Computer Aided Diagnosis and Smart Video Validation on IoT Enabled Portable Ultrasound System

K. Divya Krishna

A Thesis Submitted to
Indian Institute of Technology Hyderabad
In Partial Fulfillment of the Requirements for
The Degree of Master of Technology

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Department of Electrical Engineering

January 2016

Declaration

I declare that this written submission represents my ideas in my own words, and where ideas or words of others have been included, I have adequately cited and referenced the original sources. I also declare that I have adhered to all principles of academic honesty and integrity and have not misrepresented or fabricated or falsified any idea/data/fact/source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the Institute and can also evoke penal action from the sources that have thus not been properly cited, or from whom proper permission has not been taken when needed.

K. Divya Krishna

(Signature)

K. Divya Krishna

(K. Divya Krishna)

EE13M0003

(Roll No.)

Approval Sheet

This Thesis entitled Computer Aided Diagnosis and Smart Video Validation on IoT Enabled Portable Ultrasound System by K. Divya Krishna is approved for the degree of Master of Technology from IIT Hyderabad

22/1/16

(Dr. Sumohana Channappayya) Examiner
Department of Electrical Engineering
IITH

(Dr. C. Krishna Mohan) Examiner
Department of Computer Science Engineering
IITH

(Dr. P. Rajalakshmi) Adviser
Department of Electrical Engineering
IITH

Acknowledgements

I sincerely express my gratitude to everyone who supported me throughout the course of this work. I am thankful to my advisor Dr. P. Rajalakshmi for her aspiring guidance, invaluably constructive criticism and friendly advice during the work. I am also thankful to my friends who supported and guided me during my work. Finally, I thank Dr. Abdul Mateen Mohammed, Asian Institute of Gastroentrology, for the discussion related to diagnosis of kidney.

Dedication

I dedicate this work to my family and everyone who is being part of my success. A special feeling of gratitude to my loving dad for his endless love, support and encouragement.

Abstract

This thesis mainly focused on implementation of Computer Aided Diagnosis (CAD) on kidney ultrasound images and ultrasound video validation based on organ information as a part of backend processing in ultrasound imaging system. Ultrasound imaging has been widely used for preliminary diagnosis as it is non-invasive and has good scope for the doctors to analyze many diseases. The most recent research efforts have focused on using low-cost processor cores, such as Field Programmable Gate Arrays (FPGAs) and digital signal processors (DSP). Hence firstly, backend algorithms are implemented on kintex 7 FPGA to meet requirements of present ultrasound system. In later development, System on Chip (SoC) architecture using Zync 7000 platform is used to design an IoT enabled portable ultrasound system with on board CAD and smart ultrasound video validation algorithm.

In rural and remote areas people found struggling to access timely medical treatment. Telemedicine has come to serve rural populations, where time and the cost of travel constraints their access to the best medical care. In the recent survey, it is revealed that most of the people are suffering from kidney diseases. Lack of trained sonographers make ultrasound imaging diagnosis time consuming to detect any abnormality. Hence, we proposed computer aided automatic detection of abnormality in kidney, which is operator independent. In IoT enabled portable ultrasound system, the patient data from remote area is stored securely in the cloud and doctor can access the data from cloud. In this process most of the images/videos that have been uploaded for diagnosis by the semi skilled operators are not very informative. Smart video validation algorithm has been proposed to validate the ultrasound video based on data useful for diagnosis before sending the video to cloud.

CAD for kidney ultrasound images is implemented successfully on Kintex 7 FPGA and Zedboard, also achieved good accuracy. Region Of Interest (ROI) based compression is performed by considering different compression techniques and concluded that applying Integer Wavelet Transform (IWT) predictive coding on ROI and JPEG2K on non-ROI of image gives best PSNR value for different bit rates comparing to all other combination of methods.The proposed video validation algorithm is successfully implemented on Zedboard. It ensures that only valid video data is stored in the cloud, thereby reducing the data in cloud and network traffic. This also ensures reduction in time consumption for transmission of video.

Contents

Chapter 1

Introduction

Ultrasound imaging, also called sonography, involves exposing part of the body to high- frequency sound waves to produce pictures of the inside of the body. Compared to Computed Tomography (CT), Magnetic Resonance Imaging (MRI), Intravenous Urography (IVU) and Angiography (AG), Ultrasonography (US) is more advantageous. US examinations do not use ionizing radiation (as used in x-rays) [1]. Because ultrasound images are captured in real- time, they can show the structure and movement of the bodys internal organs, as well as blood flowing through blood vessels. US is the most widely used imaging technology worldwide and popular due to availability, speed, low cost, patient-friendliness (no radiation). Used for both diagnosis and therapeutic procedures with least invasive compared to other procedures [2]. Ultrasound is a useful way of examining many of the bodys internal organs like liver, gallbladder, spleen, pancreas, kidneys, bladder, uterus, ovaries, and fetus in pregnant patients. The image is created based on the amplitude (strength), frequency and time it takes for the sound signal to return from the area of the patient being examined to the transducer and the type of body structure the sound travels through. Ultrasound image is immediately visible on a video display screen and sonographer will freeze those images which are important for diagnosis by doctor.

In India, particularly in rural and remote areas people found struggling to access timely medical treatment. There is lack of qualified personnel in every sector of health service. Telemedicine has come to serve rural populations, where time and the cost of travel constraints their access to the best medical care [3]. In India 70% of population live in rural areas whereas 75% of qualified consultants practice only in urban centres. Hence there is need to bridge the gap between rural and urban centres. Cloud based wireless portable ultrasound imaging system with preliminary Computer Aided Diagnosis (CAD) serves this purpose. The patient data from the remote ultrasound device is stored on the cloud and can be accessed by authenticated doctors from geographically anywhere as shown in Fig 1.1.

The application of proposed method can be applied in regions where requirement of sonographers is minimal. Due to lack of medical experts there is always delay in the process of getting reports, which is not preferred in emergency. Hence there is necessity for device

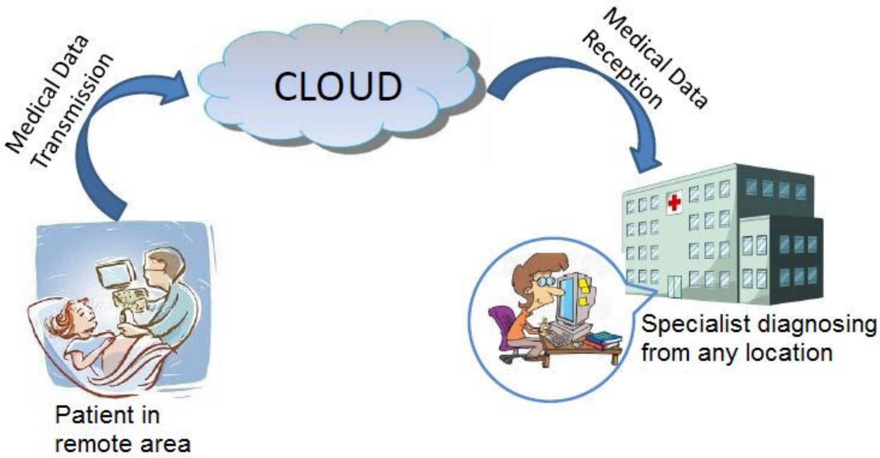


Figure 1.1: Medical data storage in cloud.

to detect the abnormality, that helps the untrained sonographers to take correct decisions. According to Indian council of medical research (ICMR), it is estimated that 77.2 million people are suffering from pre-diabetes, a condition in which the patients have high blood glucose level which is not in diabetes range but having great risk of getting diabetes. Out of 1.27 billion population 65.1 million patients are confirmed diabetes and 17 million diabetes patients are suffering from kidney problems [4]. So there is a need to design a system for preliminary diagnosis of kidney diseases, which is portable and operator independent. The best suited medical imaging modality with these advantages is our proposed IoT enabled ultrasound machine with preliminary CAD. To make the device user independent and detect the abnormality, image processing algorithm are to be ported on the device itself. Normal and abnormal case of kidney with cyst and stone are marked with yellow circle as shown in Fig. 1.2 .

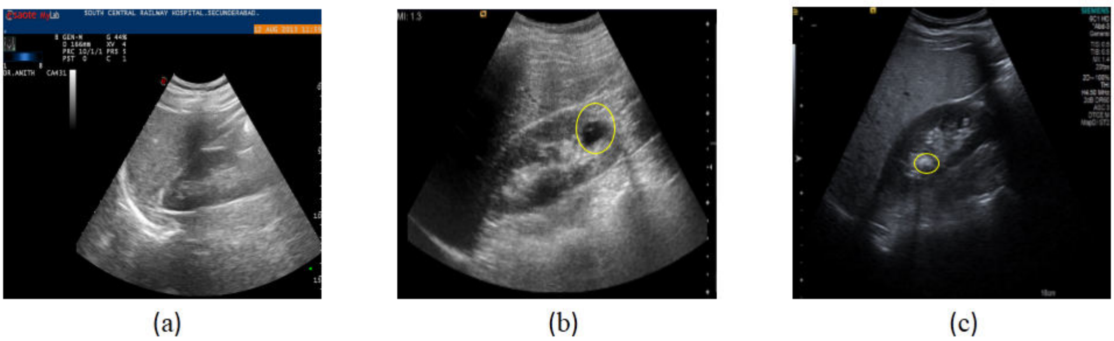
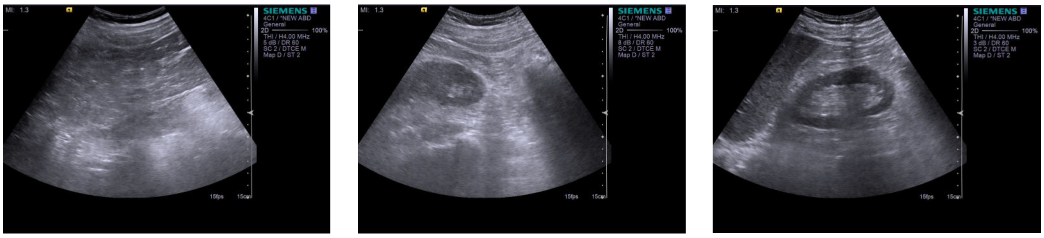


Figure 1.2: (a) Normal, (b) Cyst and (c) Stone

Due to unavailability of radiographers, there is a need for smart connected ultrasound device for remote healthcare. Our interaction with the radiologists show that most of the images/videos that have been uploaded for diagnosis by the semi skilled operators are not very informative due to lack of expertise in using the ultrasound machine itself. Hence, we proposed ultrasound video validation based on the information useful for diagnosis by discarding the data not useful for diagnosis. We classified ultrasound video frames in to three types such as no organ, partial organ and full organ as shown in Fig. 1.3. Ultrasound video frame with any information that can be perceived for diagnosis is classified as *valid frame* for transmission. Ultrasound video frame that does not convey any organ information is considered as *invalid frame* for transmission to cloud.



(a) (b) (c)

Figure 1.3: (a) No Organ (b) Partial organ (c) Full organ.

1.1 Literature Survey

For decades, research has been performed to evolve efficient US system to reduce system hardware and computational complexity while simultaneously providing a flexibility, image quality and programmability comparable to those of conventional US imaging systems. The most recent research efforts have focused on using low-cost processor cores, such as Field-Programmable Gate Arrays (FPGAs) and digital signal processors (DSP) [5][6][7][8]. For miniaturization, cost reduction and long battery life time, the US machines were developed using Application Specific Integrated Circuit (ASIC) chips [9]. In [10] a low cost portable US system is employed on single FPGA board, here they have proposed pseudo-dynamic receive beamforming and look-up table based processing to reduce the hardware complexity. [11] presents a smart US imaging system (SMUS) based on an android-OS smart phone in point-of-care diagnostic applications. Different methods were proposed to miniaturize the US imaging system to reduce the hardware complexity and providing the best quality of image [12][13][14]. Hence, we proposed a novel SoC architecture using Zync 7000 device for an IoT enabled portable US system designed at IIT Hyderabad. In our proposed design, we used a Zynq SoC, which has an FPGA integrated with an ARM cortex A9 dual core processor running at 667.66 MHz interfaced with advanced microprocessor bus architecture. The proposed design gives us flexibility to use the processing capabilities of both FPGA and a processor to efficiently implement all the modules. Parallel processing capabilities of the FPGA can be utilized to implement complex algorithms whereas the ARM processor comes with a number of widely used peripherals and controllers and it makes interfacing the external devices easy. A debian based light weight linux distribution is ported to the SoC which is used to run openCV libraries on the ARM processor. In the proposed US system we incorporated some smart algorithms that includes organ validation, which are not present in commercially available US systems. RAK411 is used as Wi-Fi module, which is interfaced with the device to transmit only valid data to the cloud.

Computer Aided Diagnosis (CAD) has become one of the major tools for medical imaging and diagnostic radiology. Medical diagnostics are highly biased to the person skill and mood. Computer assisted diagnosis can be very handy and will significantly improve the decision making without human intervention. CAD is used to provide a computer output as a second opinion to assist radiologists image interpretation and reducing the time for reading the image. In the 1980s, the concept of automated computer diagnosis was started [15], but these early attempts were not successful. Thus, it became extremely difficult to carry out a computer based analysis in medical images. Therefore, it was not easy to anticipate whether the development of CAD schemes would be successful or failure. From 2000s, there is a gradual increase in research related to CAD diagnosis of breast cancer, lung cancer, colon cancer, prostate cancer, bone metastases, coronary artery disease and congenital heart defect. Today's CAD can provide specificity up to 90 to 100% depending on the application. It may not provide accurate results every time, but can be helpful in providing preliminary diagnosis, in addition to doctors decision which is always important for final consideration. Since ultrasound has many advantages and more frequently used imaging model, we performed CAD on ultrasound images. However the interpretation of ultrasound images of any organ by computer is difficult because of large speckle noise and gray scale image. Eliminating these problems, CAD can be used as a preliminary diagnosis to assist the doctor in ultrasound scanning.

In past, [16] proposed a computer aided decision support system for kidney abnormality detection using first and second order statistical features. Detection of abnormality in kidney ultrasound images based on textural features is discussed in [17]. We proposed FPGA based CAD for normal and abnormal classification of kidney ultrasound images in [18]. Segmenting the kidney from ultrasound images is crucial in computer aided diagnosis of kidney. Many ideas are proposed in the literature to segment kidney from ultrasound images [19]. Ultrasound kidney segmentation based on textural and shape priors is proposed in [20]. [21] proposed an algorithm to automatically detect the region of interest from kidney ultrasound images using textural based classification. Automatic kidney detection using Markov random fields and active contours is proposed in [22], but a manual adjustment is required before processing the image.[23] is based on image appearance and shape variations, aligning of image is necessary which is not feasible when real time detection is considered, [24] is based on machine learning techniques, but requires manual intervention to mark the points on kidney before segmentation. The method we adopted for detecting kidney in an ultrasound images is based on Viola Jones detector [25].

From the previous literature, an overview of medical video communication systems for mhealth has been described in [26]. In [27], authors discuss about video communication framework over mobile WiMAX networks for the wireless transmission of H.264/AVC medical ultrasound video. Authors compared different video compression techniques for ultrasound videos in [28] [29]. In these papers, medical video processing for transmission is proposed assuming, there exists a valid medical video data. However in the scenario of lack of trained sonographers in remote locations, acquiring of valid medical data for further processing is itself a challenge. Before applying any segmentation algorithms [30] or computer aided detection algorithms on medical videos as proposed in [31] and [32], one need to validate the video for further processing.

1.2 Performace Measure

The algorithms developed were first tested on Matlab simulink. Later verilog code for CAD was developed and ported onto Kinetx 7 board. The results obtained are compared with matlab results to validate the algorithms performance. Later experiments for CAD were performed on Zed board that uses Xilinx Zynq 7000 all programmable SOC running with Xillinux operating sytems at 667 MHz clock frequency. Zedboard is a combination of FPGA and ARM Cortex-A9 [33]. The Zedboard has the computational capability to implement ultrasound signal processing algorithms. These features of the board make it ideal for rapid prototyping and proof of concept development.

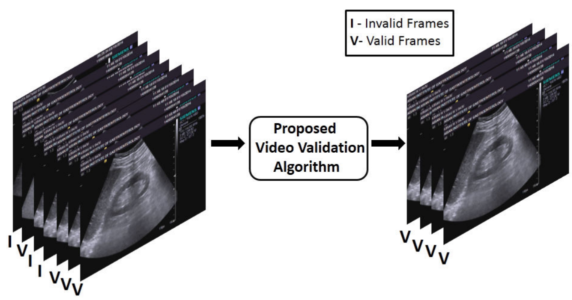


Figure 1.4: Smart Video Validation Output.

To the best of our knowledge this is the first of its kind in the teleradiology applications. The key features of proposed smart ultrasound video validation algorithm implemented on zedboard includes:

- **Smart Scanning:** To make an end device more smart, we proposed smart scanning based on organ validation. From a video, valid frames or diagnosable frames are extracted. Valid frames are those which contains most of the medical data that is used for diagnosis by a doctor. If a video contains all invalid frames without having any organ or partial organ then the device alerts semi-skilled operator to rescan. Thus this smart scanning algorithm at the ultrasound device assist the operator to capture a fully valid ultrasound video. Abstract level representation of video validation performance is as shown in Fig. 1.4.

- **Smart Transmission:** Organ validation algorithm discards the frames based on the useful information present in each frame for diagnosis. The size of video is reduced by discarding those invalid frames, which also reduces the energy required for transmission of video to cloud. For further reduction of network traffic and bandwidth consumption, the video after validation is compressed using H.264 [34]. The compressed data is stored in the cloud using WiFi, which is also integrated on board.

Chapter 2

Preliminary CAD For Kidney Ultrasound Images

Fig. 2.1 shows IoT enabled portable ultrasound system architecture. This provides medical amenities for people in rural areas by transmitting the ultrasound data to cloud, which can be later accessed by doctors from anywhere across the globe for doing diagnosis. Ultrasound imaging system consists of data acquisition, signal conditioning and image reconstruction blocks. CAD, when included in ultrasound device provides a preliminary diagnosis to patient by indicating abnormality in the organ thus providing a faster medication.

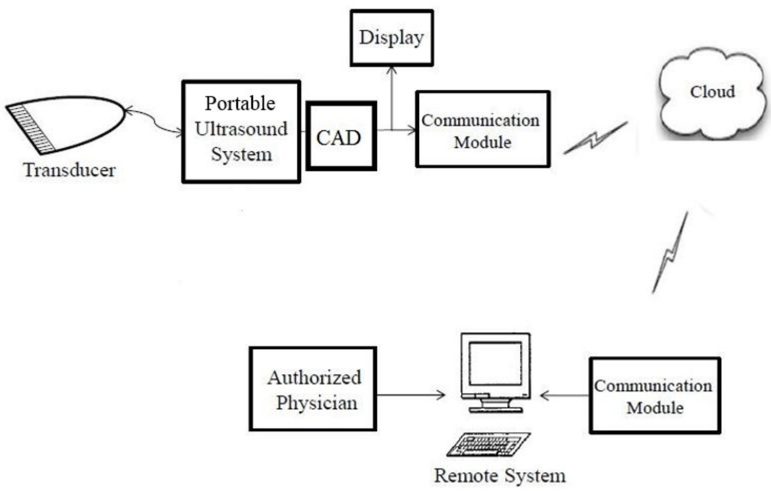


Figure 2.1: IoT enabled ultrasound imaging system with proposed CAD for preliminary diagnosis.

For our analysis, we considered US images consisting of normal, cyst and stone cases in kidney. Fig. 2.2 shows workflow of proposed CAD. After classification all images are stored in cloud and doctors can access those images by providing their own login id and password.

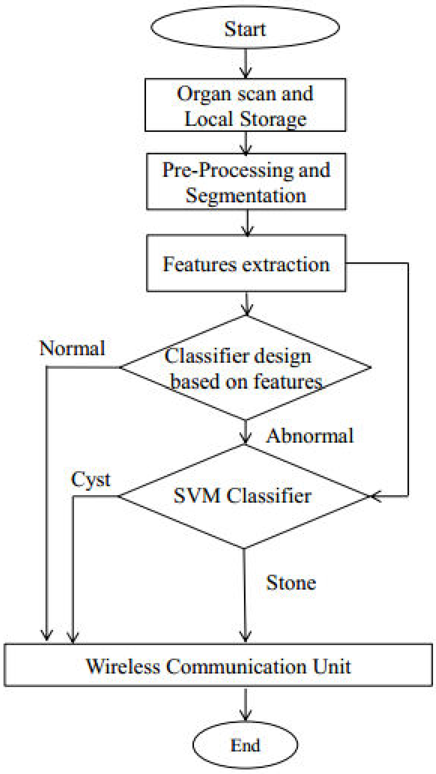


Figure 2.2: Overall work flow of proposed CAD.

2.1 Proposed CAD for Kidney US images

Block diagram for CAD implementation of the classifier for detecting the abnormality of kidney is shown in Fig. 2.3. After acquiring raw image, noise is to be reduced as it can be a major problem for segmentation [35]. Wavelet based pre-processing technique is applied to reduce the noise [36]. From denoised image, kidney region is segmented and features for further processing are extracted, out of which only few features are selected at feature selection to confirm abnormality present in the kidney [18].

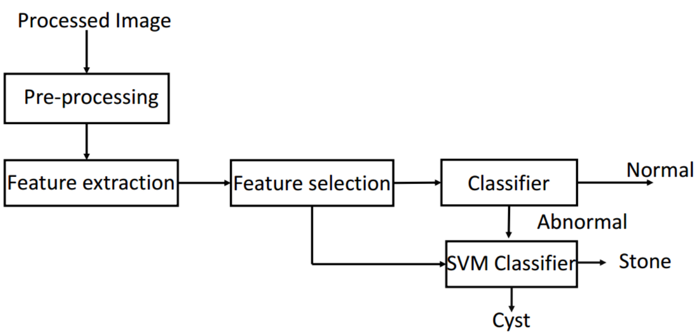


Figure 2.3: System Architecture of proposed CAD.

If any abnormal case is detected, then SVM classifier is used to detect cyst or stone in kidney. Based on the classifier decision, priority of sending patient data can be changed to high in case of emergency. Normal and Abnormal case of kidney with cyst and stone are marked with yellow circle as shown in Fig. 2.4.

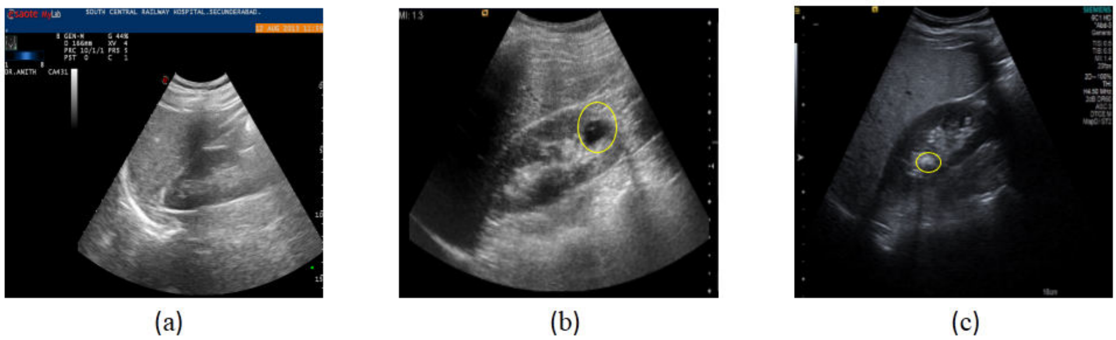


Figure 2.4: (a) Normal kidney, (b) Kidney with Cyst, (c) Kidney with Stone.

2.1.1 Pre-processing

Reduction of speckle noise is one of the operation which increase the quality of ultrasound images by retaining the important features of image. Speckle noise is deterministic and random in ultrasound image which is spatially correlated and multiplicative in nature [37]. Reducing the speckle noise in image will improve the contour of organ in US images and easier to segment. Denoising of speckles is done using threshold wavelet coefficients as shown in Fig. 2.5 by using the fact that in the wavelet domain image is sparse in nature [38].

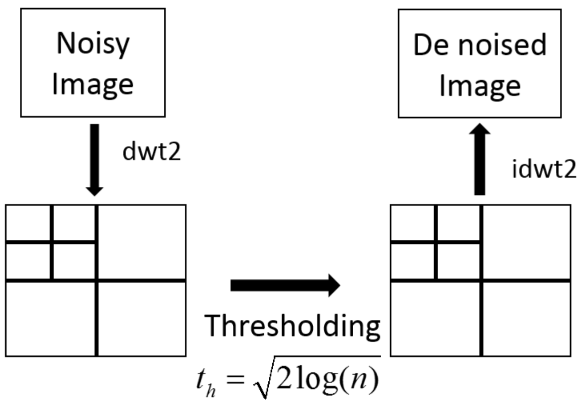
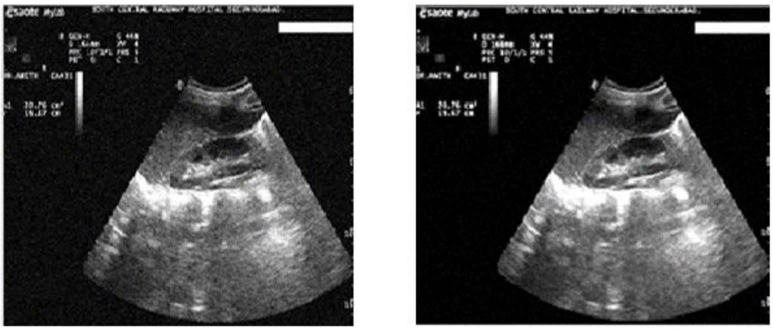


Figure 2.5: Denoising using 2 level DWT: dwt2, idwt2.

Original image has coefficients with large value and when noise is added to it, there will be coefficients having small value. Coefficients having small value are set to zero, since these values belongs to added noise. Global threshold value is selected such that values below it are set to zero and values greater than threshold are set to start from zero. Global threshold value λ is given by:

$$\lambda = \sqrt{2 * log(n)} * s$$

where n is total number of pixels in image given by $N*M$ where N, M are the dimensions of image, s is the noise variance. Discrete Wavelet Transform (DWT) of image is calculated for 3 decomposition levels and threshold is applied to these levels. Inverse Discrete Wavelet Transform (IDWT) is performed on the resultant wavelet coefficients, to obtain the denoised image. Noisy and denoised image obtained after applying global threshold is shown in Fig. 2.6(a) and Fig. 2.6(b) respectively.



(a) (b)

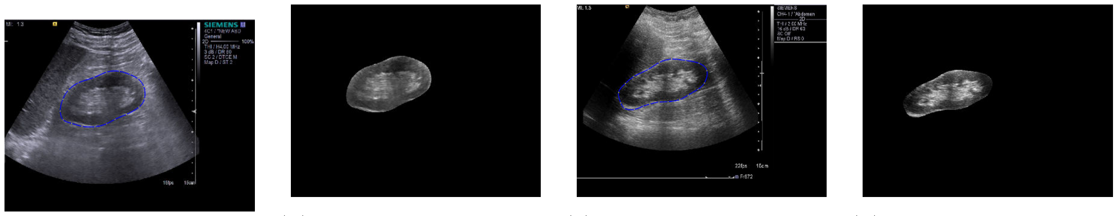
Figure 2.6: (a) Noisy image (b) Denoised image

2.1.2 Finding Region of Interest (ROI)

To do CAD on any organ, firstly we have to extract region of interest for an organ in an image, this is similar to separating the foreground (organ region) from background (region not useful for diagnosis) region from an image. Different techniques are employed from the literature survey and performance of classifier for different techniques are measured.

Manual Segmentation

Due to nonrigid nature of kidney in ultrasound images and scanning artifacts like acoustic shadowing, low contrast and low signal to noise ratio, we manually segmented the region of interest from kidney ultrasound images. For doing manual segmentation, a graphical user interface is created for extracting the region of interest of kidney. We manually marked points along the contour of kidney and cubic spline interpolation is performed between the points to get smooth contour [39]. Normal case of kidney ultrasound image is shown in Fig. 2.7(a) with segmented points marked along contour, segmented region of kidney is shown in Fig. 2.7(b). Fig. 2.7(c) shows ultrasound image with stone in kidney, corresponding segmented region is shown in Fig. 2.7(d).



(a) Normal kidney image. (b) Segmented image of (a). (c) Abnormal kidney image. (d) Segmented image of (c).

Figure 2.7: Manual segmentation of normal and abnormal case of kidney ultrasound images.

Automatic ROI

Region of interest (ROI) is the region in medical image which contains diagnostically important data for examination. Ultrasound images contains more speckle noise and data, which is not useful for diagnosis. Hence we consider only the region that is useful for transmission and diagnosis. In this study, the region of interest is generated based on algorithm proposed by Hafizah et al. [21], where they extracted ROI automatically from ultrasound kidney images based on texture analysis as shown in Fig. 2.8. Window size depends on the seed point selection and center pixel value of the region. The region inside the rectangle box is the region of interest. Applying automatic ROI algorithm on 200 images, we got an accuracy of 95% where the organ lies inside the selected window.

Figure 2.8: ROI of Kidney US image

Automatic ROI: Viola Jones Method

Kidney detection algorithm is based on Viola Jones technique [40]. Block diagram for detection of kidney region is as shown in Fig. 2.9. This algorithm is designed by giving kidney and non kidney ultrasound images as input and is trained to recognize a kidney in image. Later the designed algorithm can be used to detect kidney in any ultrasound image. According to this algorithm all the features of training data are calculated and are stored in a file, later features of new input image are evaluated and are compared with trained features to detect if there is kidney in image or not. This basic components corresponding to this algorithm include haar like features, integral image, Adaboost algorithm and cascade classifier.

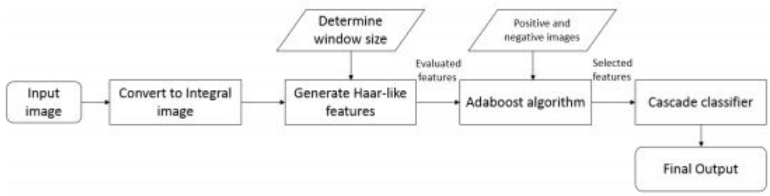


Figure 2.9: Block diagram of Viola Jones Method

Terminology **Image**: Ultrasound B-mode image obtained from siemens ultrasound machine **Window**: To denote detection rectangle where features are evaluated, size of window is less than or equal to size of image.

Reduction of Multiple detections

Multiple detections are a result of overlapping windows having same features as shown in Fig. 2.10. In this paper, we used two stage approach to reduce the multiple detection: In first stage, a window with size less than threshold of size $30X30$ is dropped, threshold is selected based on the observations of several kidney images. In second stage, multiple windows with overlapping region of interest is merged into single window by averaging the coordinates of the window. The overlapping windows in image is determined if the union of two overlapping windows have more than 75% of pixels.

2.1.3 Feature Extraction

Manual segmentation is done on the set of normal and abnormal images in presence of well trained doctor. Features required to characterize kidney are extracted from segmented region. These features can be categorized into three classes: Adaptive features, Histogram features, Haralick features [17]. Some features are adaptive that include size, location and echo texture. The term adaptive means these features vary from person to person and hence cannot be generalized. For example, echo texture varies between thin and muscled persons, thin people have dark echo texture and kidney is visible clearly but muscled people have light echo texture, latitudinal length of normal kidney varies from 8 cm for short people and 12 cm for tall people making it difficult to categorize the kidney. Comparative study is to be done between two kidneys to determine the change in size of kidney. This method also eliminates false negative of abnormality detection in case of diabetics where kidneys are usually larger in size. Intensity histogram features include mean, variance, skewness, kurtosis, energy and entropy.

Feature extraction based on the histogram gives first order statistical features of an image, which is useful for organ recognition and classification. Haralick features also called as

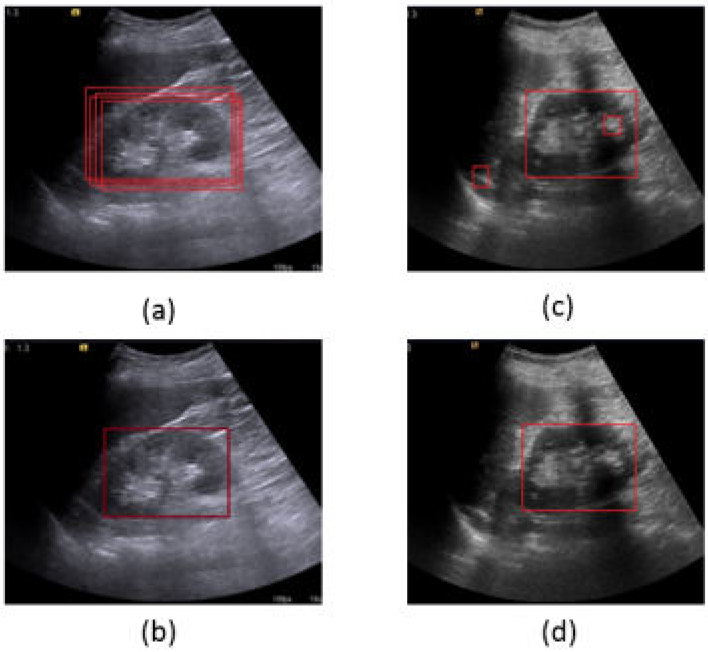


Figure 2.10: (a),(c) Multiple detection of kidney in ultrasound images, red boxes shows the region of interest for kidney. (b) Merging of Multiple detection in (a) to form single ROI. (d) Single ROI detection on (c) using threshold technique.

Gray Level Co-occurence Matrix (GLCM) features are rotational invariant features which include auto correlation, contrast, cluster prominence, correlation, cluster shade, homogeneity, dissimilarity, maximum probability, sum average, sum of squares, sum variance, difference variance, sum entropy, difference entropy, information measure of correlation and inverse difference moment normalized [41]. These are extracted from co-occurrence matrix G of dimension N_g (number of gray level) as given below, each element $P(i, j)$ gives the probability of occurrence of gray level i in the specified spatial relationship with gray level j.

$$G = \begin{pmatrix} P(1,1) & \dots & P(1,N_g) \\ \vdots & \ddots & \vdots \\ P(N_g,1) & \cdots & P(N_g,N_g) \end{pmatrix}$$

$$\mu_x = \sum_{i=1}^{N_g} iP_x(i), \qquad \mu_y = \sum_{j=1}^{N_g} jP_y(j)$$

$$\sigma^2{}_x = \sum_{i=1}^{N_g} (P_x(i) - \mu_x)^2, \quad \sigma^2{}_y = \sum_{j=1}^{N_g} (P_y(j) - \mu_y)^2$$

$$P_x(i) = \sum_{i=1}^{N_g} P(i,j), \qquad P_y(j) = \sum_{j=1}^{N_g} P(i,j),$$

$$P_{x+y}(k) = \sum_{i,j=1}^{N_g} P(i,j)$$

where μ_x, μ_y, σ_x, σ_y are the mean and standard deviation of P_x and P_y. $P_x(i)$, $P_y(j)$ is sum of i^{th} row and j^{th} column respectively.

2.1.4 Feature Selection

Histogram features gives intensity distribution of an image, which includes mean, skewness, kurtosis, variance and entropy. Similarly from Haralick features, 16 features are computed from region of interest. Out of these mean, skewness, kurtosis are selected from histogram features and cluster shade feature is selected from Haralick features [42], further standard deviation of each metric is calculated separately for all images and those with lower deviation are further selected to increase the parameter length to train classifier. We found that homogeneity, maximum probability, correlation, sum average and sum of squares have least deviation compared to the other features. Finally these 9 optimized features are considered for further classification rather than considering all the features. These features are computed in the following way.

$$Mean, \mu = \frac{1}{MN} \sum_{i,j} I(i,j) \tag{2.1}$$

$$Kurtosis = \frac{1}{MN} \sum_{i,j} \frac{(I(i,j) - \mu)^4}{\sigma^4} \tag{2.2}$$

$$Skewness = \frac{1}{MN} \sum_{i,j} \frac{(I(i,j) - \mu)^3}{\sigma^3} \tag{2.3}$$

$$correlation = \sum_{i,j=1}^{N_g} \frac{\{i \times j\} \times P(i,j) - \{\mu_x \times \mu_y\}}{\sigma_x \sigma_y} \tag{2.4}$$

$$sum\ avg = \sum_{i=1}^{2N_g} i P_{x+y}(i) \tag{2.5}$$

$$max\ prob. = \sum_{i,j=1}^{N_g} \max(P(i,j)) \tag{2.6}$$

$$cluster\ shade = \sum_{i,j} (i + j - \mu_x - \mu_y)^3 P(i,j) \tag{2.7}$$

$$homogeneity = \sum_{i,j} \frac{P(i,j)}{1 + |i - j|} \tag{2.8}$$

$$sum\ of\ squares = \sum_{i,j} (i - \mu_x)^2 P(i,j) \tag{2.9}$$

where $I(i,j)$ is the intensity value at i^{th} row and j^{th} column.

2.1.5 Normal/Abnormal classification of Kidney

The classifier is initially loaded with features having desired range as mentioned in Table 2.1. Since length is an adaptive feature it has to be trained every time depending on the patient, for example in the case of diabetics the size of kidney will be more than 12 cm but this case is to be identified as normal. Intensity and Haralick features need not be changed as they are fixed.

Table 2.1: Desired range of features extracted from normal kidney.

Sl.No	Feature	Desired Range
1	Length	8-12 cm
2	Sum Average	0.414-4.182
3	Mean	1.08-1.336
4	Skewness	2.822-7.708
5	Kurtosis	11.06-71.152
6	Correlation	0.971-0.987
7	Cluster Shade	72-243
8	Homogeneity	0.993-0.998
9	Maximum Probability	0.840-0.969
10	Sum of Squares	0.828-10.756

2.1.6 SVM Classifier

Support vector machines (SVMs) are set of supervised learning methods used for classification, regression and outlier's detection. SVM models are designed based on learning algorithms that help in analysing data and recognizing patterns. If classification is easier in a high-dimensional feature space, we would like to build a maximal margin hyperplane in that space as shown in Fig. 2.11. We choose the hyper plane such that the distance from hyper plane to nearest support vectors on both sides is maximized.

The construction depends on inner products, we will have to evaluate inner products in the feature space. This can be computationally intractable, if the dimensions become too large. Using of kernel function that lives in low dimensions, but behaves like an inner product in high dimensions [43] [44]. In training phase, SVM tries to build a model that best separates the features of two different classes. SVM uses an optimization method to identify support vectors S_i, weights W_i, and bias B that are used to classify vectors X according to the following equation:

$$Y = \sum_i W_i K(S_i, X) + B \tag{2.10}$$

where K is a kernel function. Different kernel functions like linear, polynomial, radial basis, multilayer perception are used with SVM classifier [44]. In the case of a linear kernel, K is the dot product. If $Y \geq 0$, then X is classified as a member of the first group, else it is classified as a member of the second group.

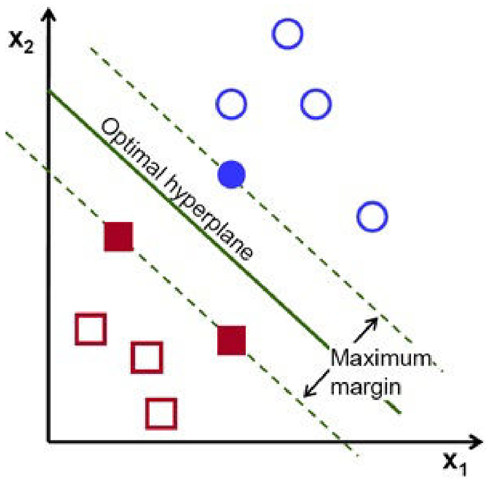


Figure 2.11: SVM Classifier Hyperplane Separation.

Linear SVM

If D be the training data of both normal and abnormal images with a set of n images, it can be represented of the form

$$D = ((x_i, y_i)|x_i \in R^p, y_i \in (-1, 1))_{i=1}^n \tag{2.11}$$

where y_i is either 1 or -1, indicating the class to which x_i belongs. x_i is two dimensional vector, of which x_1 is data set of normal images and x_2 is data set of abnormal images. Linear kernel finds hyperplane to classify the images having $y_i = 1$ and $y_i = -1$.

Polynomial kernel

Polynomial kernel is defined as

$$k(x, y) = (x^T y + c)^d \tag{2.12}$$

where x and y are vectors in the input space, i.e., vectors of features computed from training and test samples. d is set to 3 and c is set to 0 as we computed results for homogeneous kernel.

Radial basis function kernel

Radial basis function kernel is a Gaussian kernel. It has two samples x and x', representing feature vectors of given input image. It is defined as

$$k(x, x') = exp(\frac{||x - x'||^2}{2\sigma^2}) \tag{2.13}$$

$||x - x^{'}||^2$ is the squared Euclidean distance between feature vectors. σ is variance of Gaussian kernel and is set to 1 in our case.

Multi-Layer perceptron

Multi-Layer Perceptron resemble like neural network and also called as hyperbolic tangent kernel. It combines several single layer perceptron's. Each single layer perceptron uses a sigmoid shaped transfer function like the logistic or hyperbolic tangent function:

$$k(x, x^{'}) = tanh(\rho\langle x, x^{'}\rangle, \delta) \tag{2.14}$$

The simplest algorithm for training a multilayer perceptron is the backpropagation algorithm, which is as follows:

1. Select small random weights w.
2. Until halting condition:
   (a) Select a random training example.
   (b) Calculate the output of the hidden layer.
   (c) Calculate the output of the output layer.
   (d) Calculate error for output layer.
   (e) Calculate error for hidden layer.
   (f) Update weights.

2.1.7 Pseudo code for proposed CAD

Steps involved in the CAD analysis are shown in Algorithm 1. Depending on the patient's condition, longitudinal length of normal kidney is considered as threshold. Later 9 features extracted from feature selection block and length of kidney are fed as inputs to classifier block. The intervals are then compared with threshold values and if any feature exceeds the threshold limit then classifier decides that the kidney is abnormal and sends the data to cloud with high priority requesting doctor for immediate diagnosis. After classifying the image has abnormal, the image has to be further classified as cyst and stone using supervised SVM classifier.

2.1.8 Database Acquisition

We acquired images using Siemens Acuson S2000 ultrasound machine with the help of radiologist by taking patients approval on patient consent form. 508 patients including male and female gender are involved in the data collection procedure. The patients are in the age group of 14 to 60 years. Logarithmically compressed kidney images are acquired from the

Algorithm 1 Automatic Kidney Classification.

Initial: Set *Threshold* values

Set *abnormal_count*=0 ;

```
procedure DECISION MAKER(Extracted Features)
    Comment: Calculate mean, skewness, kurtosis, correlation, cluster shade, homogeneity, maximum probability, sum of squares, sum average intervals.
    Calculate length of normal kidney;
    Set length.threshold = length of normal kidney;
    Calculate Data.length_interval;
    if Data.length ≠ length.threshold then
            Decide the kidney is abnormal;
            Send data with high priority;
        Calculate Data.mean_interval;
        Calculate Data.skewness_interval;
        Calculate Data.krutosis_interval;
        Calculate Data.correlation_interval;
        Calculate Data.clustershade_interval;
        Calculate Data.homogeneity_interval;
        Calculate Data.maximumprobability_interval;
        Calculate Data.sumofsquares_interval;
        Calculate Data.sumaverage_interval;
        if Data exceeds Threshold then
            Transmit the data immediately;
            Set abnormal_count=1;
            Apply SVMClassifier;
            if classifier_output = 0 then
                Cyst is present in Kidney ;
            else
                Stone is present in Kidney;
            end if
        else
            Set length.threshold parameter;
            abnormal_count = 0;
        end if
    else
        Decide the patient is normal;
        Transmit the data;
    end if
end procedure
```

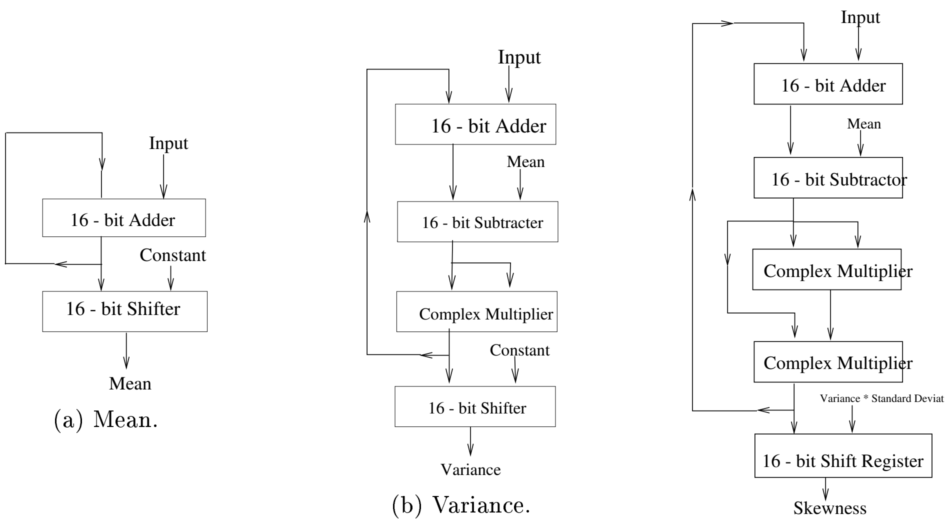
machine. The database consists of 250 normal, 138 stone and 120 cyst kidney images. The condition of kidneys are analyzed and confirmed from the Radiologist, which is used as a ground truth in training and testing the algorithm.

2.2 Hardware Complexity Analysis of proposed CAD algorithm

Hardware complexity of our algorithm for abnormality detection is calculated in terms of gates and transistors [45] [46]. Complexity of each feature extraction and classification are calculated as follows.

2.2.1 Complexity analysis for feature extraction

Table 2.1 shows the selected features for feature extraction. Fig. 2.12(a) to Fig. 2.14 shows the hardware architecture of selected features. Input to the system for calculating mean, skewness, kurtosis is image pixel values and for cluster shade, correlation, maximum probability, sum of average, sum of square, homogeneity is gray level co-occurrence matrix values. Fig. 2.12(a) shows hardware architecture to determine mean. It requires 16 bit adder to add pixel values based on element index and 16 bit shifter to divide result of adder by total number of pixels. Fig. 2.12(b) shows hardware architecture to determine variance. It requires 16 bit adder to add total pixels, 16 bit subtractor for subtracting values of mean from pixel values, Multiplier for squaring the result obtained and 16 bit shifter to divide the final value by total number of pixels. Fig. 2.12(c) shows hardware architecture to determine skewness. It requires 16 bit adder to add total pixel values, 16 bit subtractor for subtracting values of mean from pixel values, two multipliers to compute cube of the result obtained and finally 16 bit shifter for dividing the result by constant. Constant in this case is the product of total number of pixels in image and ^{3} which is product of variance and standard deviation. Fig. 2.13(a) shows hardware architecture to determine Kurtosis. It requires



(a) Mean.

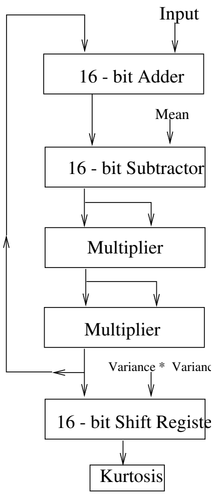
(b) Variance.

(c) Skewness.

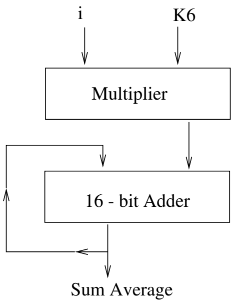
Figure 2.12: Hardware architecture of a) Mean b) Variance c) Skewness.

16 bit adder to add total pixel values, 16 bit subtractor for subtracting values of mean from pixel values, two multipliers to compute 4th power of the result obtained and finally

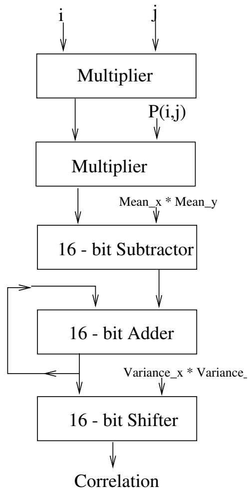
16 bit shifter for dividing by constant which is product of variance2 and total number of pixels in image. Fig. 2.13(b) shows hardware architecture to determine sum average. It requires multiplier to find product of constant K6 defined as mutual probability obtained from GLCM features and index of image, both inputs are given to multiplier and later 16 bit adder is used to add the values obtained from multiplier. Fig. 2.13(c) shows hardware architecture to determine correlation. It requires two multipliers one to multiply indices of image and result is later multiplied with elements in GLCM matrix at other multiplier. The result obtained is subtracted from product of mean x and mean y, which are mean along x and y axis using 16 bit subtractor. One 16 bit adder is used to add value obtained along all the indices and later divided by product of var x and var y which are variance along x and y axis respectively.



(a) Kurtosis.



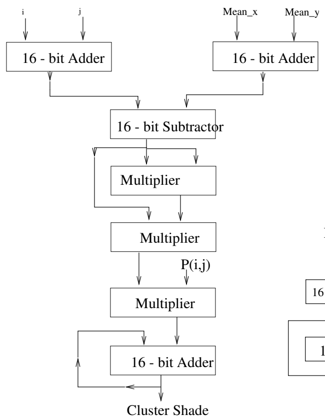
(b) Sum Average.



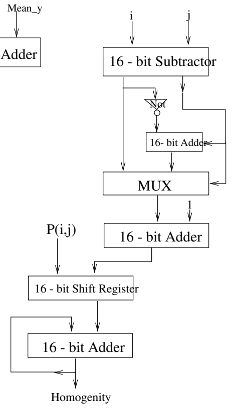
(c) Correlation.

Figure 2.13: Hardware architecture of a) Kurtosis b) Sum Average c) Correlation.

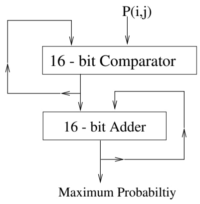
Fig. 2.14(a) shows hardware architecture to determine cluster shade. It requires two 16 bit adders, one to add sum of indices and other to add mean values along x and y axis. The result of these adders is given to 16 bit subtractor. Cube of the result is obtained from two multipliers, other multiplier is used to take the product of result and probability value at that particular index obtained using GLCM matrix. Finally 16 bit adder is used to sum the values from each element. Fig. 2.14(b) shows hardware architecture to determine homogenity. It requires one 16 bit subtractor to find the difference between index of pixel values, one Not gate, one 16 bit adder and one MUX to ensure result of subtractor is always positive. One 16 bit adder is used to add constant one to the result obtained from MUX. The obtained value divides probability value of each element obtained from GLCM matrix with result obtained from adder using 16 bit shifter and later 16 bit adder is used to sum the results from every element. Fig. 2.14(c) shows hardware architecture to determine maximum probability. It requires 16 bit comparator to find maximum probability and 16 bit adder to sum of maximum probability of each element in the GLCM matrix. Fig. 2.14(d) shows hardware architecture to determine sum of squares. It requires 16 bit subtractor to find difference between element index along x axis and mean along x axis. Square of result is obtained using multiplier and result is multiplied with probability obtained from GLCM matrix using other multiplier. Finally 16 bit adder is used to find the sum of every element and get the final result.



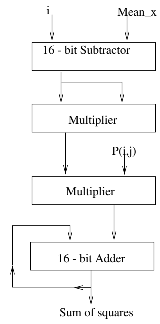
(a) Cluster Shade.



(b) Homogeneity.



(c) Correlation.



(d) Sum of squares.

Figure 2.14: Hardware architecture of a) Cluster Shade b) Homogenity c) Maximum Probability d) Sum of squares.

Table 2.2: Hardware complexity of arithmetic blocks in terms of logic gates.

	NAND	AND	OR	NOR	XOR	NOT
16-Bit Adder	32	8	4	-	16	-
16-Bit Subtractor	-	31	15	-	48	-
16-Bit Multiplier	160	40	8	-	-	16
16-Bit Shifter	-	-	-	-	96	-
Comparator	-	64	16	16	-	32
MUX	4	-	-	-	-	-

Table 2.2 shows the hardware complexity of arithmetic blocks in terms of logic gates. From Table 2.3 we could calculate number of logic gates required to compute the hardware complexity in terms of logic gates. Hence a total of 5120 gates are required to compute the features from kidney images. Table 2.4 gives total number of Complementary Metal Oxide Semiconductor (CMOS) transistors required to implement the feature extraction which are 22,420 in this case.

Table 2.3: Hardware complexity of features calculation in terms of logic gates.

Features	NAND	AND	OR	NOR	XOR	NOT
Mean	32	8	4	-	112	-
Variance	192	79	27	-	160	16
Skewness	352	119	35	-	160	32
Kurtosis	352	119	35	-	160	32
Cluster Shade	576	175	51	-	96	48
Homogeneity	100	56	27	-	192	1
Correlation	352	119	35	-	160	32
Sum of square	352	119	35	-	64	32
Sum of Average	192	48	12	-	16	16
Max Prob	32	72	20	16	16	32
Total Gates	2532	914	281	16	1136	241

Table 2.4: Hardware complexity in terms of CMOS Transistors.

Type of logic gate	No. of CMOS Transistors
NAND	10128
AND	5484
OR	1686
XOR	96
NOR	4544
NOT	482

2.2.2 Complexity analysis for classifier

Classification based on range of features

Fig. 2.15 shows hardware architecture of our proposed classifier. Ten 16 bit comparators are required to compare results obtained from feature selection block with dynamic range as mentioned in Table 2.1 obtained using training data. One 10 input OR gate is used to determine if any one of the input is out of the dynamic range, so as to classify the image as abnormal.

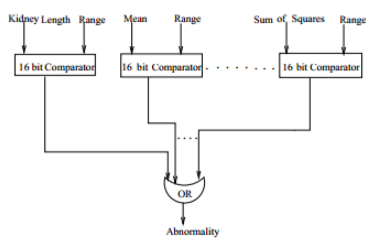


Figure 2.15: Classification based on range of features.

Table 2.5 gives the total number of gates required to calculate the hardware complexity of classifier in terms of logic gates. Total of 1289 gates are required to classify kidney images as normal or abnormal. From Table 2.6 we can say that 6134 CMOS transistors are required to implement our classifier algorithm.

Table 2.5: Hardware complexity of arithmetic blocks in terms of logic gates.

	NAND	AND	OR	NOR	XOR	NOT
16-Bit Comparator	-	640	160	160	-	320
10-Input OR Gate	-	-	9	-	-	-
Total Gates	-	640	-169	160	-	320

Table 2.6: Hardware complexity in terms of CMOS Transistors for classifier.

Type of logic gate	No. of CMOS Transistors
AND	3840
OR	1014
NOR	640
NOT	640

Thus from the above analysis, we see that a total of 6409 gates i.e. 28,554 transistors are required to implement our CAD algorithm which includes feature extraction and classification on hardware. Looking at the resources available in the recent computing platforms, this algorithm can be easily implemented.

SVM Classsifier

Fig. 2.16 shows hardware architecture of SVM classifier. On total 32 16-bit adders, 31 multipliers, and one comparator is required to design a classifier. Where alpha, b are constant values obtained from the classifier design. Input vector contains feature values of the segmented region to be classified.

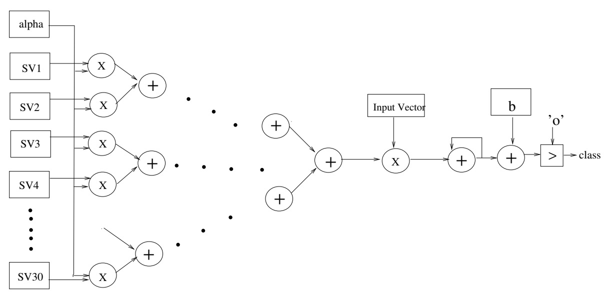


Figure 2.16: Hardware architecture of SVM Classifier

Table 2.7 gives the total number of gates required for hardware complexity analysis of classifier in terms of logic gates. Total of 8,992 gates are required to design a classifier to determine whether the frame of a video is valid or invalid. From Table 2.8 it is observed that 39,840 CMOS transistors are required to implement classifier algorithm. Therefore, total

Table 2.7: Hardware complexity of classifier in terms of logic gates.

Type of logic gate	NAND	AND	OR	NOR	XOR	NOT
Total Gates	5984	1560	392	16	512	528

Table 2.8: Hardware complexity in terms of CMOS Transistors for SVM classifier.

Type of logic gate	No. of CMOS Transistors
NAND	23936
AND	9360
OR	2352
NOR	64
XOR	3072
NOT	1056

of 42,080 CMOS transistors are required to implement feature extraction and classification algorithm on the hardware.

2.3 Results

2.3.1 Implementation on FPGA

The proposed CAD for preliminary diagnosis of kidney has been implemented on kintex-7 FPGA board. For normal and abnormal classification of the image, 150 normal and 150 abnormal image features are analyzed, 100 normal and 100 abnormal images are used for testing the LUT approach. To detect the cyst or stone in an abnormal kidney image, SVM classifier is trained with 75 cyst and 75 stone kidney images and tested with 45 cyst and 63 stone images.

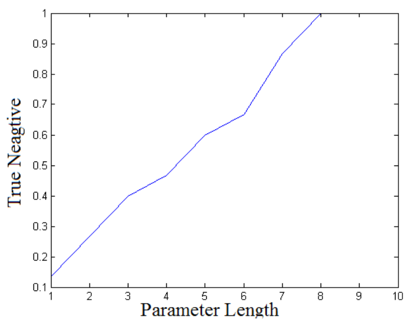


Figure 2.17: True negative for variable parameter length.

Ultrasound images are effected by speckle noise, hence wavelet based denoising technique is used to reduce speckles. Noisy image and denoised image obtained after applying global threshold is shown in Fig. 2.6. Kidney region is segmented from the denoised image as shown in Fig. 2.7. To avoid artifacts from nonkidney images, features are extracted from segmented image. From segmented region, 6 intensity histogram and 16 Haralick features are extracted, out of which only 9 are selected based on standard deviation. Length of a kidney, mean, skewness, kurtosis, correlation, cluster shade, homogeneity, sum of average, sum of squares and maximum probability are selected. If the range of values lie in the desired range as given in Table. 2.1 then kidney is considered to be normal otherwise it is classified as abnormal.

		Patients with Kidney abnormality (as confirmed by sonographer)		
		Condition Positive	Condition Negative	
Kidney Imaging Test Outcome	Test outcome Positive	True Positive (TP) = 100	False Positive (FP) = 0	Positive predictive value = 100 %
	Test outcome Negative	False Negative (FN) = 28	True Negative (TN) = 72	Negative Predictive Value = 72 %
		Sensitivity ≈78.14%	Specificity = 100 %	

(a) Confusion matrix for parameter length = 4 in LUT algorithm.

		Patients with Kidney abnormality (as confirmed by sonographer)		
		Condition Positive	Condition Negative	
Kidney Imaging Test Outcome	Test outcome Positive	True Positive (TP) = 100	False Positive (FP) = 0	Positive predictive value = 100 %
	Test outcome Negative	False Negative (FN) = 0	True Negative (TN) = 100	Negative Predictive Value = 100 %
		Sensitivity = 100 %	Specificity = 100 %	

(b) Confusion matrix for parameter length = 10 in LUT algorithm.

Figure 2.18: Confusion Matrices fro normal or abnormal classification.

Fig. 2.17 indicates the plot of true negatives resulted in LUT approach for different parameter length. Since true negative is constant, from parameter length 8, 9 and 10, we have considered parameter length of 10 which includes longitudinal length of kidney that is variable depending on person so as to detect abnormality efficiently. The accuracy of LUT algorithm increased with increase in parameter length. The classification efficiency of LUT approach for parameter length of 4 in classifying normal and abnormal kidney cases is shown in Fig. 2.18(a), the algorithm resulted with an accuracy of 86%. The LUT approach classified with an accuracy of 100% for parameter length 10 as shown in Fig. 2.18(b). Xilinx Simulink model used to implement LUT approach on FPGA with those 10 features is shown in Fig. 2.19. The algorithm display text N for normal, A for abnormal case as shown in Fig. 2.20 and Fig. 2.21 respectively.

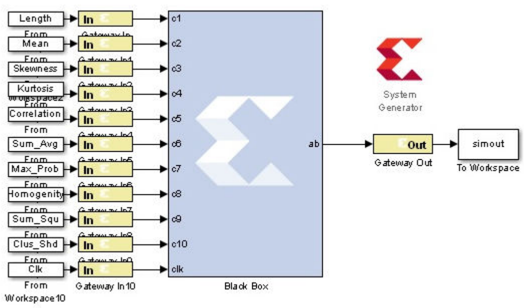


Figure 2.19: Simulink model for Normal/Abnormal classification on FPGA.

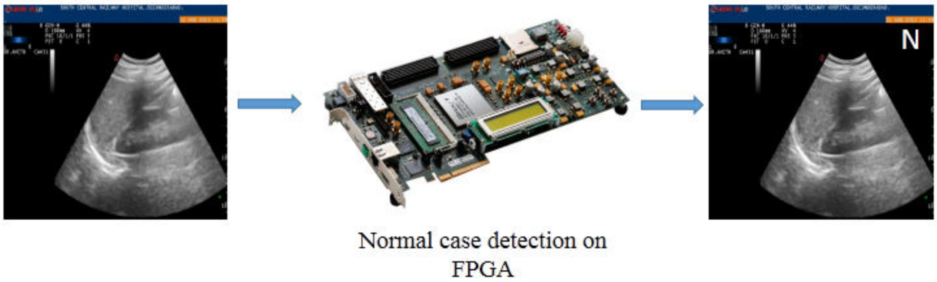


Figure 2.20: Normal case being detected on FPGA.

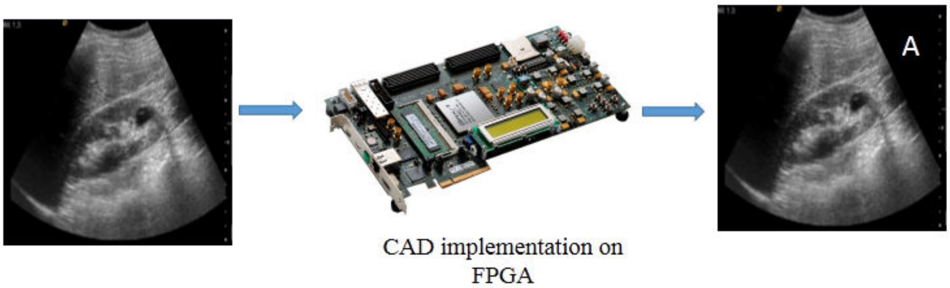


Figure 2.21: Abnormal case being detected on FPGA.

If the kidney is classified as abnormal, then we further classify abnormal kidney as cyst or stone. SVM classifier with same set of features (used for classifying normal and abnormal) are used for detecting the kidney abnormality. SVM classifier is tested with 108 images (45 cyst and 63 stone images). The results are analyzed by plotting the confusion matrix for each SVM kernel. Linear, RBF, polynomial and MLP kernels are used with SVM and confusion matrix for each kernel are shown in Fig. 2.22a, Fig. 2.22b, Fig. 2.23a and Fig. 2.23b respectively.

		Patients with Kidney abnormality (as confirmed by sonographer)		
		Cyst	Stone	
Kidney Imaging Test Outcome	Cyst	36	9	Positive predictive value = 80%
	Stone	27	36	Negative Predictive Value = 57.14%
		Sensitivity =57.14%	Specificity = 80%	

(a) Confusion matrix for linear kernel in abnormality detection.

		Patients with Kidney abnormality (as confirmed by sonographer)		
		Cyst	Stone	
Kidney Imaging Test Outcome	Cyst	18	27	Positive predictive value = 40%
	Stone	18	45	Negative Predictive Value = 71.42%
		Sensitivity =50%	Specificity = 62.5%	

(b) Confusion matrix for radial basis function kernel in abnormality detection.

Figure 2.22: Confusion Matrix for kernels.

		Patients with Kidney abnormality (as confirmed by sonographer)		
		Cyst	Stone	
Kidney Imaging Test Outcome	Cyst	18	27	Positive predictive value = 40%
	Stone	36	27	Negative Predictive Value = 42.86%
		Sensitivity ≈33.5%	Specificity = 80%	

(a) Confusion matrix for polynomial kernel in abnormality detection.

		Patients with Kidney abnormality (as confirmed by sonographer)		
		Cyst	Stone	
Kidney Imaging Test Outcome	Cyst	45	0	Positive predictive value = 100%
	Stone	2	61	Negative Predictive Value = 96.82%
		Sensitivity ≈95.74%	Specificity = 100%	

(b) Confusion matrix for MLP kernel in abnormality detection.

Figure 2.23: Confusion Matrix for kernels.

From Table. 2.9 we can conclude that among linear, RBF, polynomial and MLP kernels, MLP resulted with highest accuracy of 98.14% while polynomial kernel has least accuracy with 41.66%. Xilinx simulink model for SVM classifier to detect cyst or stone in kidney image is shown in Fig. 2.24. If the cyst is present in a kidney then it indicates with text C and for stone with text S on displayed image as shown in Fig. 2.25 and Fig. 2.26 respectively.

Table 2.9: Accuracy using different kernels in SVM classifier.

Kernel	Accuracy(%)
Linear	66.67
RBF	58.33
Polynomial	41.66
MLP	98.14

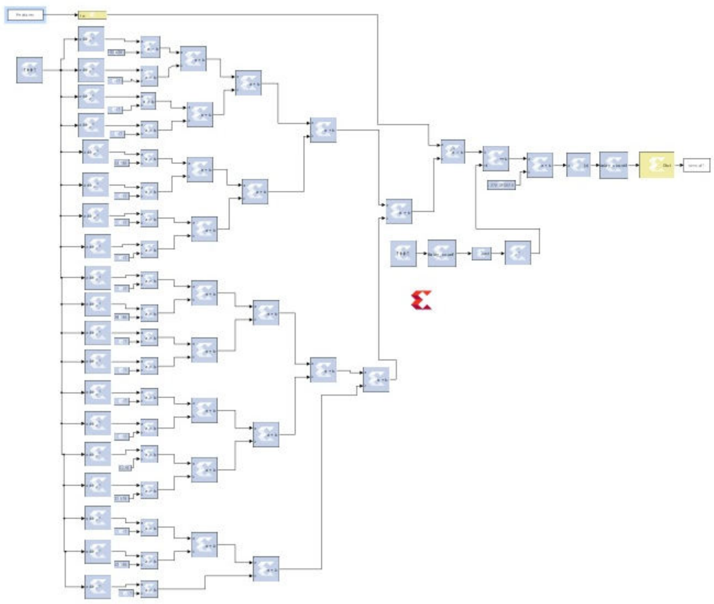


Figure 2.24: Simulink model for cyst/stone classification on FPGA.

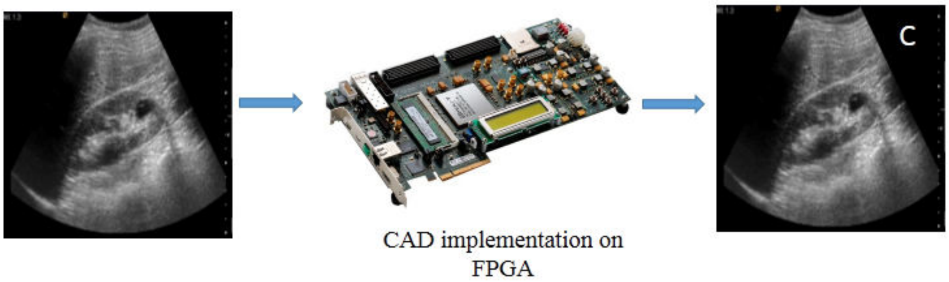


Figure 2.25: Abnormal case with cyst being detected on FPGA.

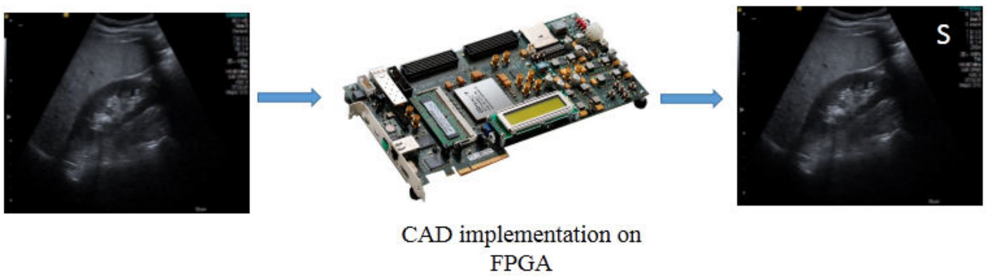


Figure 2.26: Abnormal case with stone being detected on FPGA.

2.3.2 Implementation on zedboard

The proposed CAD for kidney ultrasound images is implemented on zedboard and efficiency of classification is measured for three different methods of finding ROI. As we discussed in section 2.1.2, where ROI is segmented manually and two more methods to find automatically. Table. 2.10 shows the accuracy of classification using three different methods.

Table 2.10: Accuracy of classification using three different methods of finding ROI.

Methods	Normal/Abnormal classification	Cyst/Stone classification
Manual Segmentation	100%	98.14%
Automatic ROI	92.2%	89.6%
Viola Jones	98.65%	96.34%

2.4 Conclusion

In this paper we proposed an algorithm for abnormality detection of kidney ultrasound images for FPGA based IoT enabled ultrasound system. This includes wavelet based noise removal and segmentation of kidney region, feature extraction and selection, and supervised classification. Experimental results show that the designed LUT based classifier, classifies normal and abnormal images without any error. Detected abnormality is further classified as cyst or stone in kidney using SVM classifier. Various kernels are used with SVM classifier out of which multi-layer perceptron kernel gave an accuracy of 98.14% in classifying whether abnormal images corresponds to cyst or stone. Providing such information helps radiologist to suggest immediate precaution and also monitor disease progression. Thus the proposed code is successfully implemented on FPGA using manual segmentation and using automatic ROI implemented on zedboard. Using Viola Jones method for segmentation, we achieved best accuracy that lies closely to manual segmentation.

Chapter 3

ROI Based Compression Techniques

The methodology of the ROI based compression technique for ultrasound images is shown in Figure 3.1. The region which is diagnostically more important is selected from the image and lossless compression technique is applied on this to avoid the loss of data. On the other hand lossy compression is applied on the non-ROI part of the image which is not useful for diagnosis.

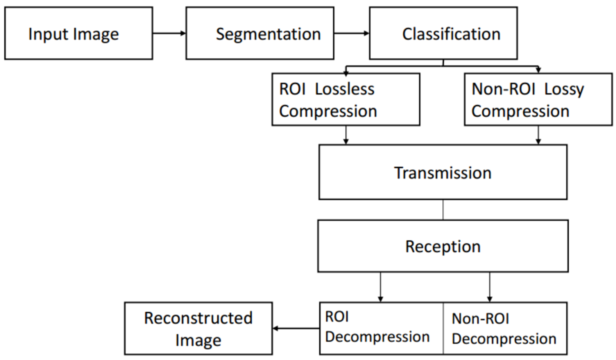


Figure 3.1: Flow Diagram of Proposed ROI Based Compression Technique

Figure 3.2 shows ROI and non-ROI part of kidney ultrasound image. Region Of Interest (ROI) is a part of medical image which contains the data useful for diagnosis and any loss of data in this region will give false report. Non-Region Of Interest (non-ROI) part of the medical image will not contain any useful information for diagnosis and it can also be transmitted with some loss of data. These two regions are compressed and transmitted to the channel with less bandwidth and received at the receiver side. The received image is decompressed and both regions are fused to get the reconstructed image. The final reconstructed image is enhanced using histogram equalization for more quality.

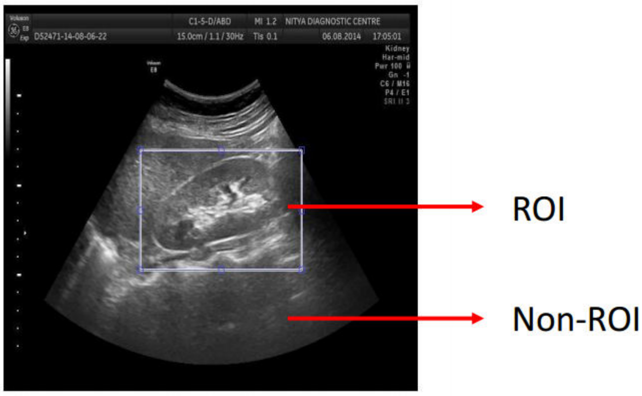


Figure 3.2: ROI and non-ROI of an Ultrasound Kidney Image

3.0.1 Region of Interest (ROI) Classification

Ultrasound imaging is the one with more speckle noise and having distorted edges where normal person can not detect and scan the body without proper training. Different people may have different perception to an image considering physiological and psychological features, and each expert has their own way of diagnosis. Fully automatic detection of organ in ultrasound images will not provide the exact location of organ in the presence of any abnormality and acquired data may not be always useful for diagnosis. In this study, the region of interest is generated based on algorithm proposed by Hafizah et al. [21], where they developed ROI automatically from the ultrasound images based on texture analysis. In this experiment, the input image dimension is fixed according to which the window size is fixed and placement of window exactly around the organ is done by selecting the seed point. As a result, the central area, the renal sinus of the kidney appears more brighter than other part of the image, and in the texture analysis, it is also considered as the common region detected in kidney images. Rectangular ROI is generated after applying some pre-processing methods for reduction of speckle noise and finding the seed point. Block diagram for automatic ROI proposed is shown in Figure 3.3.

3.0.2 ROI: Lossless Compression

After finding ROI from an image, lossless compression is applied for that region. Integer wavelet transform and Predictive coding techniques are used to perform lossless compression on the ROI [47]. This is used to remove spatial redundancy and coding redundancy. When Integer Wavelet Transform (IWT) is applied on the image, it decomposes the image into four subbands using haar filter in the lifting scheme and filter coefficients are

$$h1=[-1\ 9\ 9\ 1]/(16*1.5)$$
$$h2=[0\ 0\ 1\ 1]/(-4*1.5)$$

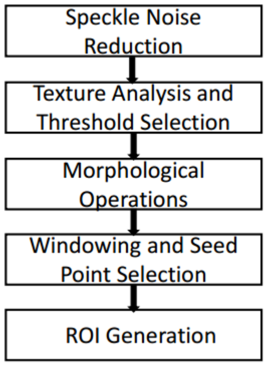


Figure 3.3: Flow Diagram for ROI Generation

The low frequency subband is at highest level and has miniature version of original image with approximate coefficients, which contain maximum information of image that can be called as significant coefficient part. The other subbands contain information about lines, edges and boundaries which typically correspond to high frequency part of image. Human Visual System (HVS) is sensitive to disturbances or losses in low frequency portions of image rather than high frequency component loss. After applying IWT, image is subjected to lossless predictive coding and transfered by highly compressing high frequency components. Block diagram of system is shown in Figure 3.4.

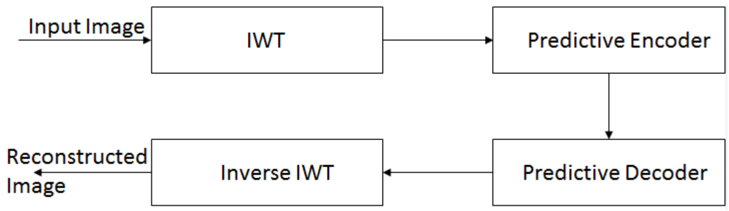


Figure 3.4: IWT and Predictive Coding

Predictive Encoder: In predictive coding [47], a predicted sample is calculated by predicting block by using a second order relation. The second order equation uses previous two samples of the image to calculate present predicted sample, it is rounded off. Then the difference between the predicted sample and present sample is encoded and transmitted. In the process of predictive coding input image is passed through a predictor where it is predicted with its two previous values as shown in Figure 3.5.

$$\widehat{f}(n) = \alpha f(n-1) + \beta f(n-2)$$

Where $\widehat{f}(n) =< f(n-1) >$, $\widehat{f}(n)$ is the rounded output of predictor, $f(n-1)$ and $f(n-2)$ are previous values, α and β are second order predictor coefficients ranging from 0 to 1.

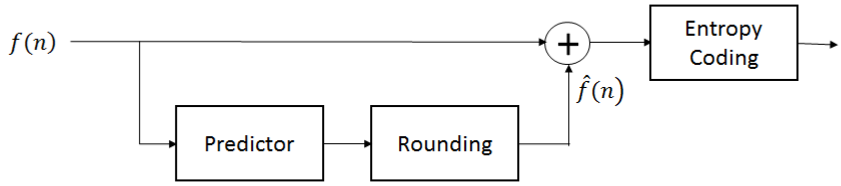


Figure 3.5: Predictive Encoder

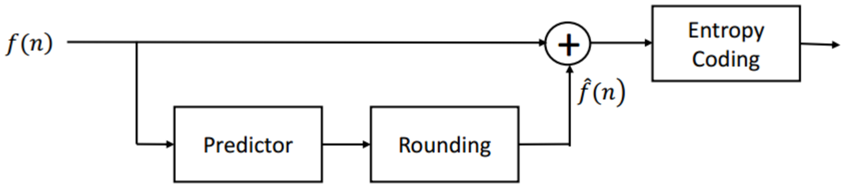


Figure 3.6: Predictive Decoder

Predictive Decoder: Received sample also undergoes predictive detection and the first sample is taken as reference. Then the next sample is predicted from previous samples. Now the received sample and the difference is added to the predicted sample and the pixel is reconstructed as shown in Figure3.6.

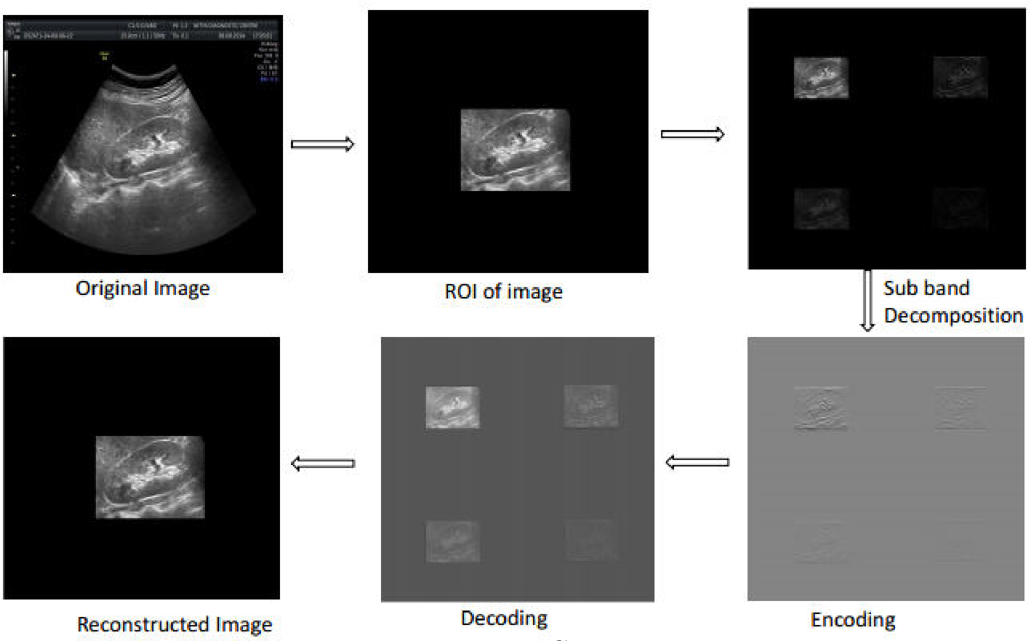


Figure 3.7: Lossless Compression Result

3.0.3 Non-ROI: Lossy Compression

Set Partitioning in Hierarchical Trees (SPIHT) and JPEG2000 techniques are used for compression of non-ROI region. Images processed by using wavelet based techniques are having good visual quality. SPIHT [48] uses hierarchical tree stucture on image which is transformed using wavelet-transform. Usually image information is present in low frequency components of transformed image and decreasing variance is observed as we move from low frequency to high frequency subband in image. Spatial approximate-similarity is observed in connecting sub-bands. As we move down in the pyramid, the coefficients are in better magnitude allignment in same spatial direction. Spatial Oriented Tree (SOT) defines spatial relation on hierarchical pyramid as shown in Figure 3.8. Each node in tree is represented by a pixel and its co-ordinates. Parent node and decendents of the pixel and the pixel form 2*2 adjacent pixel group. Pixels at highest level are tree roots. Except for one node the remaining three nodes have offspring of group size 2*2 pixels and so on forming a hierarchial tree stucture.

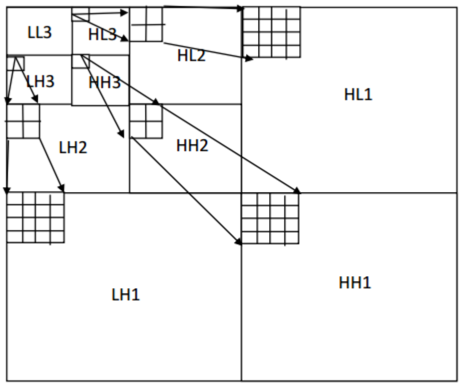


Figure 3.8: Spatial relationship between different levels

JPEG2000 can be used as both lossless and lossy compression at the same time in one encoded bit stream, and is more advanced than JPEG in which DCT is replaced with wavelet transform [49]. This standard defines a new image coding scheme based on wavelet tehnology, in addition to that ROI coding is possible to secure the image. At first, wavelet transform is applied to image followed by quantization and entropy coding then the compressed data is sent to the channel. From receiver side the compressed data decoded and inverse quantized followed by inverse wavelet transform.

3.1 Image quality metrics

The quality of image degrades as it passes through storing, compressing and transmitting processes. Image quality assessment measures the digital image degradation to improve the resultant quality of image. Subjective and objective methods are used for evaluation of image quality. Subjective method is time consuming and expensive where as objective method is simple and independent of user interface. Peak Signal to Noise Ratio, Structural SIMilarity are the metrics considered for comparison of different compression methods [50].

3.1.1 PSNR

Peak signal to noise ratio is the ratio between reference image and distorted image. Higher PSNR value implies distorted image is closer to original image with high image quality. It is calculted as shown in equation

$$PSNR = 10\log_{10}\left(\frac{255^2}{MSE}\right) dB$$

where MSE is the mean sqaure error of an image

3.1.2 SSIM

Structural similarity index is the metric used for measuring the similarity between two images. SSIM is much more better than MSE, PSNR for quantifying the image. It considers image degradation as recognized change in structural information which ranges from -1 to 1. SSIM measured between two windows x and y of size N*N is

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

Where μ_x, μ_y are average of x and y, ${\sigma_x}^2$, ${\sigma_y}^2$ are varaince of x and y, σ_{xy} is covaraince between x and y, $c_1 = (K_1L)^2$, $c_2 = (K_2L)^2$ are two variables to stabilize the division with weak denominator, and L the dynamic range of the pixel-values ($K_1 = K_2 = 0.5$).

3.1.3 Compression Ratio

Compression ratio is used to measure the ability of data compression by considering the ratio of size of compressed image to the size of original image. Relation between bits per pixel (bpp) and compressed image is given by

$CompressionRatio(CR\%) = \frac{8}{bpp}$

Compression ratio decreases with increase in bits per pixel.

3.1.4 Transmission Energy

Energy spent on transmitting the encoded image refering to characteristics of MICA2 motes is given by

$E_{T_x} = (Numberofbytesfortransmission)Xe_{T_x}$

Where e_{T_x} is the energy required to transmit 1 byte of data [51].

3.2 Results and Analysis

Different combination of image compression techniques such as SPIHT, JPEG2000 and IWT followed by predictive coding are applied on ROI and non-ROI parts of image. Figure 3.7 shows the lossless compression using IWT for ROI, Figure 3.9 refers to SPIHT compression for non-ROI, and Figure.3.10 represents reconstructed image combining both the regions. From Table 3.1, it is observed that applying IWT for ROI and JPEG2000 for non-ROI (IWT/JPEG2K) gives high Peak Signal to Noise Ratio (PSNR) at all bit rates compared to any other method. Figure 3.11 shows the performance analysis plot of PSNR Vs Bit Rate for all methods. It is observed that Structural SIMilarity (SSIM) at 0.5 bits per pixel is 0.9974 for IWT/JPEG2K as shown in Table 3.2, which is better than any other techniques. Table 3.3 shows energy spent for transmission of encoded image in millijoules (mJ) for different methods with PSNR value of 52dB, approximately calculated related bits per pixel from Table 3.1. IWT/JPEG2K consumes less energy for transmission of encoded image than any other methods. From the analysis, it is clear that IWT followed by predictive coding for region of interest and JPEG2000 for non region of interest gives better performance compared to other existing techniques.

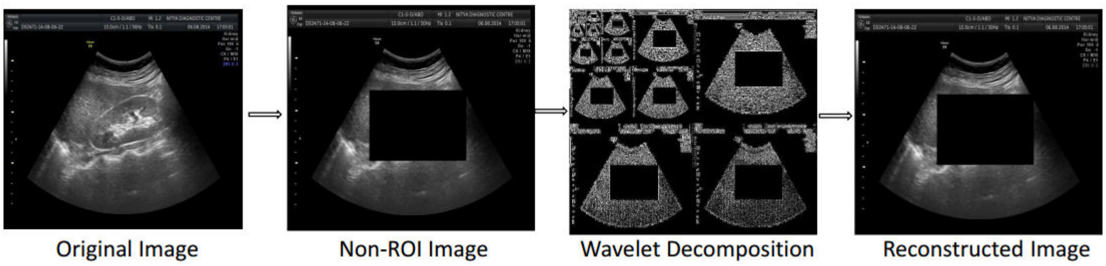


Figure 3.9: Lossy Compression using SPIHT

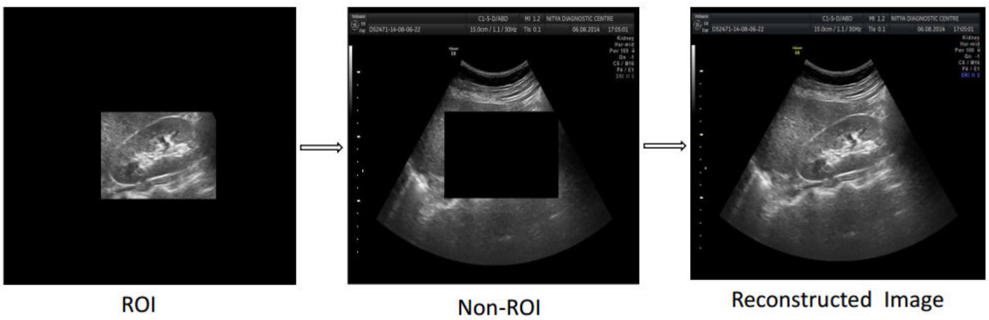


Figure 3.10: Connecting both regions

Table 3.1: Calculation of PSNR for different Bit Rates for each techniques

Bit Rate	CR%	JPEG2K/SPIHT	SPIHT/JPEG2K	IWT/SPIHT	IWT/JPEG2K
0.1	80	37.19	38.14	39.19	41.47
0.2	40	42.65	43.63	44.72	47.42
0.3	26.67	46.72	47.40	48.96	51.63
0.4	20	49.56	49.89	52.05	54.23
0.5	16	51.23	51.40	54.04	55.39
0.6	13.33	52.06	52.21	55.79	56.31
0.7	11.43	52.58	52.72	57.13	58.38
0.8	10	52.98	53.17	58.38	59.96
0.9	8.89	53.39	53.61	59.99	61.58

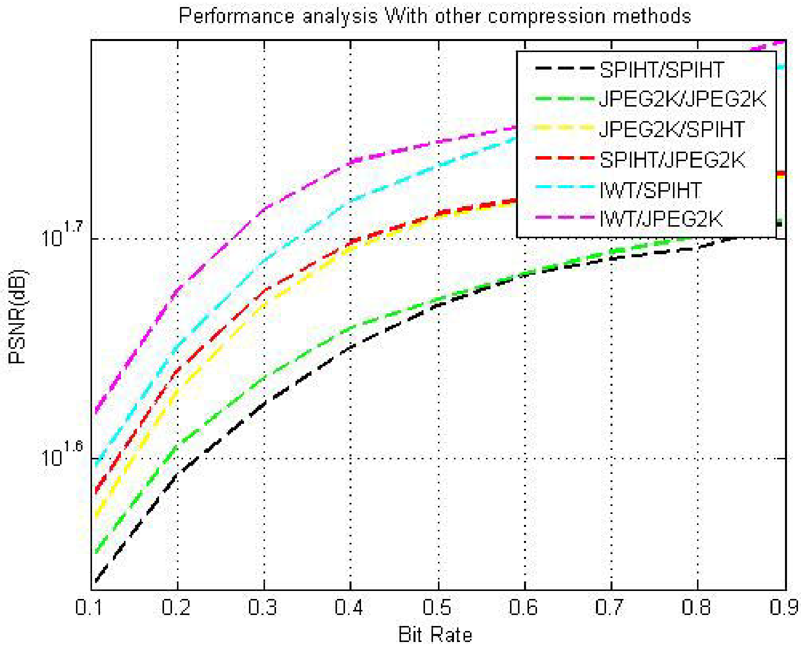


Figure 3.11: Comparision of PSNR for different compression schemes

Table 3.2: SSIM value at 0.5bpp for different methods

Methods	SSIM
SPIHT/SPIHT	0.9930
JPEG2K/JPEG2K	0.9933
JPEG2K/SPIHT	0.9938
SPIHT/JPEG2K	0.9940
IWT/SPIHT	0.9967
IWT/JPEG2K	0.9974

Table 3.3: Energy spent on transmission of image with PSNR=52dB for different methods

Methods	Bit Rate	Energy spent on transmitting (mJ)
JPEG2K/SPIHT	0.7	670.924
SPIHT/JPEG2K	0.6	575.078
IWT/SPIHT	0.4	383.386
IWT/JPEG2K	0.3	287.539

3.3 Conclusion

In this work, medical image compression is done based on region of interest. Region of interest is generated automatically and lossless compression is applied, where as lossy compression is applied for non-ROI. SPIHT, JPEG2K are applied on non-ROI and IWT with predictive coding is applied on ROI for no loss of medical data. From the experimental results, applying IWT predictive coding on ROI and JPEG2K on non-ROI of image gives best PSNR value for different bit rates comparing to all other combination of methods. From SSIM index metric, quality of original and reconstructed image is more similar for this method. Energy required for transmission of compressed image is less for proposed method. Thus, this compression technique can be applied for remote medical health care system efficiently for transmission of images with less bandwidth.

Chapter 4

Smart Ultrasound Video Validation

4.1 Proposed US System Architecture

The architecture of proposed smart connected portable ultrasound system for remote health monitoring is shown in Fig. 4.1. It consists of data acquisition module, signal conditioning and image reconstruction module. Ultrasound video validation and compression is added to the system before transmission of video to cloud. After organ scanning, our proposed method for smart transmission is applied to validate whether the scanned video is useful for diagnosis or not. Transmitting only useful set of video frames can reduce network traffic and improve the battery life time of the portable ultrasound device.

Fig. 4.2 shows our own custom designed ultrasound board combines both the analog and digital processes involved in processing the ultrasound data into a single Printed Circuit Board (PCB) which is powered with Zynq-7000 AP SoC as the central processing system. The board supports 32 channels. It consists of crutial ultrasound components Analog-front-end, Beam former and high voltage pulsar which are interfaced with the FPGA of Zynq-7000 SoC. The board requires a standard 5V@8A supply for supporting the Zynq board and on board peripherals and a supply of 15V-70V, 500mA for exciting the transducer. The software architecture implemented in the device is flexible and is easy to add new features and algorithms using Xilinx tools.

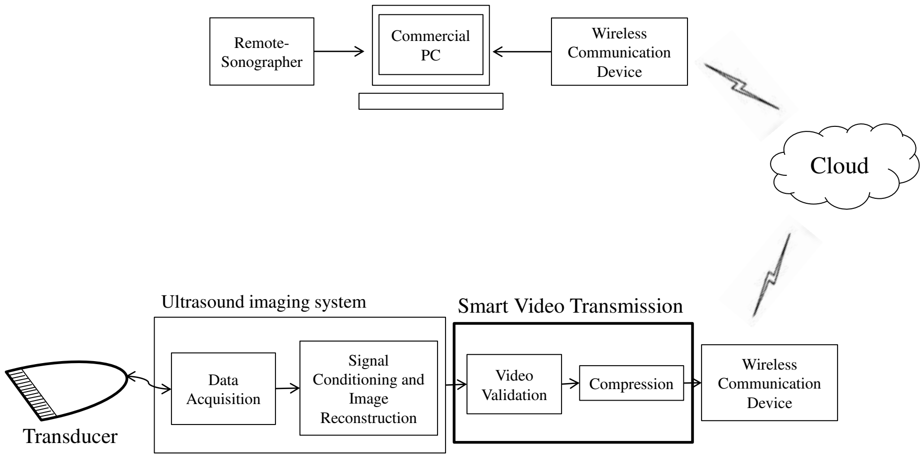


Figure 4.1: Proposed architecture of Smart Connected Ultrasound Imaging System (SCUIS) for remote health monitoring.

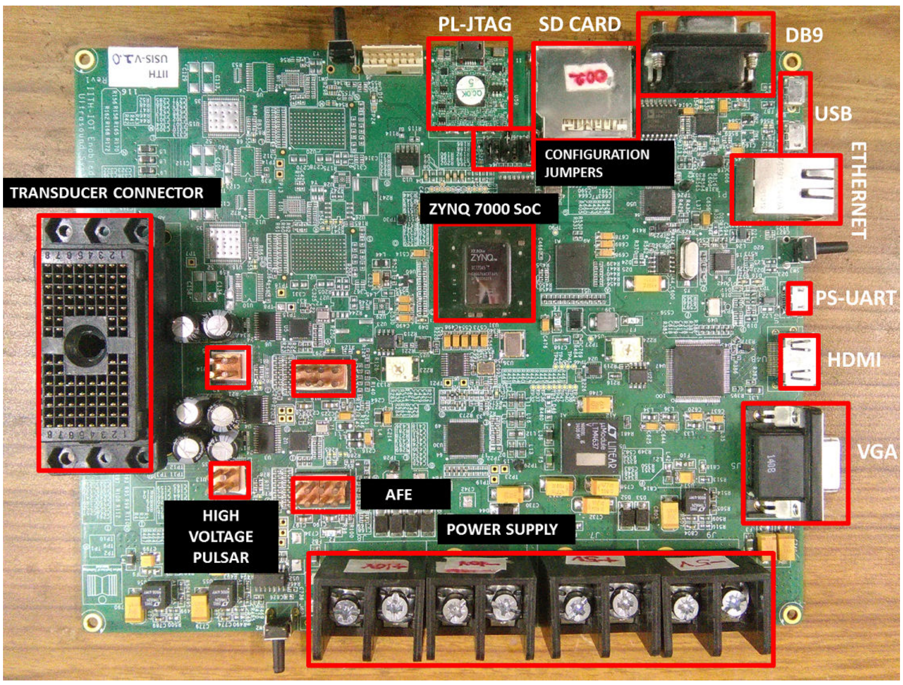


Figure 4.2: Custom designed Ultrasound Board.

4.1.1 External Interface-RAK411 As Wi-Fi Module

RAK411 is used as Wi-Fi module as shown in Fig. 4.3, it is available in QFN package and can be embedded onto the board directly. RAK411 as Wi-Fi module interfaced with the device and features of this module support IEEE 802.11 b/g/n protocols, four wire SPI interface with a maximum clock frequency support of 16 MHz, a maximum throughput of 2Mbps using SPI protocol and printed ceramic antenna on board.

Figure 4.3: RAK411 as Wi-Fi Module.

Although the speed of the communications module is 2Mbps, the actual speed that can be achieved depends on the link speed of the network the device is connected to. The average speeds that the device could achieve are as follows.

- 400kbps when connected to Wi-Fi router connected to a LAN.
- 296kbps when connected to a mobile phone acting as Wi-Fi access point sharing the mobile data connection.
- 154kbps when the device is connected to the network using a 2G internet dongle.
- 833kbps when the device is connected to the network using a wired internet connection.

4.2 Ultrasound Video Validation for Smart Transmission

In back end processing, after reconstruction of an image or video from the US system the smart validation algorithm runs on the processor. Workflow for smart ultrasound video transmission is shown in Fig. 4.4. Captured video is converted into frames depending on the frame rate of the video, where N_F is the total number of frames generated in the video. We classified ultrasound video frames in to three types such as no organ, partial organ and full organ as shown in Fig. 4.5. Ultrasound video frame with any information that can be perceived for diagnosis is classified as *valid frame* for transmission. Ultrasound video frame that does not convey any organ information is considered as *invalid frame* for transmission to cloud. Fig. 4.5(a) shows a frame with no organ i.e., invalid frame, frames with partial organ and full organ are treated as valid frame as shown in Fig. 4.5(b) and Fig. 4.5(c) respectively. V_f gives valid frame count and F_count gives frame count, initially these are assigned to zero. Features are extracted from each frame and classifier is trained with feature values of three classes of classification. Frames that are not useful for diagnosis are discarded, while valid frames that are useful for diagnosis are stored in buffer and V_f count is increased. F_count increased by one for every frame. The process of feature extraction, classification continues until $F\_count = N\_f$. If $V\_f = 0$ i.e., there are no valid frames in a video, then it alerts the sonographer to rescan. If the condition of $V\_f = 0$ is false then video contains only valid frames and transmitted to cloud after appropriate compression.

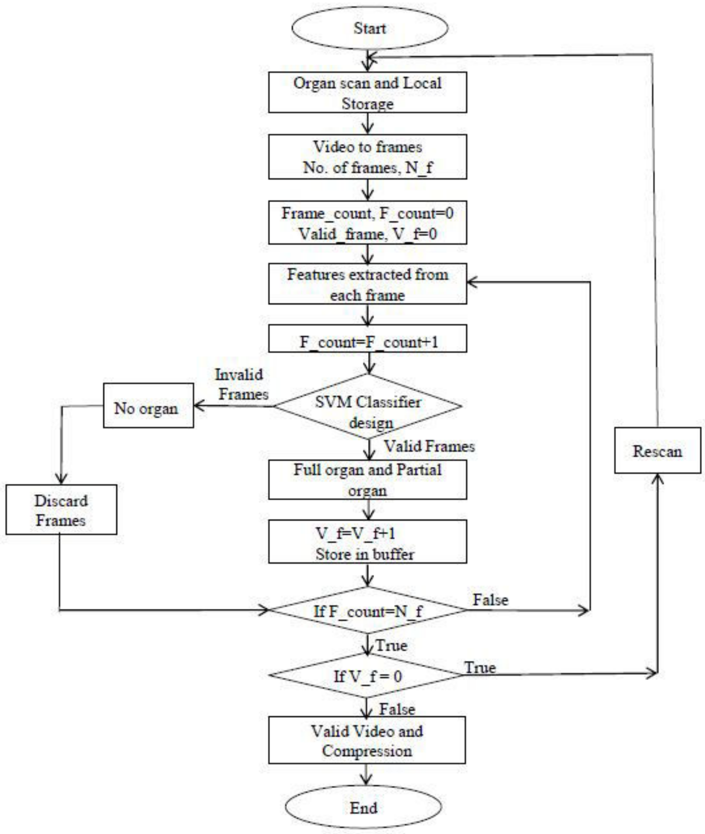
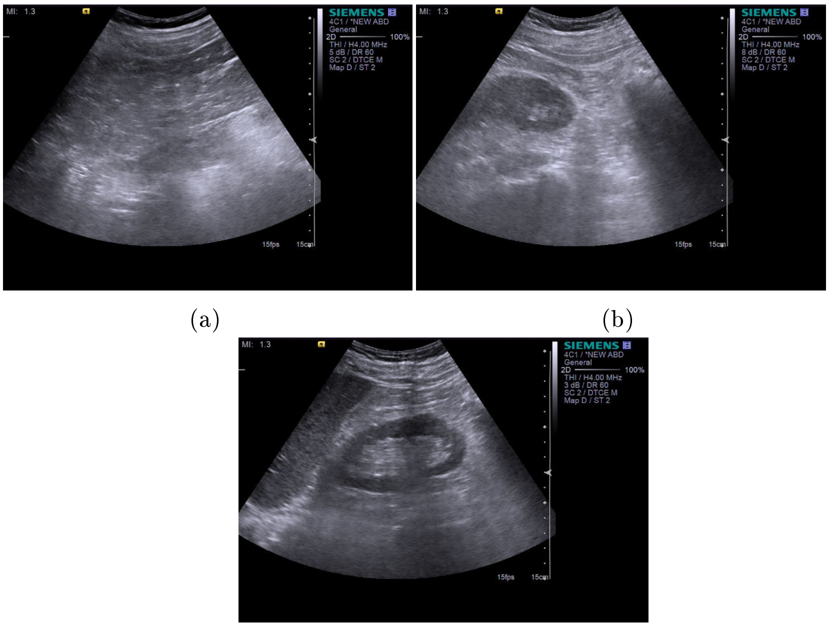


Figure 4.4: Overall work flow of Proposed Ultrasound Video Validation followed by Compression for Smart Transmission.

4.2.1 Implementation Proposed Algorithm in Detail

The block diagram of proposed organ validation algorithm is shown in Fig. 4.6. Ultrasound video stream is divided into set of frames and region of interest is found at each frame and features are extracted from that region. After extraction of features, only selected features are considered for further classification. Classifier decides the current frame as no organ or full organ or partial organ frame. Partial and full organ frames are considered as valid frames and stored in buffer, where as no organ frames are discarded.



(a) (b)

(c)

Figure 4.5: (a) No Organ (b) Partial organ (c) Full organ.

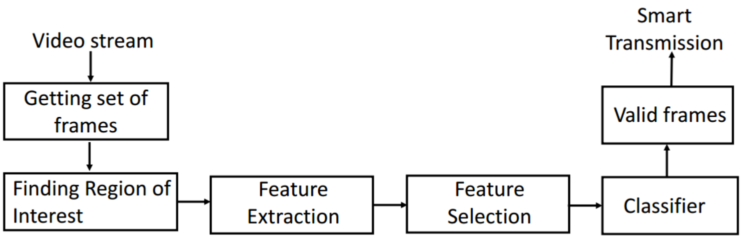


Figure 4.6: Block diagram for video validation algorithm.

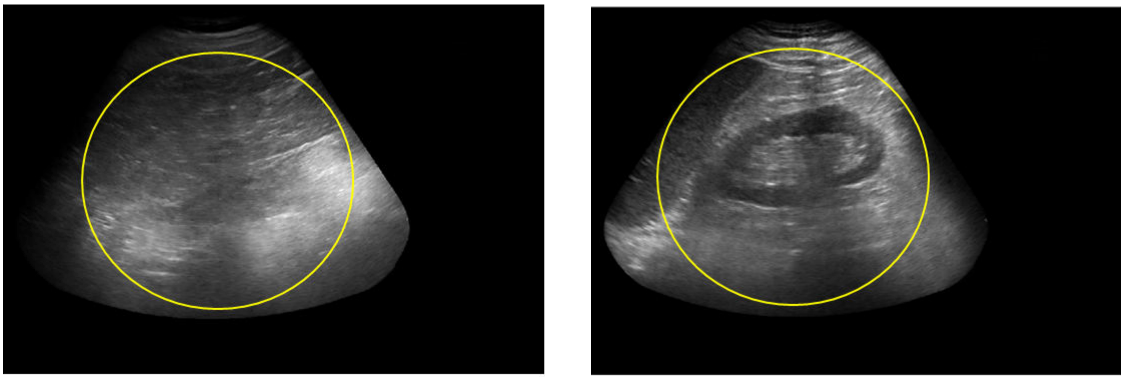
4.2.2 Database Acquisition

We acquired videos using Siemens Acuson S2000 ultrasound machine with the help of radiologist by taking patients approval on patient consent form. 80 patients including male and female gender are involved in the data collection procedure. The database consists of 20 normal, 20 abnormal kidney US videos. The condition of kidneys are analyzed and confirmed from the Radiologist, which is used as a ground truth in training and testing the algorithm. Out of 40 videos, 10 normal and 10 abnormal cases of kidney videos are used to train the classifier. 350 valid and 350 invalid frames are used to train the classifier by combining both normal and abnormal cases. Remaining 20 videos are used to test the classifier. In the testing phase, 300 valid and 300 invalid frames are used.

4.2.3 Finding Region of Interest

Frames are extracted from the video and considered as individual images. Region of interest (ROI) is the region in medical image which contains diagnostically important data for examination. In ultrasound images the imaging plane is in cone shaped where data required for diagnosis lies inside it. Circle shape with appropriate radius is chosen to enclose the ROI such that it covers the maximum area at bottom of cone where medical data is present.

Before finding ROI the image is normalized and median filter is used to remove background noise and only cone shaped imaging plane is obtained. The radius and center of circle is selected based on the resolution of the image. The size of image is $1024X712$, 1024 columns and 712 rows. Image is scanned along X-axis to detect first and last nonzero pixel values. Let the coordinates of first and last nonzero pixel values along X-axis be (x_i, y) and (x_j, y), where $i, j \in [1, 1024]$. i, j take values from 1 to 1024, where i starts from minimum value 1 and j starts with 1024. The midpoint of first and last nonzero pixel value is the required x-coordinate of center of circle. Similarly, by scanning along Y-axis the y-coordinate of center of circle is determined. Radius of length 250 pixels is considered for ROI, that exactly fits to the acquired image. ROI for valid and invalid frame is as shown in Fig. 4.7



(a) Invalid Frame　　(b) Valid Frame

Figure 4.7: ROI for valid and invalid frame.

4.2.4 Feature Extraction

Texture is one of the important characteristic to determine the presence of any organ or finding the region of interest in an image. In most of the medical images like CT scan, MRI, and ultrasound they differ by image texture. There are many metrics in image processing to quantify the perceived texture of an image [52][53]. Features are extracted to represent an image in reduced and compact form in order to speed up the decision making process for further classification. Intensity histogram features, Gray Level Co-occurrence Matrix (GLCM) features, Gray Level Run Length Matrix (GLRLM) are used in our algorithm for classifying each frame as valid or invalid.

Intensity Histogram Features

These are first order statistical features used mainly for organ recognition and classification. It includes mean, variance, skewness, kurtosis, energy and entropy.

GLCM

GLCM features includes auto correlation, correlation, contrast, cluster shade, cluster prominence, dissimilarity, entropy, energy, homogeneity, maximum probability, sum of squares, sum variance, sum average, sum entropy, difference entropy, difference variance, information measure of correlation, inverse moment normalized, and inverse difference normalized. Harlick features are also called as GLCM features because, features are extracted from gray level co-occurrence matrix G with dimension of $N_g * N_g$ where number of gray levels are N_g [41] [42]. Gray level co-occurrence matrix gives the spatial relationship between the pixels i.e, probability of occurrence of each gray level value in specified spatial relation with other gray level P(i,j) as shown below.

$$G = \begin{pmatrix} P(1,1) & \dots & P(1,N_g) \\ \vdots & \ddots & \vdots \\ P(N_g,1) & \cdots & P(N_g,N_g) \end{pmatrix}$$

$$\mu_x = \sum_{i=1}^{N_g} iP_x(i), \qquad \mu_y = \sum_{j=1}^{N_g} jP_y(j)$$

$$\sigma^2{}_x = \sum_{i=1}^{N_g} (P_x(i) - \mu_x)^2, \quad \sigma^2{}_y = \sum_{j=1}^{N_g} (P_y(j) - \mu_y)^2$$

$$P_x(i) = \sum_{i=1}^{N_g} P(i,j), \qquad P_y(j) = \sum_{j=1}^{N_g} P(i,j),$$

$$P_{x+y}(k) = \sum_{i,j=1}^{N_g} P(i,j)$$

where μ_x, μ_y, σ_x, σ_y are the mean and standard deviation of P_x and P_y. $P_x(i)$, $P_y(j)$ is sum of i^{th} row and j^{th} column respectively.

GLRLM

The Gray Level Run Length Matrix (GLRLM) method [54] extracts higher order statistical texture features. Let N_g be the number of gray levels, R be the longest run, and N be the number of pixels in the image. The GLRLM is a two dimensional matrix of (N_g*R) elements in which each element $P(k,l|\theta)$ gives the total number of occurrences of runs having length k of gray level l, in a given direction. Seven statistical texture features to be extracted from the Gray Level Run Length matrices. These are: Short Runs Emphasis (SRE), Long Runs Emphasis (LRE), Gray Level Non-uniformity (GLN), Run Length Non-uniformity (RLN), Run Percentage (RP), Low Gray Level Runs Emphasis (LGRE), and High Gray Level Runs

Emphasis (HGRE).

4.2.5 Feature Selection

After extraction of features from region of interest, we get 6 histogram features, 19 GLCM features, and 7 GLRLM features. From these 32 features we need to optimize some features to reduce the feature vector size.

Genetic Algorithm (GA) Based Method

Genetic algorithms [55] are search algorithms, which are based on Charles Darwin's principle of 'survival of the fittest'. Initially a set of candidate solutions are created and their corresponding fitness values are calculated. This set of solutions is referred to as a population and each solution as an individual. The individuals with the best fitness values are combined randomly to produce offspring's which make up the next population. To do so, individual are selected and undergo cross-over and also are subject to random mutations. This process is repeated again and again and many generations are produced that should create better and better solutions. For feature selection, the individuals are subsets of predictors that are encoded as binary; a feature is either included or not in the subset. The fitness values are some measure of model performance, such as the RMSE or classification accuracy. One issue with using GAs for feature selection is that the optimization process can be very aggressive and their is potential for the GA to overfit to the predictors.

In addition to 6 histogram features and 19 GLCM features, 7 GLRLM features are also calculated and GA is applied. From the algorithm, out of 32 features 10 features namely skewness, kurtosis, correlation, cluster shade, cluster prominence, homogeneity, sum of square, gray level non-uniformity, run length non-uniformity, and high gray level run emphasis have been chosen. These 10 features are computed as shown below, which are further used for classification.

$$Skewness = \frac{1}{MN}\sum_{i,j}\frac{(I(i,j)-\mu)^3}{\sigma^3} \tag{4.1}$$

$$Kurtosis = \frac{1}{MN}\sum_{i,j}\frac{(I(i,j)-\mu)^4}{\sigma^4} \tag{4.2}$$

$$Correlation = \sum_{i,j=1}^{N_g}\frac{\{i \times j\} \times P(i,j) - \{\mu_x \times \mu_y\}}{\sigma_x \sigma_y} \tag{4.3}$$

$$Cluster\ shade = \sum_{i,j}(i+j-\mu_x-\mu_y)^3 P(i,j) \tag{4.4}$$

$$Homogeneity = \sum_{i,j}\frac{P(i,j)}{1+|i-j|} \tag{4.5}$$

$$Cluster\ prominence = \sum_{i,j} (i + j - \mu_x - \mu_y)^4 P(i,j) \tag{4.6}$$

$$Sum\ of\ squares = \sum_{i,j} (i - \mu_x)^2 P(i,j) \tag{4.7}$$

$$GLN = \frac{\sum\limits_{k=1}^{N_g} (\sum\limits_{l=1}^{R} p(k,l|\theta))^2}{\sum\limits_{k=1}^{N_g} \sum\limits_{l=1}^{R} p(k,l|\theta)} \tag{4.8}$$

$$RLN = \frac{\sum\limits_{l=1}^{R} (\sum\limits_{k=1}^{N_g} p(k,l|\theta))^2}{\sum\limits_{l=1}^{R} \sum\limits_{k=1}^{N_g} p(k,l|\theta)} \tag{4.9}$$

$$HGRE = \frac{\sum\limits_{k=1}^{N_g} \sum\limits_{l=1}^{R} k^2 p(k,l|\theta)}{\sum\limits_{k=1}^{N_g} \sum\limits_{l=1}^{R} p(k,l|\theta)} \tag{4.10}$$

4.2.6 Classification Methods and Performance Measure

Different classification methods have been employed to compare the performance of each classifier. K-nearest neighbor (KNN) [56] [57], Decision Tree classifier [58], Naive Bayes classifier (NB) [59] and Support Vector Machine (SVM) [60] classifiers are used for analysis. Selected features are extracted from the region of interest for each frame in video and fed to the classifier. Each classifier is trained with valid and invalid classes of images. In the field of machine learning, confusion matrix is used to measure the performance of classifier in terms of number of true positives, true negatives, false positives and false negatives. A confusion matrix contains information about actual and predicted classifications done by a classification system. Performance of such systems is commonly evaluated using *Accuracy* .

Table 4.1: Confusion matrix for detection of valid/invalid frames.

		Predicted class	
		Valid	Invalid
	Valid	300	0
Actual class	Invalid	10	290

Fig. 4.8 shows performance of different classifiers trained with features selected from genetic algorithm. Performance is measured using accuracy of classifier by varying the size of training data set. Training data size is varied from 80 to 700 images with equal number of valid and invalid images. It is observed that SVM classifier has highest accuracy compared

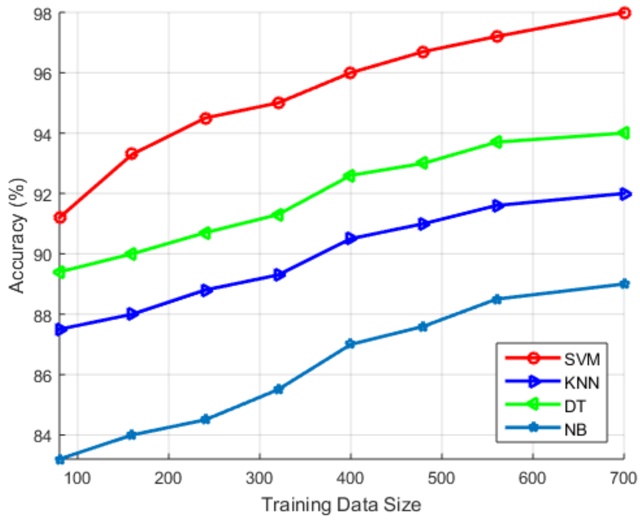


Figure 4.8: Performance of classifier Vs training data set size using genetic algorithm for feature selection

to all other classifiers and confusion matrix for SVM classifier is as shown in Table 4.1. 300 valid frames are detected as valid and out of 300 invalid frames 290 frames are detected correctly and remaining 10 frames which are invalid but detected as valid. By analyzing the confusion matrix, SVM classifier has an accuracy of 98.33%, which is more compared to other classifiers. Our empirical results suggest that, given sufficient training data, SVMs tend to be the best classifier. Although the accuracy of classifier is not 100% there is no loss of data that is used for diagnosis since, all valid frames are detected correctly and invalid frames detected as valid frame will not much effect the examination by doctor. After Video validation, further H.264 compression is applied on a video with only valid frames [34].

4.2.7 Pseudo Code for the Proposed Smart Video Validation Algorithm

Ultrasound video is converted into frames and features like skewness, kurtosis, correlation, cluster shade, cluster prominence, homogeneity, sum of square, gray level non-uniformity, run length non-uniformity, and high gray level run selected from genetic algorithm are extracted from each *Frame*. SVM classifier is applied on each frame and resulting output of classifier i.e., *classifier_output* can be 0 or 1. If the *classifier_output* is 0 then the frame is called *valid_frame* (frame with full organ and partial organ). Those frames with full organ and partial organ are valid frames, where they are stored in a buffer. If the *classifier_output* is 1 then the frame is called *invalid_frame* (frame with no organ), these are invalid frames and discarded from the video. All valid frames stored in a $buffer$ are combined to form a $Valid\ Video\ Output$. Steps involved in proposed organ validation in an ultrasound video is shown in Algorithm 2.

Algorithm 2 Smart Ultrasound Video Validation

Initial: Converting *Video to Frames*

Set $N\_f = Numberofframes$;

Input:$Frame(1), Frame(2), ..., Frame(N\_f)$

Output:$Valid\ Video\ Output$

Set $classifier\_output = 0$;

Set $F\_count = 0$;

Set $V\_f = 0$;

Set $buffer = 0$;

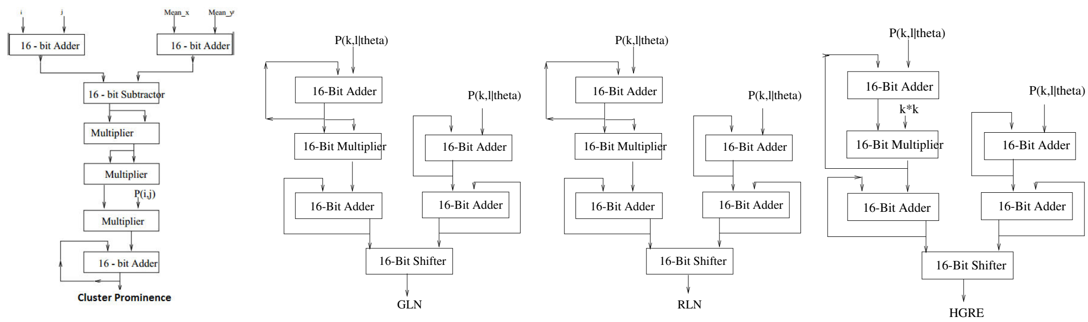
```
while F_count ≠ N_f do
    F_count=F_count + 1;
    Take Frame(F_count);                          ▷ Calculate features for each frame
    Calculate Data.skewness_interval;
    Calculate Data.krutosis_interval;
    Calculate Data.correlation_interval;
    Calculate Data.clustershade_interval;
    Calculate Data.homogeneity_interval;
    Calculate Data.clusterprominence_interval;
    Calculate Data.sumofsquares_interval;
    Calculate Data.GLN_interval;
    Calculate Data.RLN_interval;
    Calculate Data.HGRE_interval;
    Apply SVMClassifier;          ▷ classifier_output can be 0(valid_frame) or 1(invalid_frame)
    if classifier_output = 0 then
        V_f = V_f+1 ;
        Store Frame(F_count) in buffer;
    else
        Discard invalid_frame;
    end if
end while
if V_f=0 then
    RESCAN
else
    Valid Video Output=buffer;
end if
```

4.3 Hardware complexity of video validation

In this section, the hardware complex analysis of the proposed organ validation algorithm is carried for portability on computing platforms. We discussed the hardware complexity involved in extraction of features and classification, which are major blocks in the proposed organ validation algorithm.

Complexity analysis for feature extraction

Hardware analysis of variance, skewness, kurtosis, cluster shade, sum of squares, homogeneity, correlation are explained in section 2.2.1. Hardware architecture to measure cluster prominence is shown in Fig. 4.9a. One 16-bit adder is used to add sum of indices and other to add mean values along x and y axis. The result of these adders is given to 16-bit subtractor. Two multipliers are used to obtain 4^{th} power, other multiplier is used to take the product with probability value at that particular index obtained using GLCM matrix. Finally 16-bit adder is used to sum the values from each element. Hardware architecture to measure GLN is shown in Fig. 4.9b. Input to 16-bit adder is gray level run length matrix elements, output value is squared using multiplier. The squared values are added and divided by sum of all elements of the matrix using 16-bit shifter.



(a) Hardware architecture of cluster prominence.

(b) Hardware architecture of GLN.

(c) Hardware architecture of RLN.

(d) Hardware architecture of HGRE.

Figure 4.9: Hardware architecture of a) Cluster Prominence b) GLN c) RLN d) HGRE.

Hardware architecture to measure RLN is shown in Fig. 4.9c. Input to 16-bit adder is gray level run length matrix elements, output value is squared using multiplier. The squared values are added and divided by sum of all elements of the matrix using 16-bit shifter. Hardware architecture to measure HGRE is shown in Fig. 4.9d. Input to 16-bit adder is gray level run length matrix elements, output value is multiplied with squared value of row index. The resultant values from multiplier are added and divided by sum of all elements of the matrix using 16-bit shifter.

From Table 4.2 we could calculate number of logic gates required to compute the features for hardware complexity analysis. Hence a total of 6649 gates are required to compute the features from each frame of video. Table 4.3 gives total of 31,744 number of Complementary Metal Oxide Semiconductor (CMOS) transistors required to implement the feature extraction.

Table 4.2: Hardware complexity of features calculation in terms of logic gates.

Features	NAND	AND	OR	NOR	XOR	NOT
Variance	192	79	27	-	160	16
Skewness	352	119	35	-	160	32
Kurtosis	352	119	35	-	160	32
ClusterShade	576	175	51	-	96	48
Sum of Square	352	119	35	-	64	32
Homogeneity	100	56	27	-	192	1
Correlation	352	119	35	-	160	32
Cluster Prominence	576	175	51	-	96	48
GLN	288	72	24	-	160	16
RLN	288	72	24	-	160	16
HGRE	288	72	24	-	160	16
Total Gates	3529	1098	341	-	1408	273

Table 4.3: Hardware complexity in terms of CMOS Transistors.

Type of logic gate	No. of CMOS Transistors
NAND	14116
AND	6588
OR	2046
XOR	8448
NOT	546

Complexity analysis for classifier

Hardware complexity for classifier is discussed in section 2.2.2. Table 2.7 gives the total number of gates required for hardware complexity analysis of classifier in terms of logic gates. Total of 8,992 gates are required to design a classifier to determine whether the frame of a video is valid or invalid. From Table 2.8 it is observed that 39,840 CMOS transistors are required to implement classifier algorithm. Therefore, total of 71,584 CMOS transistors are required to implement feature extraction and classification algorithm on the hardware.

4.4 Performance Analysis of proposed smart transmission algorithm

The performance analysis of proposed smart transmission algorithm has been implemented on ARM processor in Zync-7000 architecture. In our analysis we considered 40 videos with different sizes and frame rate. Fig. 4.10 shows time required to run the algorithm on processor, for different size of videos. Data compression and power consumed for transmitting are used as evaluation metric for smart transmission algorithm. The performance of proposed method for video validation and compression of frames in ultrasound video is discussed in the following sections.

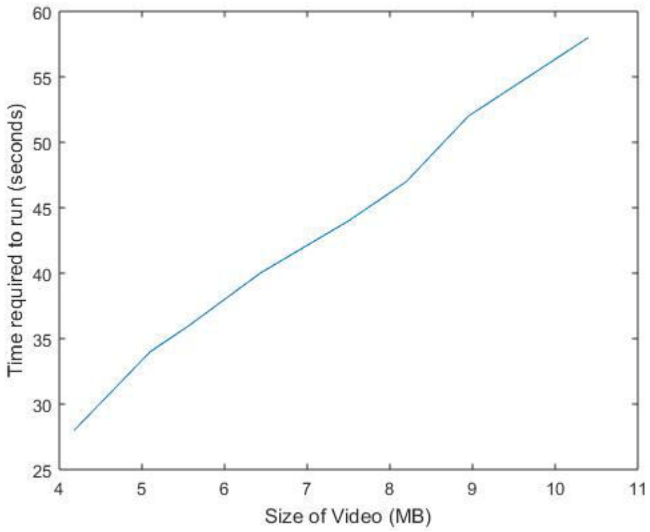


Figure 4.10: Time taken to run proposed algorithm on zedboard.

4.4.1 Video Validation

Region of interest is detected and selected features are extracted based on genetic algorithm, and four different classifiers are used for testing. From Fig. 4.8, SVM classifier is observed to be more accurate with accuracy of 98.33% in detecting the organ frames. Hence, SVM classifier with 10 features selected from genetic algorithm is used to further validate the video. This algorithm can be applied to ultrasound video consisting of any organ with normal and abnormal cases. The result obtained after organ validation in a video is as shown in Fig. 4.11.

4.4.2 Analysis on size of video

Fig. 4.12 shows normal case of kidney ultrasound video having 20 frames before validation and after applying proposed video validation algorithm, frames are reduced to 13 by discarding invalid ones. Original size of video is 4.81 MB, then the size of video is reduced to 2.91 MB after applying proposed algorithm. 20 videos are used for testing and it is observed that there is reduction of size from 38% to 42% compared to original size. Without applying the organ validation algorithm and compressing the video directly using H.264, size is reduced to 829KB. Applying H.264 compression on video after validation, size of encoded video is reduced to 495KB.

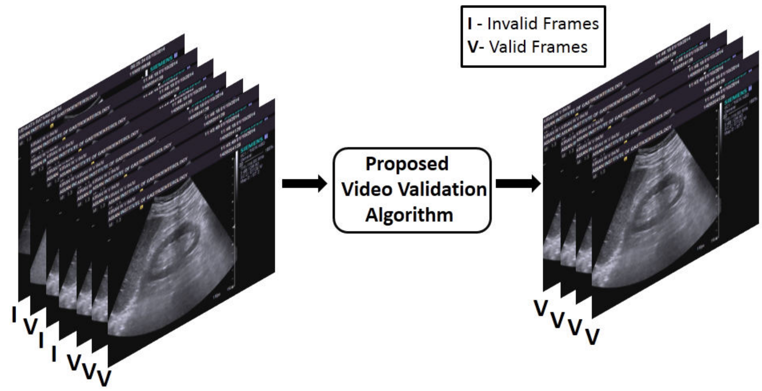


Figure 4.11: Smart Video Validation Output.

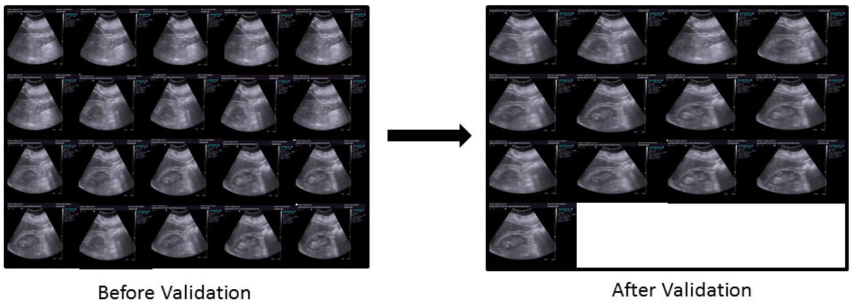


Figure 4.12: Ultrasound video frames of healthy kidney organ.

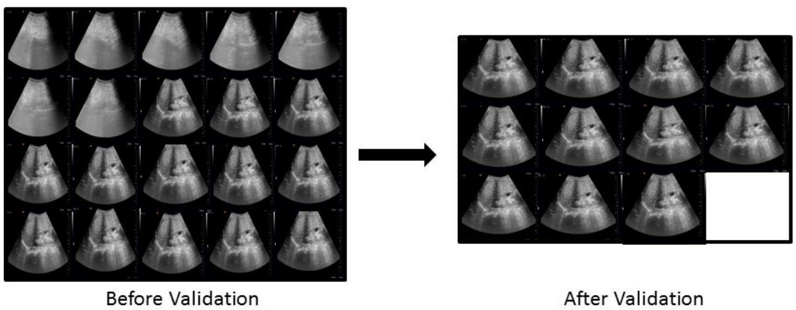


Figure 4.13: Ultrasound video frames of diseased kidney organ.

Compression of video after validation then there is a reduction of 85% to 90% in size compared with original size of video. Similarly, Fig. 4.13 shows abnormal case of kidney ultrasound video having cyst in kidney, after validation of video, 20 frames are reduced to 11. Fig. 4.14 and Fig. 4.15 shows similar analysis to both healthy and diseased liver.

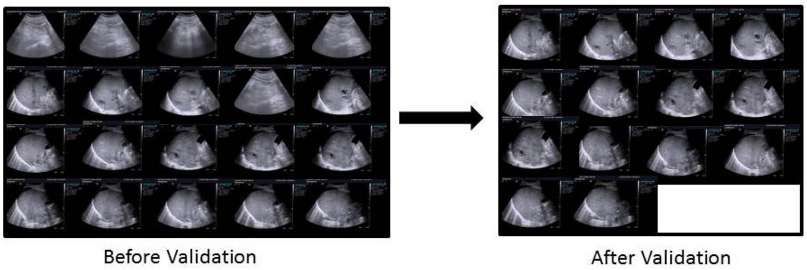


Figure 4.14: Ultrasound video frames of healthy liver organ.

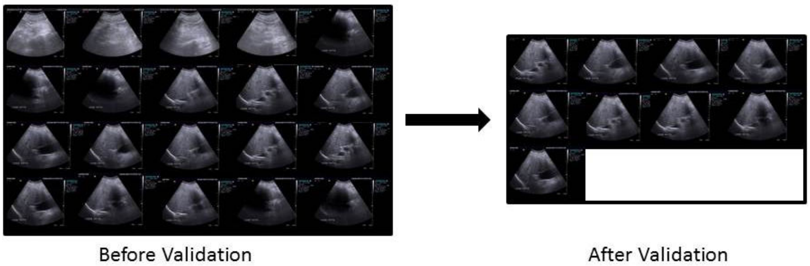


Figure 4.15: Ultrasound video frames of diseased liver organ.

Performance of algorithm is similar irrespective of the organ, whether it is normal or abnormal and can be applied to any organ. Variation in the size of video is plotted for four different scenarios considering three videos as shown in Fig. 4.16.

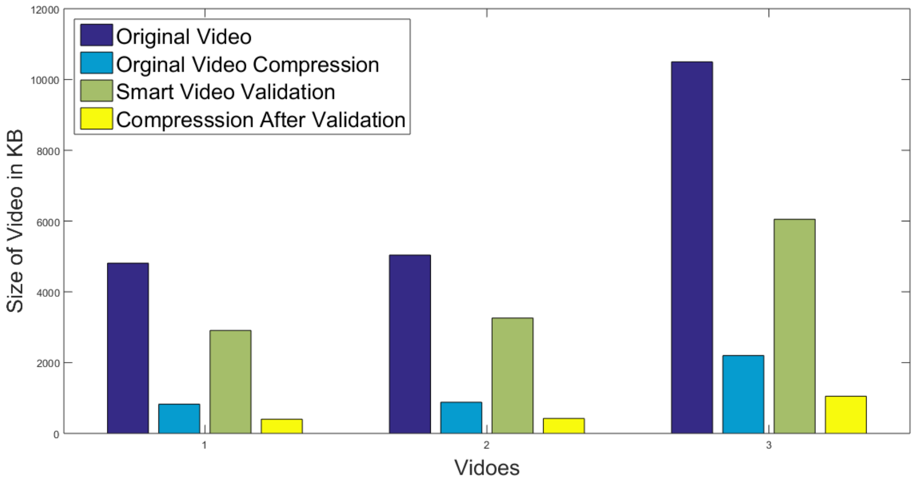


Figure 4.16: Size of video plotted for four different scenarios of three videos.

4.4.3 Time taken to transmit video to cloud using RAK411 as Wi-Fi Module

Time constraint is very important factor to consider while transmission of any data to cloud/server. To transmit the video to cloud through RAK411 as Wi-Fi module interfaced with US board, it takes 296kbps when connected to a mobile phone acting as Wi-Fi access point sharing the mobile data connection. Fig. 4.17 shows time taken to transmit the encoded video without validation and after validation. From this analysis, time consumed for transmission of 829KB is 2.8 seconds, after validation of video it requires only 1.25 seconds. Transmission time is reduced to more than 50% using the proposed smart validation algorithm.

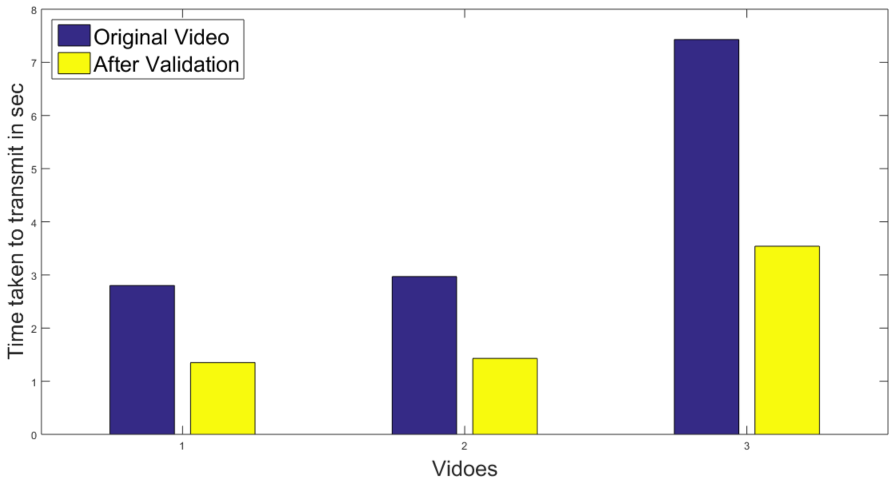


Figure 4.17: Time taken to transmit video using RAK411 as Wi-Fi module for four different scenarios.

4.5 Conclusion

We proposed a novel SoC architecture using Zync 7000 processor for a portable US system with on board video validation algorithm. In our proposed design, we used a zynq SoC which has an FPGA integrated with an ARM cortex A9 dual core processor. To make device more smart, ultrasound video validation algorithm followed by compression for smart transmission is implemented on ARM Cortex A9 dual core processor. RAK411 as Wi-Fi module is interfaced with the US board for transmission of video to cloud for teleradiology applications.

After classification of each frame, invalid frames are discarded from the video by which the video size is reduced to more than 38% compared to original size. Applying H.264 compression on valid video stream, the size of video is further reduced to more than 85% and time consumed for transmission is reduced to more than 50% . Thus, the proposed organ validation algorithm ensures that only valid video data is stored in the cloud, thereby reducing the data in cloud. This also ensures reduction in energy consumption for transmission of video, so that the lifetime of device can be increased.

Chapter 5

Summary and Discussion

My research work has been focused on two main areas (i) Computer Aided Diagnosis for kidney (ii) Smart ultrasound video validation. In this context, the thesis report also includes ROI based compression techniques that best suite ultrasound imaging system, kidney detection algorithm to localize the position of kidney, and reduction in the dimensionality of features required to detect whether the kidney is normal or abnormal and valid or invalid. The algorithms are tested on various platforms and a large database to provide proof of concepts and support the claims made. In the course of these studies, several interesting insights were obtained. Field trails have been performed in the presence of doctor and reviewed our results under his presence. This chapter concludes my thesis in brief and discusses in future scope in this direction.

5.1 Computer Aided Diagnosis for Kidney

We proposed a fully-automated kidney abnormality detection system based on automated feature selection and supervised classification. Different techniques are used to find ROI to extract features from segmneted region. The Viola Jones algorithm gives best accuracy to detect kidney automatically. The texture features are computed from detected kidney and SVM classifier is used for classifying the kidney image as normal or abnormal. The simulations are verified on Xilinx Zed board, which has a capability of implementing entire ultrasound scanning system. The processor on Xilinx Zed board, can perform the computer aided diagnosis on the system and hence provides best solution for integrating both signal processing and diagnosis on a single system. Providing such information helps sonographers to suggest immediate precaution and also monitor disease progression. Thus the proposed technique aids prelimianry CAD for kidney on ultrasound systems.

5.2 Smart Ultrasound video validation

We proposed a novel SoC architecture using Zync 7000 processor for a portable US system with on board video validation algorithm. In our proposed design, we used a zynq SoC which has an FPGA integrated with an ARM cortex A9 dual core processor. To make device more smart, ultrasound video validation algorithm followed by compression for smart transmission is implemented on ARM Cortex A9 dual core processor. RAK411 as Wi-Fi module is interfaced with the US board for transmission of video to cloud for teleradiology applications. Video is validated based on the organ information present in each frame and discards the frames not useful for diagnosis. This method reduces the size of video thereby reducing the data storage in cloud to avoid network trafficking. This also ensures reduction in energy consumption for transmission of video, so that the lifetime of device can be increased.

5.3 Future Scope

As a future work, we would like to extend our algorithms to wide scope of usage. Extending the CAD to many few more organs to automatically to detect abnormality.

References

[1] Ewertsen, Caroline, et al. "Real-time image fusion involving diagnostic ultrasound." American Journal of Roentgenology 200.3 (2013): W249-W255.

[2] Mack, Michael J. "Minimally invasive and robotic surgery." Jama 285.5 (2001): 568-572.

[3] Wootton, Richard, John Craig, and Victor Patterson. Introduction to telemedicine. Royal Society of Medicine Press, 2006.

[4] "More than 77 million people in India have pre-diabetes," The Hindu, Coimbatore, January 27, 2014.

[5] H. Lee, H. Sohn, C. Yoon, Y. Yoo, and T.K. Song, "Software-based Hand-Held Ultrasound Color Doppler Imaging System," IEEE Ultrasonics symposium, pp.1844-1847, 2009.

[6] B. P. Nelson, and K. Chason, "Use of ultrasound by emergency medical services: a review," International Journal of Emergency Medicine, vol. 1, no. 4, pp. 253-259, 2008.

[7] J.J. Hwang, and J. Quistgaard, "A handheld ultrasound array imaging device," in Proc. SPIE, vol. 3664, pp. 194-201, 1999

[8] F. K. Schneider, A. Agarwal, Y. M. Yoo, T. Fukuoka, and Y. Kim, "A fully programmable computing architecture for medical ultrasound machines," IEEE Trans. Inf. Technol. Biomed., vol. 14, no. 2, pp. 538âĂŞ540, 2010.

[9] Kang, Jeeun, et al. "A System-on-Chip Solution for Point-of-Care Ultrasound Imaging Systems: Architecture and ASIC Implementation." IEEE Transactions on Biomedical Circuits and systems, 2015.

[10] Kim, Gi-Duck, et al. "A single FPGA-based portable ultrasound imaging system for point-of-care applications." Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on 59.7 (2012): 1386-1394.

[11] Ahn, Sewoong, et al. "Smartphone-based portable ultrasound imaging system: Prototype implementation and evaluation." Ultrasonics Symposium (IUS), 2015 IEEE International. IEEE, 2015.

[12] a. agarwal, F. K. schneider, y. M. yoo, and y. Kim, "Image quality evaluation with a new phase rotation beamformer," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 55, no. 9, pp. 1947-1955, 2008.

[13] B. G. Tomov and j. a. jensen, "compact FPGA-based beamformer using oversampling 1-bit a/d converters," IEEE Trans. Ultrason. Ferroelectr. Freq. Control, vol. 52, no. 5, pp. 870-880, 2005.

[14] H. H. Han, j. j. Kim, and T.-K. song, "Hardware-efficient methods for eliminating of signal distortion in sigma-delta based ultrasound beamformer," Ultrason. Imaging, vol. 31, no. 2, pp. 101-119, 2009.

[15] Doi, Kunio. "Computer-aided diagnosis in medical imaging: historical review, current status and future potential." Computerized medical imaging and graphics 31.4-5 (2007): 198-211.

[16] Raja, K. Bommanna, M. Madheswaran, and K. Thyagarajah. "A hybrid fuzzy-neural system for computer-aided diagnosis of ultrasound kidney images using prominent features." Journal of Medical Systems, vol. 32.1, pp. 65-83, yr-2008.

[17] Hafizah, Wan Mahani, Eko Supriyanto, and Jasmy Yunus. "Feature extraction of kidney ultrasound images based on intensity histogram and gray level co-occurrence matrix." Modelling Symposium (AMS), 2012 Sixth Asia. IEEE, 2012.

[18] Krishna, K. Divya, et al. "FPGA based preliminary CAD for kidney on IoT enabled portable ultrasound imaging system." e-Health Networking, Applications and Services (Healthcom), 2014 IEEE 16th International Conference on. IEEE, 2014.

[19] Noble, J. Alison, and Djamal Boukerroui. "Ultrasound image segmentation: a survey." Medical Imaging, IEEE Transactions on vol.25.8, pp. 987-1010, yr. 2006

[20] Jun Xie; Yifeng Jiang; Hung-Tat Tsui, "Segmentation of Kidney From Ultrasound Images based on Texture and Shape Priors," Medical Imaging, IEEE Transactions on , vol. 24.1, pp.45-57, yr. 2005.

[21] Hafizah, Wan Mahani, and Eko Supriyanto. "Automatic region of interest generation for kidney ultrasound images." Proceeding of Recent Researches in Applied Informatics and Remote Sensing (2011): 70-75.

[22] Martn-Fernndez, Marcos, and Carlos Alberola-Lopez. "An approach for contour detection of human kidneys from ultrasound images using Markov random fields and active contours." Medical Image Analysis, vol. 9.1, pp. 1-23, yr. 2005.

[23] Ali, Asem M., Aly A. Farag, and Ayman S. El-Baz. "Graph cuts framework for kidney segmentation with prior shape constraints." Medical Image Computing and Computer-Assisted Intervention MICCAI Springer Berlin Heidelberg, pp. 384-392. yr. 2007.

[24] Prevost, Raphael, et al. "Kidney Detection and Segmentation in Contrast-Enhanced Ultrasound 3D Images." Abdomen and Thoracic Imaging. Springer US, pp. 37-67, yr. 2014. 44

[25] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." Computer Vision and Pattern Recognition 2001. IEEE Computer Society Conference on. Vol. 1. IEEE, yr. 2001.

[26] Panayides, Andreas S., Zinonas C. Antoniou, and Anthony G. Constantinides. "An Overview of mHealth Medical Video Communication Systems." Mobile Health. Springer International Publishing, 2015. 609-633.

[27] Panayides, Andreas, et al. "High-resolution, low-delay, and error-resilient medical ultrasound video communication using H. 264/AVC over mobile WiMAX networks." Biomedical and Health Informatics, IEEE Journal of 17.3 (2013): 619-628.

[28] Panayides, Andreas S., et al. "An Effective Ultrasound Video Communication System Using Despeckle Filtering and HEVC." Biomedical and Health Informatics, IEEE Journal of 19.2 (2015): 668-676.

[29] Cavero, Eva, et al. "SPIHT-based echocardiogram compression: clinical evaluation and recommendations of use." Biomedical and Health Informatics, IEEE Journal of 17.1 (2013): 103-112.

[30] Quellec, Gwenole, et al. "Real-time segmentation and recognition of surgical tasks in cataract surgery videos." Medical Imaging, IEEE Transactions on 33.12 (2014): 2352-2360.

[31] Wang, Shijun, et al. "Seeing is believing: Video classification for computed tomographic colonography using multiple-instance learning." Medical Imaging, IEEE Transactions on 31.5 (2012): 1141-1153.

[32] Melendez, Jaime, et al. "A Novel Multiple-Instance Learning-Based Approach to Computer-Aided Detection of Tuberculosis on Chest X-Rays." Medical Imaging, IEEE Transactions on 34.1 (2015): 179-192.

[33] zedboard [Online]. Available: http://zedboard.org/support/documentation/1521.

[34] Yu, Hongtao, Zhiping Lin, and Feng Pan. "Applications and improve- ment of H. 264 in medical video compression." Circuits and Systems I: Regular Papers, IEEE Transactions on 52.12 (2005): 2707-2716.

[35] Achim, Alin, Anastasios Bezerianos, and Panagiotis Tsakalides. "Novel Bayesian multiscale method for speckle removal in medical ultrasound images." Medical Imaging, IEEE Transactions on 20.8 (2001): 772-783.

[36] Wei, Dong, Umesh Rajashekar, and Alan C. Bovik "3.4 Wavelet Denoising for Image Enhancement."

[37] Achim, Alin, Anastasios Bezerianos, and Panagiotis Tsakalides. "Novel Bayesian multiscale method for speckle removal in medical ultrasound images" Medical Imaging, IEEE Transactions on 20.8 (2001): 772-783.

[38] MR.Hitesh S. ASARI, ASS.PROF. Ami shah "Research Paper On Reducion Of Speckle Noise In Ultrasound Imaging Using Wavelet And Contourlet Transform", Journal of Information, Knowledge and Research in Electronics and Communication Engineering, Nov 12, Volume - 02, Issue - 02.

[39] R. Keys, (1981). "Cubic convolution interpolation for digital image processing". IEEE Transactions on Acoustics, Speech, and Signal Processing 29 (6): 1153-1160. doi:10.1109/TASSP.1981.1163711

[40] Bharath, R., et al. "FPGA-Based Portable Ultrasound Scanning System with Automatic Kidney Detection." Journal of Imaging 1.1 (2015): 193-219.

[41] R. M. Haralick, K. Shanmugan and I. Dinstein, "Textural Features for Image Classification", *IEEE Transactions on Systems:Man,and Cybernetics SMC*, Vol. 3, 1973, pp. 610- 621. doi:10.1109/TSMC.1973.4309314

[42] Eleyan, Alaa, and Hasan Demirel. "Co-occurrence matrix and its statistical features as a new approach for face recognition." Turkish Journal of Electrical Engineering and Computer Sciences 19.1 (2011): 97-107.

[43] Scholkopf, Bernhard, and Alexander J. Smola. Learning with kernels: Support vector machines, regularization, optimization, and beyond. MIT press, 2002.

[44] Camps-Valls, Gustavo, and Lorenzo Bruzzone. "Kernel-based methods for hyperspectral image classification." Geoscience and Remote Sensing, IEEE Transactions on 43.6 (2005): 1351-1362.

[45] Anjuli, Satyajit Anand. "2-Bit Magnitude Comparator Design Using Different Logic Styles." International Journal of Engineering Science Invention 2.1 (2013): 13-24.

[46] Acharyya, Amit, et al. "Coordinate rotation based low complexity ND FastICA algorithm and architecture." Signal Processing, IEEE Transactions on 59.8 (2011): 3997-4011.

[47] Gupta, Vineeta, A. G. Rao, and Krishna Mohan Pandey. "Integer Wavelet Transform and Predictive Coding Technique for Lossless Medical Image Compression." International Journal of Technological Exploration and Learning (IJTEL) Volume 2 Issue 4 (August 2013)

[48] Guo, Shu-Mei, et al. "Performance Improvement of Set Partitioning Embedded Block Algorithm for Still Image Compression." Modern Advances in Applied Intelligence. Springer International Publishing, 2014. 270-278.

[49] Jha, Mukund, Amarjeet Singh Yumnam, and Dilip Raju. "Comparison between image codecs: JPEG and JPEG2000." Computing for Sustainable Global Development (INDIACom), 2014 International Conference on. IEEE, 2014.

[50] Hore, Alain, and Djemel Ziou. "Image quality metrics: PSNR vs. SSIM." Pattern Recognition (ICPR), 2010 20th International Conference on. IEEE, 2010.

[51] Hemalatha.R, Radha.S, Raghuvarman.N, Soumya.B, and Vivekanandan.B. "Energy Efficient Image Transmission over Bandwidth Scarce WSN using Compressed Sensing ." International Conference on IT and Intelligent Systems (ICITIS'2013) August 28-29, 2013 Penang (Malaysia).

[52] Tai, Shen-Chuan, Zih-Siou Chen, and Wei-Ting Tsai. "An automatic mass detection system in mammograms based on complex texture features." Biomedical and Health Informatics, IEEE Journal of 18.2 (2014): 618-627.

[53] Hafizah, Wan Mahani, Eko Supriyanto, and Jasmy Yunus. "Feature extraction of kidney ultrasound images based on intensity histogram and gray level co-occurrence matrix." Modelling Symposium (AMS), 2012 Sixth Asia. IEEE, 2012.

[54] Krishnan, K. Raghesh, and R. Sudhakar. "Automatic Classification of Liver Diseases from Ultrasound Images Using GLRLM Texture Features." Soft Computing Applications. Springer Berlin Heidelberg, 2013. 611-624.

[55] Babatunde, Oluleye, et al. "Comparative Analysis of Genetic Algorithm and Particle Swam Optimization: An Application in Precision Agriculture." Asian Journal of Computer and Information Systems (ISSN: 2321-5658) 3.01 (2015).

[56] Ji, Rongrong, et al. "Spectral-spatial constraint hyperspectral image classification." Geoscience and Remote Sensing, IEEE Transactions on 52.3 (2014): 1811-1824.

[57] Cover, T., Hart, P. , "Nearest-neighbor pattern classification", Information Theory, IEEE Transactions on, Jan. 1967, pp. 21-27.

[58] Duda, Richard O., Peter E. Hart, and David G. Stork. Pattern classification. John Wiley and Sons, 2012.

[59] McCann, Sancho, and David G. Lowe. "Local naive bayes nearest neighbor for image classification." Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012.

[60] Mountrakis, Giorgos, Jungho Im, and Caesar Ogole. "Support vector machines in remote sensing: A review." ISPRS Journal of Photogrammetry and Remote Sensing 66.3 (2011): 247-259.