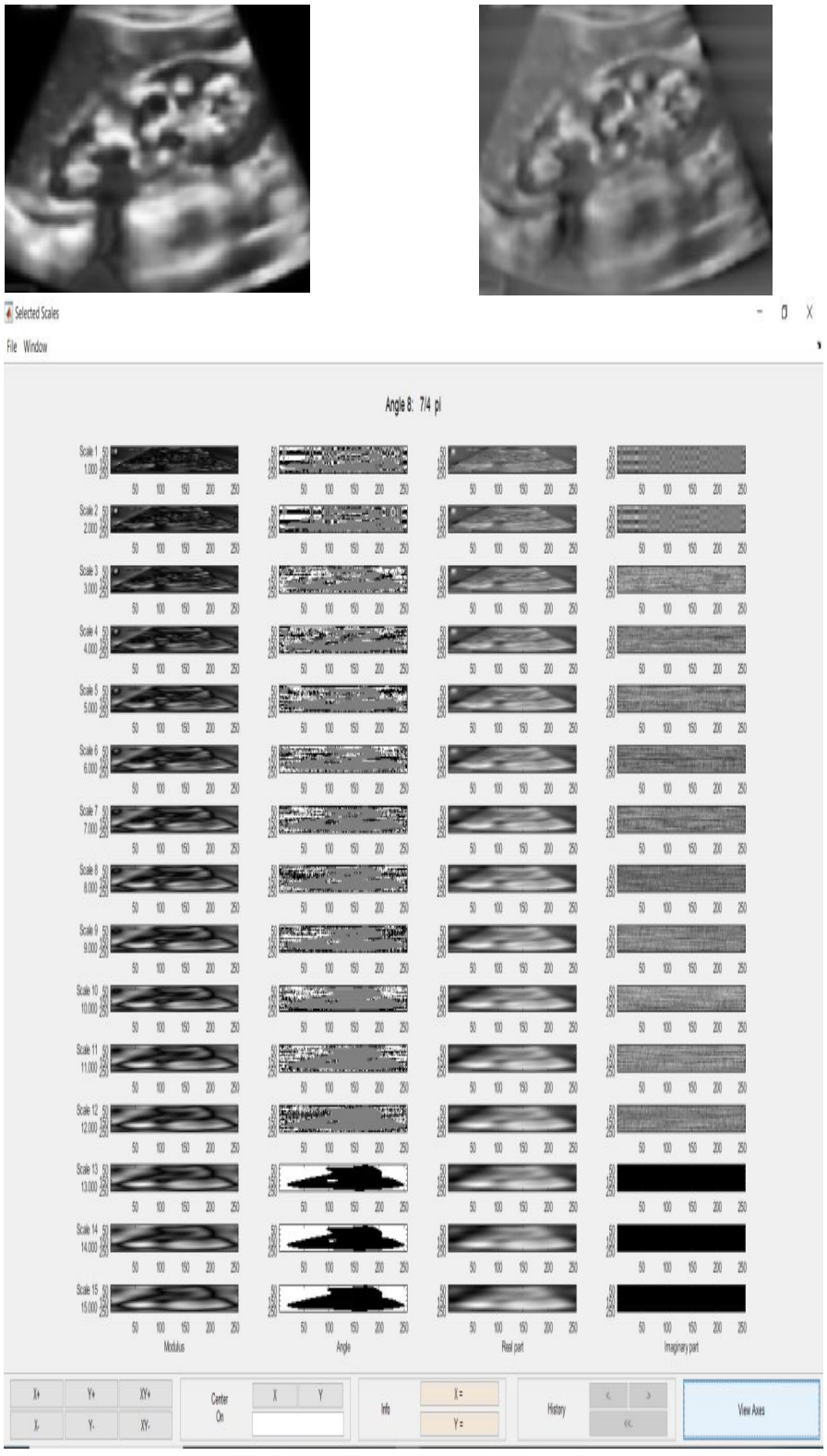

Figure 4.5. Original pre-processed image (top left), 2D GDWT transformed image (top right), image at different scale, translation and angular rotation values and corresponding transformed images.

4.4 Semantic Segmentation Results

This is the final part of the kidney segmentation scheme where the deep learning models (U-Net, Seg-Net and DeepLabV3) are trained, validated and tested with different deep learning model parameters including learning rate, batch size and epoch level. The models were implemented on a GPU resource due to the computational requirement to run the deep learning models. Accordingly, the best model parameters were found out through several training experiments. Table 4.1 lists the best parameter combinations that offered the best model training accuracy for all the three deep learning models.

Table 4.1. Used CNN training parameters.

Training Options	Stochastic gradient descent with momentum
Learning Rate Schedule	Piecewise
Learning Rate	0.01
Learning Rate Drop Period	10
Max Epochs	20
Mini Batch Size	8
Used hard ware resource	GPU

After training for 2 hours with batch size of 8 and epoch size of 20, a validation accuracy of 87.93%, 88.10% and 95.79% for U-net, Seg-net and DeepLabV3 deep learning models, respectively, showing a superior performance by the DeepLabV3 model compared to the others. These results were obtained by applying the models on the Ultrasound images which were pre-processed by the schemes discussed in the previous chapter. The same models were also tested on raw (un-processed) Ultrasound images and all the models offered much inferior results. Figure 4.6, Figure 4.7 and Figure 4.8 present the training and validation accuracies as well as the corresponding losses for each of the three models. Table 4.2 summarizes the final testing performance of the models in terms of Global accuracy, Mean accuracy, Mean IOU, Weighted IOU, BF scores with and without image pre-processing. The accuracies per each of the four distinct classes (Kidney stone, Acoustic shadow, Kidney morphology and the rest of the background) is summarized in Table 4.3. Note that testing was done on 50 Ultrasound images (20% of the total data set).

Figure 4.6.Comparision of validation accuracy between pre-processed and un-processed image for DeepLabV3 model.

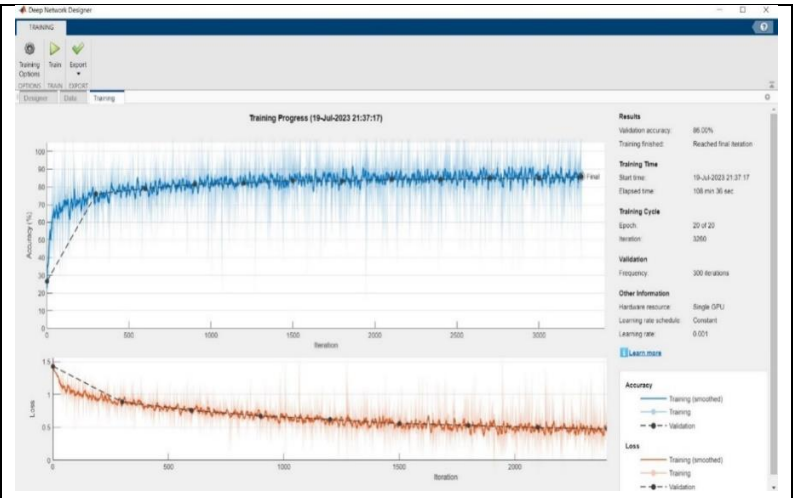

Figure 4.7. Comparison of validation accuracy between pre-processed and un-processed image for Seg-Net model.

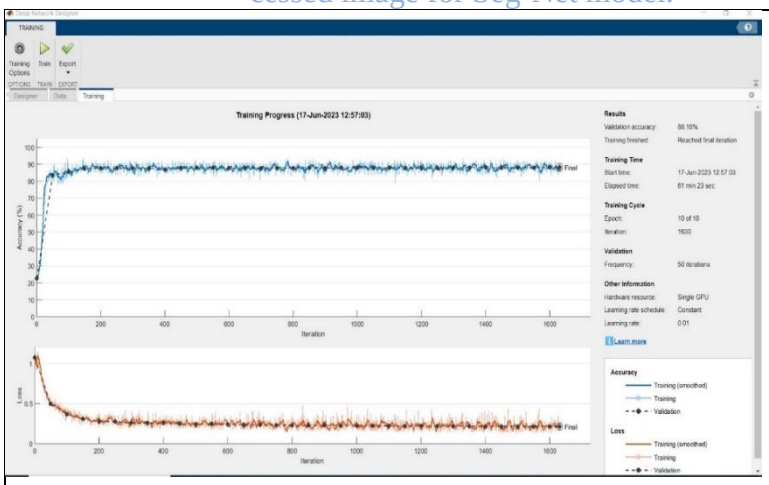

Figure 4.8. Comparison of Validation accuracy between pre-processed and un-processed image for U-Net model.

Table 4.2. Overall evaluation of DeepLabV3, Seg-Net, and U-Net models.

Performance matrices results for DeepLabV3, U-Net and Seg-Net models.						
	Processed image			Un-processed image		
Evaluating Matrices	DeepLabV3	Seg-Net	U-Net	DeepLabV3	Seg-Net	U-Net
Global accuracy	0.95281	0.93125	0.88148	0.86006	0.8503	0.81348
Mean accuracy	0.80133	0.5633	0.25	0.62777	0.45255	0.39201
Mean IOU	0.72847	0.48288	0.22037	0.502	0.38952	0.32555
Weighted IOU	0.91321	0.8799	0.77701	0.76461	0.74267	0.68695
BF scores	0.82543	0.71797	0.6569	0.50706	0.38134	0.35946

Table 4.3. Evaluation of classification performance using accuracy and IOU.

Accuracy and mean IOU results per each class						
Class names	Deep labV3 model		Seg-net model		U-net model	
	Accuracy	IOU	Accuracy	IOU	Accuracy	IOU
Kidney stone	0.79193	0.6924	0.62774	0.096321	0	0
Acoustic shadow	0.72259	0.62197	0.008244	0.008158	0	0
Kidney morphology	0.74157	0.62097	0.081347	0.070625	0	0
Outer region	0.97976	0.94929	0.98473	0.92328	1	0.88148

4.5 Confusion Matrix

According to results presented above, DeepLabV3 model offered the best results. In order to compute the specificity, sensitivity and accuracy of the model, a confusion matrix was derived, shown in Figure 4.9.

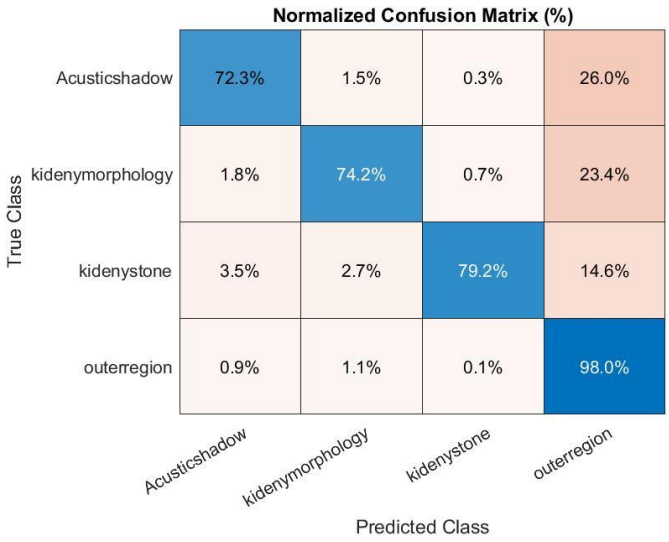

Figure 4.9. Confusion matrix generated for the DeepLabV3 model.

Based on the confusion matrix, the sensitivities, specificities and accuracies were computed per each of the four classes. Based on the calculations, the sensitivities, and specificities or kidney stone, acoustic shadow, kidney morphology and the rest of the background are 0.792, 0.9963; 0.72227, 0.9793; 0.742, 0.9823; 0.98, 0.786, respectively for the DeepLabV3 model. The overall segmentation accuracy of the model was calculated to be 0.945.

4.6 Final Semantically Segmented Results

Figure 4.10 presents the final segmentation outputs generated using the three deep learning models for selected ultrasound test images. Each of the four classes are labeled using different colors: yellow for kidney stone, blue for acoustic shadow, cyan for kidney morphology and red for the rest of the background showing the superiority of the DeepLabV3 model compared to the rest of the models. The test images used in Figure 4.10 are composed of Ultrasound images with one (rows 3 and 5) or multiple/two (rows 2, 4 and 6) kidney stones as well as those images scanned from control subjects with no manifestation of kidney stones (row 1). The best deep learning model used in the current study, DeepLabV3, was able to show stone signatures in all except one (row 3) of the cases. There were instances where the model was able to find the acoustic shadow but miss one of the stones (row 2). In instance, the model correctly found a stone with its acoustic shadow but miss the second stone and its shadow (row 4). In the rest instances, the model was able to detect the stone/s and their acoustic shadows correctly. The proposed deep learning scheme has showed good promises in automatically detecting kidney stones based on Ultrasound images and differentiating it from the acoustic shadows, the kidney and the rest of the background. It is clear that the pre-processing steps implemented in the current study had significantly improved the performance of all three deep learning models tested in the current study, though the performance of DeepLabV3 is much superior to the rest models. The overall accuracy of the deep learning models preceded by the pre-processing methods (filtering, image enhancement and directional wavelet transform) was acceptable in the sense that there are still rooms for improvement.

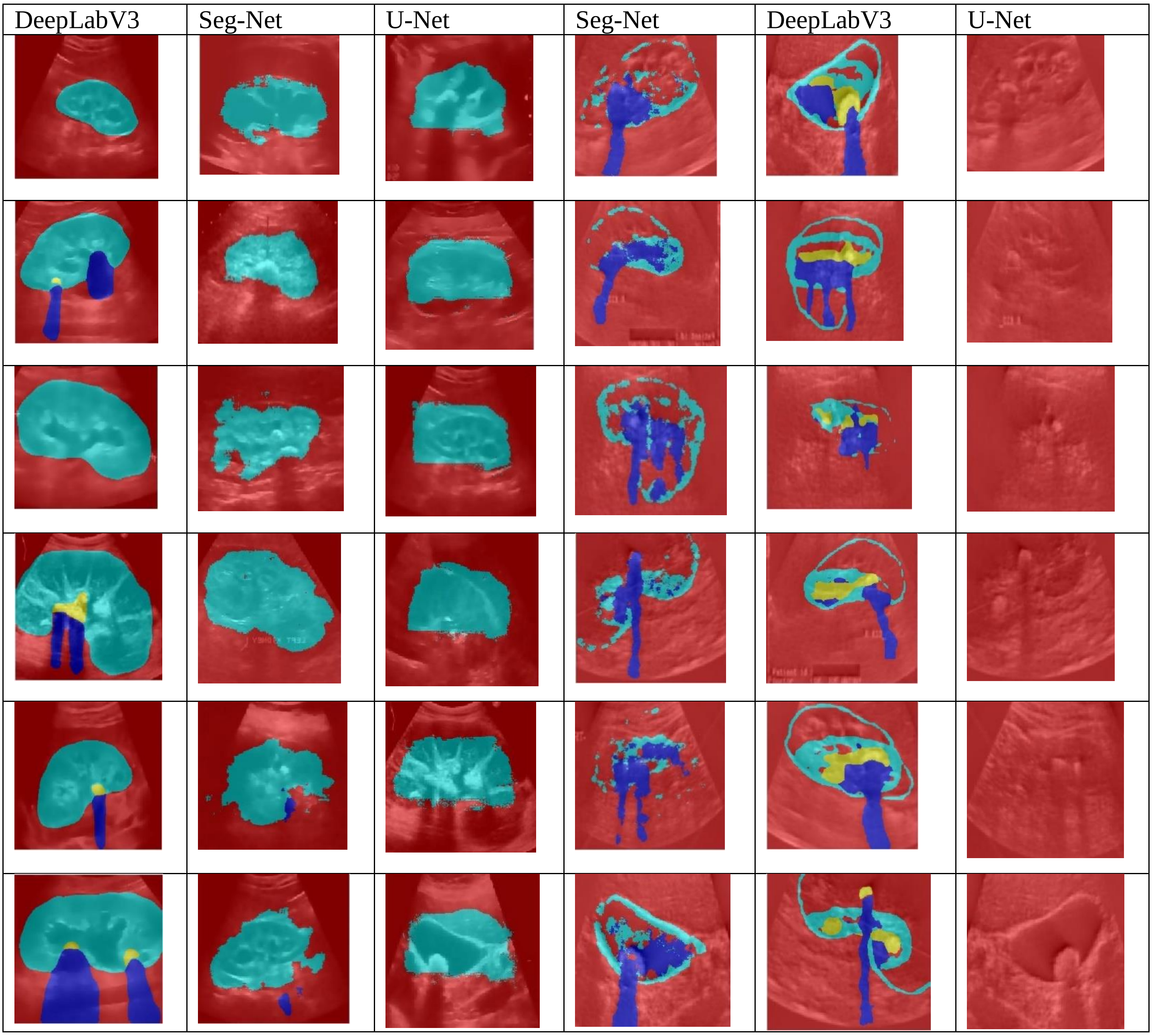

Figure 4.10. Semantically segmented images: with pre-processing (C 1:3) and without (C 3:6).

5 CONCLUSION

This study has tried to develop and implement a new deep learning approach for segmenting kidney stones identified on Ultrasound images. Given the fact that Ultrasound images come with noise and other artifacts, a sequence of pre-processing steps were considered. The results derived are clear indications that pre-processing has indeed played vital roles in improving the effectiveness of the deep learning schemes. Different filters and image enhancement methods were tested during pre-processing and their effectiveness was checked based on useful quantitative matrices. Accordingly, Anisotropic filter for noise reduction (focusing on speckle types) as well as a hybrid CLAHE-RMSHE for image enhancement were found to work best.

Three popular deep learning models, namely U-Net, Seg-Net and DeepLabV3 were used or segmenting the ultrasound images assuming multiple classes (four classes including kidney stones, acoustic shadow, kidney morphology and the rest of the background). The global accuracy the best model, DeepLabV3, stood at 95.3% with mean IOU value of 72.8%. Compared to performances reported in the literature by previous scholars who developed different kidney stone detection algorithms, the proposed method offered results that showed great promises. A previous study that made use of Thresholding reported a mean IOU of 60.0% and accuracy of 85% in detecting kidney stones. Another study that exploited the a region growing approach reported an IOU of 70.0% and accuracy of 90.0%. Another one made use of ege detection to detect kidney stones and reported an IOU of 75.0% and accuracy of 92.0%. There was another study that exploited the deep learning models and obtained an IOU of 87.7% and accuracy of 91.2%. The performance of this last study indeed superior to the results obtained in the current study. The major problem with all these methods listed above is that they made use CT images to develop, train and test their algorithms. There were also some that made use of X-ray images. It was one of the major intents in the current study to develop and test a kidney stone detection scheme based on imaging modalities that do not make use ionizing radiation. Availability of the imaging devices was another criterion. In that regard, there were some studies reported in the literature (as explained in the literature review in this thesis) that tried to exploit information available on Ultrasound images.

One of these studies, for example, reported a global accuracy of 92.57% using Ultrasound images to detect kidney stones. Another study also reported a global accuracy of 92.1% and mean IOU of 87.7% for Ultrasound based kidney stone segmentation based on a dataset composed of 100 images.

Based on these reports, the performance level of the method that is proposed in the current study can be considered very promising. There are, however, rooms for improvements. The overall class based (considering the four classes) performance of the proposed method is still low, particularly in terms of its sensitivity. Whether the issue is with the inadequacy of the pre-processing methods used or the capability of the deep learning models themselves requires further investigation. Furthermore, the efficacy of the algorithm needs to be checked on a larger dataset before any generalizations could be derived. There is also a trend in the technological advancement of the Ultrasound imaging protocol itself that could be exploited as the new versions of the imaging technology which come with unprecedented spatial resolution and reasonable temporal resolution. Hence, more useful image details could allow us capture the images with a better quality, which in effect could enhance the overall performance of the segmentation models. The much-needed clinical implications of the study could also be studied. Such and similar other issues are beyond the scope of the current study and require much further investigation.

6 RECOMMENDATION

In terms of classification performance, the computed sensitivity values were not as high while the specificities are quite promising. Whether the shortcoming is with the adequacy of the pre-processing methods used or the capability of the deep learning models themselves requires further investigation. Another important issue is the size and quality of the image dataset used in the current study. Acquiring a good quality and large number of datasets for deep learning and computer vision applications can seriously influence the robustness of the model. The total number of Ultrasound images acquired were only 284 and 34 of these could not satisfy the inclusion criteria, and all were collected from two local hospitals. This size is indeed considered low for most deep learning applications. Having good data, archival systems in our local hospitals could enhance the data collection process as well as research data availability and adequacy. This is to mean the proposed stone detection algorithm might need to be tested on a larger dataset where its efficacy can be further checked.

7 REFERENCES

[1] V. N. Ratkalkar and J. G. Kleinman, “Mechanisms of stone formation,” Clinical reviews in bone and mineral metabolism, vol. 9, no. 3-4, pp. 187-197, 2011.

[2] R. K. Hsu and N. R. Powe, “Recent trends in the prevalence of chronic kidney disease: not the same old song,” Current opinion in nephrology and hypertension, vol. 26, no. 3, pp. 187-196, 2017.

[3] W. S. McDougal , A. J. Wein, L. R. Kavoussi, A. W. Partin and C. A. Peters, Campbell-walsh urology 11th Edition Review E-Book, Campbell, California: Elsevier Health Sciences, 2015.

[4] D. Andualem and G. Gidena, “Admission patterns and management of urolithiasis a hospital based study in Tikur Anbessa Specialized Hospital (TASH),” East and Central African Journal of Surgery, vol. 19, no. 3, pp. 29-34, 2014.

[5] I. Sorokin, C. Mamoulakis, K. Miyazawa, A. Rodgers , J. Talati and Y. Lotan, “Epidemiology of stone disease across the world,” World J Urol, vol. 35, no. 9, p. 1301–1320 , 2017.

[6] S. Mohammed, B. Yohannes , A. Tegegne and . K. Abebe, “Presentation and surgical outcome at a tertiary care hospital in Ethiopia,” Research and Reports in Urology, Dec 8,2020.

[7] A. Sertsu, W. Teshager, F. Gelana and A. K. Tura, “A cross-sectional study desig on Prevalence of chronic kidney disease and associated factors among patients visiting renal unit of St. Paul’s Hospital Millennium Medical College, Addis Ababa, Ethiopia,” SAGE Open Medicine, vol. 10, 2022.

[8] S. M. Penny , Examination Review for Ultrasound: Abdomen and Obstetrics & Gynecology., ippincott Williams & Wilkins, 2022 Sep 1.

[9] W. Brisbane, M. R. Bailey and M. D. Sorensen, “An overview of kidney stone imaging techniques,” Nature Reviews Urology, vol. 13, no. 11, pp. 654-662, Nov 2016.

[10] P. ICoR, “recommendations of the International Commission on Radiological Protection,” Ann. ICRP , 1990.

[11] R. Semelka, D. Armao, J. Elias and W. Huda , “Imaging strategies to reduce the risk of radiation in CT studies, including selective substitution with MRI,”

ournal of Magnetic Resonance Imaging: an Official Journal of the International Society for Magnetic Resonance in Medicine, vol. 25, no. 5, pp. 900-909, 2007.

[12] M. Straub and A. Hesse, “"Metabolic Evaluation and Metaphylaxis of Stone Disease",” Therapeutic Energy Application in Urology: Standards and Recent Developments, vol. 72, 2005.

[13] M. Kim, Y. Jihye, C. Yongwon, S. Keewon , J. Ryoungwoo , j. B. Hyun and K. Namkug , “Deep learning in medical imaging,” Neurospine, vol. 16, no. 4 , p. 657, 2019.

[14] G. Kossoff , “Basic physics and imaging characteristics of ultrasound,” World journal of surgery, vol. 24, no. 2, pp. 134-42, 2000 Feb.

[15] A. Kapoor and T. Singh , “Speckle reducing filtering for ultrasound images,” in International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud, 2017.

[16] S. Tania and R. Rowaida, “A comparative study of various image filtering techniques for removing various noisy pixels in aerial image,” International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 9, no. 3, pp. 113-24, 2016 Mar.

[17] S. Sheela , M. Sumathi and G. Priya , “Comparative analysis of various filtering techniques for Speckle noise suppression in,” International Journal of Control Theory and Applications, vol. 9, no. 2, pp. 419-33, 2016.

[18] M. E. Alonso, “Speckle noise reduction in SAR images using different similarity measures,” Remote Sensing, vol. 7, no. 1, p. 111, 2015.

[19] M. Gupta , H. Taneja and L. Chand, “Performance enhancement and anal-ysis of filters in ultrasound image denoising,” Procedia computer science, vol. 132, pp. 643-652, 2018.

[20] Bushberg, T. Jerrold and M. B. John, The essential physics of medical imaging, Lippincott Williams & Wilkins, 2011.

[21] F. Benzarti and H. Amiri, “Speckle noise reduction in medical ultrasound images,” arXiv preprint, pp. 1305-1344, 2013 May 6.

[22] O. V. Michailovich and A. Tannenbaum , “Despeckling of medical ultrasound images.,” ieee transactions on ultrasonics, ferroelectrics, and frequency control, vol. 53, no. 1, pp. 64-78, 2006.

[23] P. S. Hiremath , T. Prema and S. Badiger , “Speckle noise reduction in medical ultrasound images,” Advancements and breakthroughs in ultrasound imaging, vol. 1, no. 8, pp. 1-8, 2013 Jun 5.

[24] D. L. Donoho , “Denoising by soft-thresholding,” IEEE Trans Inform, vol. 41, no. 3, pp. 613-627, 1995.

[25] L. Gagnon and A. Jouan, “Speckle filtering of SAR images a comparative study between complex wavelet based and standard filters,” Proc SPIE, vol. 3169, p. 80– 91, 1997.

[26] Z. Shi and K. B. Fung , “A Comparison of Digital Speckle Filters,” in Proceedings of IGRASS’94, Pasadena, 1994 Aug 8.

[27] A. Thakur and R. S. Anand, “Image quality based comparative evaluation of wavelet filters in ultrasound speckle reduction.,” Digital Signal Processing, vol. 15, no. 5, pp. 455- 465, 2005.

[28] Y. Yue , M. M. Croitoru, A. Bidani , J. B. Zwischenberger and J. W. Clark, “Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images.,” IEEE transactions on medical imaging, vol. 25, no. 3, pp. 297-311, 2006 Feb 27.

[29] P. A. Jean, “Two dimensional directional wavelets and image processing,” in InProceedings IWISP'96, 1996 Jan 1.

[30] J. P. Antoine , R. Murenzi and P. Vandergheynst , “Directional wavelets revisited: Cauchy wavelets and symmetry detection in patterns,” Applied and Computational Harmonic Analysis, vol. 6, no. 3, pp. 314-45, 1999 May 1.

[31] X. Yang , J. Zhang , B. Peng and S. You , “An adaptive edge enhancement method based on histogram matching,” in IEEE International Conference on Computational and Information Sciences, 2010.

[32] X. Li and D. C. Liu , “Ultrasound image enhancement using dynamic filtering.,” in In Fourth International Conference on Image and Graphics , 2007.

[33] M. N. Khan and A. Altalbe, Experimental evaluation of filters used for removing speckle noise and enhancing ultrasound image quality., vol. 73, Biomedical Signal Processing and Control., 2022 Mar 1, pp. 103-399.

[34] P. Singh and R. Shree , “A new homomorphic and method noise thresholding based despeckling of SAR image using anisotropic diffusion,” Journal of King Saud University-Computer and Information Sciences, vol. 32, no. 1, pp. 137-148, 2020 Jan 1.

[35] J. Jaybhay and R. Shastri, “A study of speckle noise reduction filters,” signal & image processing: An international Journal (SIPIJ), vol. 3, pp. 71-80, 2015 Jun6.

[36] F. Baselice, G. Ferraioli , M. Ambrosanio and V. Pascazio , “Enhanced Wiener filter for ultrasound image restoration,” Computer methods and programs in biomedicine, vol. 153, pp. 71-81, 2018 Jan 1.

[37] G. Mageshkumar, S. Suthagar and K. S. Tamilselvan, “Performance comparison of adaptive filters for speckle noise reduction in SAR images,” in International Conference on Intelligent Computing and Communication for Smart World (I2C2SW), 2018.

[38] S. Pradeep and P. Nirmaladevi, “A review on speckle noise reduction techniques in ultrasound medical images based on spatial domain, transform domain and CNN methods,” IOP Publishing,InIOP conference series: materials science and engineering, vol. 1055, no. 1, 2021 Feb 1.

[39] E. Ponmani and P. Saravanan , “Image denoising and despeckling methods for SAR images to improve image enhancement performance,” Multimedia Tools and appliccation, vol. 80, no. 17, pp. 26547-69, 2021 Jul.

[40] P. Perona and J. Malik , “Scale-space and edge detection using anisotropic diffusion,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 12, no. 7, pp. 629-639, 1990.

[41] Y. Zhang and ., et al. , “Speckle reduction anisotropic diffusion (SRAD) filter for ultrasound images.,” Journal of Medical Imaging,, vol. 3(1), no. , 011003., (2010)..

[42] S. Jain and R. Ray, “Non-linear diffusion models for despeckling of images,achievements and future challenges,” IETE Technical Review, vol. 37, no. 1, pp. 66-82, 2020 Jan 2.

[43] W. Hafizah and E. Supriyanto , “Comparative evaluation of ultrasound kidney image enhancement techniques,” International Journal of Computer Applications., vol. 21, no. 7, pp. 15-9, 2011.

[44] Z. Hu , B. Liu and C. R. Wang, “Image enhancement algorithm combines maximum gray frequency restrict with dynamic histogram equalization,” 电子与信息学报, vol. 31, no. 6, pp. 1327-31, 2009 Jun 19.

[45] A. M. Grigoryan and S. A. Sos, “Gradient based histogram equalization in grayscale image enhancement,” In Mobile Multimedia/Image Processing, Security, and Applications, vol. 10993, pp. 132-142, 2019 May 13.

[46] W. M. Abdullah Al, M. H. Kabir , M. A. Dewan and O. Chae , “A dynamic histogram equalization for image contrast enhancement,” IEEE transactions on consumer electronics, vol. 53, no. 2, pp. 593-600, 2007 May.

[47] C. Zuo , Q. Chen and X. Sui, “Range limited bi-histogram equalization for image contrast enhancement,” Optik, vol. 124, no. 5, p. 425–431, 2013.

[48] C. Sun, S. Ruan, M. C. Shie and T. W. Pai , “Dynamic contrast enhancement based on histogram specification,” IEEE Transactions on Consumer Electronics, vol. 51, no. 4, p. 1300–1305, 2005.

[49] Y. T. Kim, “Contrast enhancement using brightness preserving bihistogram equalization,” IEEE Trans. Consumer Electron, vol. 43, no. 1, p. 1–8, 1997.

[50] Y. Wang, Q. Chen and B. Zhang , “Image enhancement based on equal area dualistic sub-image histogram equalization method,” IEEE Tran. Consumer Electron, vol. 45, no. 1, p. 68–75, 1999.

[51] S. D. Chen and A. R. Ramli , “Minimum mean brightness error bihistogram equalization in contrast enhancement,” IEEE Trans. Consumer Electron, vol. 49, no. 4, p. 1310–1319, 2003.

[52] X. Liu , L. Song , S. Liu and Y. Zhang, “A Review of Deep-Learning-Based Medical Image Segmentation Methods,” Sustainability , vol. 13, no. 3, pp. 2416-2425, 2021.

[53] Y. LeCun, Y. Bengio and G. Hinton, “Deep learning,” Intelligent Control and Automation,, vol. 7, no. 4, pp. 436-444, 15 November 2016.

[54] L. Geert, K. Thijs , E. B. Babak , A. Arnaud, S. Adiyoso , C. Francesco , G. Mohsen, . A. Jeroen, M. W, L. van der , v. G. Bram and I. S. Clara, “A survey on deep learning in medical image analysis,” Medical Image Analysis, vol. 42, pp. 60-88, 2017.

[55] K. M. Black , H. Law, A. Aldoukhi , J. Deng and K. R. Ghani , “Deep learning computer vision algorithm for detecting kidney stone composition,” BJU international, vol. 25, no. 6, pp. 920-940, 2020 Jun1.

[56] M. Längkvist , J. Jendeberg, P. Thunberg , A. Loutfi and M. Lidén, “Computer aided detection of ureteral stones in thin slice computed tomography volumes using convolutional neural networks,” Comput Biol Med, vol. 97, pp. 153-160, 2018.

[57] L. Fitri, . F. Haryanto, H. Arimura, C. YunHao, K. Ninomiya, R. Nakano, M. Haekal, Y. Warty and U. Fauzi, “Automated classification of urinary stones based on microcomputed tomography images using convolutional neural network,” Phys. Med, vol. 78, pp. 201-208, 2020.

[58] H. Xiang, Q. Chen, Y. Wu, D. Xu, S. Qi, J. Mei, Q. Li and X. Liu, “Urine calcium oxalate crystallization recognition method based on deep learning,” 2019.

[59] Y. Kazemi and S. A. Mirroshandel , “A novel method for predicting kidney stone type using ensemble learning,” Artif. Intell. Med, vol. 84, pp. 117-126, 2018.

[60] P. Chaitanya and K. Rajesh , “Detection of chronic kidney disease by using artificial neural networks and gravitational search algorithm,” Innovations in Electronics and Communication Engineering, Springer, pp. 441-448, 2019.

[61] M. B. Suresh and M. R. Abhishek , “Kidney stone detection using digital image processing techniques,” Third International Conference on Inventive Research in Computing Applications (ICIRCA), pp. 556-561, 2021.

[62] M. Kobayashi, J. Ishioka , Y. Matsuoka , Y. Fukuda , Y. Kohno , K. Kawano, S. Morimoto, R. Muta , M. Fujiwara, N. Kawamura and T. Okuno, “Computer-aided diagnosis with a convolutional neural network algorithm for automated detection of urinary tract stones on plain x-ray,” BMC Neurol, vol. 21, no. 1, pp. 1-10, 2021.

[63] Y. Y. Liu , Z. H. Huang and K. W. Huang , “Deep learning model for computer-aided diagnosis of urolithiasis detection from kidney–ureter–bladder images,” Bioengineering, vol. 9, no. 2, p. 811, 2022.

[64] L. M. Hsu , S. Wang , L. Walton, T. W. Wang , S. H. Lee and Y. Y. Shih , “3D U-Net Improves Automatic Brain Extraction for Isotropic Rat Brain Magnetic Resonance Imaging Data,” Frontiers in Neuroscience, vol. 15, 2021.

[65] A. Nelson, R. Marco , R. Marco, C. Tiago, O. Davi, B. Ana, F. Eduardo, F. Simas, L. A. Wagner, O. de , S. Leizer, L. Roberto and S. Monteiro, “Low-latency Perception in Off-Road Dynamical Low Visibility Environments,” arXiv preprint, no. arXiv:2012.13014, 2012.

[66] C. C. Liang, P. George, S. Florian and A. Hartwig, “Rethinking atrous convolution for semantic image segmentation.,” arXiv preprint, no. 1706.05587, 2017 Jun 17.

[67] Z. Zhang, X. Wang and C. Jung , “DCSR: Dilated convolutions for single image super-resolution,” IEEE Transactions on Image Processing, vol. 28, no. 4, pp. 1625-35, 2018 Oct 21.

[68] M. Agarwal and R. Mahajan, “Medical image contrast enhancement using range limited weighted histogram equalization,” Procedia Computer Science, vol. 125, pp. 149-56, 2018 Jan 1.

[69] S. Gupta and R. Porwal, “Appropriate contrast enhancement measures for brain and breast cancer images,” International journal of biomedical imaging, vol. Volume 2016 , no. 4710842, pp. 2-4, 2016.

[70] N. M. Sasi and V. K. Jayasree , “Contrast limited adaptive histogram equalization for qualitative enhancement of myocardial perfusion images,” Engineering, vol. 5, no. 10, pp. 326-31, 2013 May.

[71] Y. Mousania and S. Karimi, “A novel improved method of RMSHE-based technique for mammography images enhancement,” in The Selected Papers of The First International Conference on Fundamental Research in Electrical Engineering, Singapore, 2019.

[72] P. Jagatheeswari, S. Kumar and M. Rajaram, “Contrast Stretching Recursively Separated Histogram Equalization for Brightness Preservation and Contrast Enhancement,” in 2009 International Conference on Advances in Computing, Control, and Telecommunication Technologies, Bangalore,India, 2009.

[73] J. P. Antoine , “The continuous wavelet transform in image processing,” CWI Q, vol. 11, no. 4, pp. 323-45, 1998 Dec 1.

[74] J. P. Antoine and R. Murenzi , “Two-dimensional directional wavelets and the scale-angle representation,” Signal processing, vol. 52, no. 3, pp. 259-81, 1996 Aug 1.

[75] O. M. Babiker, Master's thesis on 2D wavelet analysis of the free surface with subsurface turbulence, Norway,oslo: NTNU, 2021.

[76] W. Nash, L. Zheng and N. Birbilis, “Deep learning corrosion detection with confidence.,” npj Mater Degrad, vol. 6, no. 26, 2022.