Sadanand L. Shelgaonkar* and Anil B. Nandgaonkar

Deep Belief Network for the Enhancement of Ultrasound Images with Pelvic Lesions

DOI 10.1515/jisys-2016-0112

Received July 16, 2016; previously published online May 20, 2017.

Abstract: It is well known that ultrasound images are cost-efficient and exhibit hassle-free usage. However, very few works have focused on exploiting the ultrasound modality for lesion diagnosis. Moreover, there is no reliable contribution reported in the literature for diagnosing pelvic lesions from the pelvic portion of humans, especially females. While few contributions are found for diagnosis of lesions in the pelvic region, no effort has been made on enhancing the images. Inspired from the neural network (NN), our methodology adopts deep belief NN for enhancing the ultrasound image with pelvic lesions. The higher-order statistical characteristics of image textures, such as entropy and autocorrelation, are considered to enhance the image from its noisy environment. The alignment problem is considered using skewness. The proposed method is compared with the existing NN method to demonstrate its enhancement performance.

Keywords: Pelvic lesions, ultrasound, deep belief network, enhancement, noise.

1 Introduction

A lesion is any region in the tissue that is caused by damage, injury, or disease. It may occur in any part of the body, especially the mouth, skin, brain, or any tumor regions. Melanocytic nevus, seborrheic keratosis, actinic keratosis, squamous cell carcinoma, basal cell carcinoma, and melanoma are the types of lesions [38]. Formerly, traditional techniques such as cervicography and visual inspection with acetic acid [34] were used for performing lesion diagnosis. However, those techniques were too expensive and laboratory intensive [5, 24].

Computer-aided diagnosis has been in demand in recent years [25]. Various types of imaging techniques, such as computed tomography [33], magnetic resonance imaging [8], mammography, ultrasonography [17], and positron emission tomography [11], have been developed for the easy diagnosis of lesion types. These imaging techniques help in diagnosing various types of cancers in the pelvic region of the human female. Digital colposcopy with automated image analysis techniques [16, 19, 22, 23, 26, 27] was used as a convenient technique for diagnosing cervical cancer in the early days. However, this traditional technique has failed in locating the specific area of interest [9]. Thus, an advanced automated domain-specific image analysis method has been used for detecting cervical neoplasias [28].

Imaging techniques have lot of disadvantages, such as harmful radiation, low specificity, complicated system, and low image resolution [4, 14]. Hence, there is a need for innovative location-specific imaging techniques that enable accurate diagnosis with low operational cost, minimized radiation, high sensitivity, and easy handling [18].

This paper proposes an image enhancement algorithm using deep belief network (DBN), which selects the optimal band of filtering based on higher-order features such as autocorrelation, entropy, and skewness. The DBN performance is ensured based on non-reference-based quality measurement, called edge-based structural similarity (ESSIM). The ESSIM has the feature of estimating the quality of enhancement without

*Corresponding author: Sadanand L. Shelgaonkar, Pacific Academy of Higher Education and Research University, Udaipur, Rajasthan, India, e-mail: shelgaonkarsadanand@gmail.com

Anil B. Nandgaonkar: Dr. Babasaheb Ambedkar Technological University, Lonere, Raigad, Maharashtra, India

any reference image. The similar circumstance appears in real-time problems, and hence the DBN is permitted to train based on the ESSIM outcome.

1.1 Our Contributions

The contributions of this paper are as follows:

- (1) Contribution 1: An effort is made to enhance ultrasound images with lesions of the pelvic region, as such work is not prevalent, to the best of our knowledge.
- (2) Contribution 2: Inspired from neural network (NN), we employ DBN for enhancing the lesion images of the ultrasound modality.
- (3) Contribution 3: A new paradigm for constructing the training library is proposed using the higher-order statistics of texture features and the decomposing mode of filter.

1.2 Article Outline

The paper is organized to the following sections: Section 2 reviews the literature, and Section 3 illustrates the texture analysis on the images to support image enhancement. Section 4 details the procedure for training library construction and the information about DBN. Section 5 discusses the results, and Section 6 concludes the paper.

2 Literature Review

2.1 Related Works

In 2010, Alush et al. [2] introduced novel steps for extracting and segmenting the class-specific object automatically by using class-specific boundaries, in addition to focusing on the determination of lesions in the uterine cervix images. They exploited the Markov random field to model the watershed segmentation map of the input image, and evaluated whether the region is a lesion or not. Also, they used the local pair-wise factors that depend on supervised learning for visual word distribution to indicate whether the arc is a part of the lesion or not.

In 2011, Park et al. [28] developed a framework using domain-specific automated image analysis. The proposed framework was used for detecting the cancerous as well as the precancerous lesions in the uterine cervix. Thus, for automatic diagnosis, they employed domain-specific diagnostic features using conditional random fields and a novel window-based performance assessment strategy. The new window-based performance scheme solved the problem of image misalignment.

In 2013, Lee and Won [18] developed a tactile sensation imaging system for capturing the lesions of breast cancer. They applied the principle of total internal reflection for lesion capturing, and introduced a novel method for evaluating the lesion size, depth, and elasticity. They applied two algorithms, namely neural-network-based inversion algorithm and three-dimensional finite-element-model-based forward algorithm, for evaluating the lesions and validated the results with realistic tissue phantom. Harmouche et al. [13] developed a fully automatic probabilistic method for classifying the lesions of multiple sclerosis. Based on the multimodal magnetic resonance imaging intensities, they built the regional likelihood distributions for each tissue class in the segmented brain. They applied Markov random fields for ensuring smoothness of the local class by using the neighboring voxel information. They used various metrics, such as dice overlap, positive predictive rates, and sensitivity for evaluating and comparing the voxel and the lesion classification.

In 2015, Zhan et al. [39] exploited T1 and fluid-attenuated inversion recovery image modalities for the automatic segmentation of white matter lesions. The T1 image was segmented into cerebrospinal fluid, gray matter, and white matter using the brain tissue segmentation method. Also, they defined the threshold for the preliminary identification of abnormalities in the tissues. The boundaries of white matter lesions were extracted using the level set method, which is based on the local Gaussian distribution energy.

2.2 Problem Definition

A literature review on the diagnostic methodologies of lesion presence in multimodal images reveals two primary research gaps (Table 1). First, adequate contributions have been made toward diagnostic methods,

Table 1: Summary of Literature Review.

Author [Citation]	Adopted methodology	Advantages	Disadvantages
Zhan et al. [39]	Level set method	Adaptive over the shapes under disturbed and homogenous environment	Inability to handle complex structures of lesions
Harmouche et al. [13]	Markov Random Field	Establishes relativity among the spatial information, removes the noise and strengthens the cohesiveness	Fails to determine global minimal points
Park et al. [28]	Conditional random field	Handles the misalignment due to maximum entropy	Iterative scaling often causes poor enhancement
Alush et al. [2]	Watershed segmentation	Characterizes the regional intensity	Suffers from over segmentation
Lee and Won [18]	Neural Network	Efficient for image enhancement	Long and complex time matched parameter measurements
Romagnolo et al. [31]	ROMA	Assesses the likelihood of ovarian mass malignancy.	Cannot be used without an independent clinical or radiological evaluation and unfit for women who have high rheumatoid factor
Roma and Kelly [30] Xin et al. [36]	PAX8 Expectation maximization algorithm	Accurate lesion classification, reliable Simplicity and ease of implementation	Low sensitivity Slow convergence, inability to provide estimation to the asymptotic variance-covariance matrix of the maximum likelihood estimator
Badawy et al. [3]	US modality	Quick and painless, identifies the lesion clearly, no health issues,	False-positive results, many types of cancers cannot be detected, requires skilled operator
Mahajan et al. [21]	Multiparametric MRI	Diagnostic accuracy	Spectral contamination of MR spectra
Yavariabdi et al. [37]	Deformable slice to volume registration method	Ability to decouple the plane selection and inplane deformation parts, less computational time	Inconsistency problem
Brocker et al. [5]	Electron MRI	Non-invasive, no radiation, less allergic reaction	Expensive, cannot detect all types of cancers
Epstein et al. [10]	US	Quick and painless, identifies the lesion clearly, no health issues	False-positive results, many types of cancers cannot be detected, requires skilled operator
	MRI	Non-invasive, no radiation, less allergic reaction	Expensive, cannot detect all types of cancers

but not on improving the diagnostic precision. Improved diagnostic precision can be achieved only if the image is informative. However, the imaging techniques often produce unwanted signals and, as a result, the images easily become corrupted by noise. The corrupted image loses many of its useful information and, thus, the diagnosis will not be precise enough to meet the practical constraints. Second, lesion diagnosis has been done in many human parts. Lesions in the pelvic region have not been considered sufficiently in the literature. Despite the fact that the existing diagnostic methodologies are significant for medical applications, they suffer from serious drawbacks. In Ref. [39], the level set method was adopted for the segmentation of white matter lesions from the brain tissue. The enhancement of tissue visualization was indirectly performed using the Gaussian distribution function, as the majority of the noise follows the characteristics of the Gaussian function. Though the performance of the level set method is not affected by local intensities, it suffers due to complex geometrical structures. Hence, handling of pelvic lesions under a low-quality scenario remains challenging with the level set method-based segmentation.

Markov random field was deployed in Ref. [13] for the classification of multiple sclerosis lesions. The Markov random field was primarily employed for classification purposes, yet it removed the noise from the subjected image because of its probabilistic characteristics. Its probabilistic nature established relativity among the spatial data of the image and, hence, noisy data were skipped off from the image. As a result, the classification was done in a noise-mitigated environment. However, the Markov random field suffers due to its nature of sticking with the local optimal points to establish the relativity. Hence, noise removal cannot be achieved to a substantial level, degrading the quality of classification. Conditional random field, which is a member of the family of Markov random field, was exploited in Ref. [28] to determine the lesions in the uterine cervix. As the imaging systems often produce misalignment in the resultant image, a window-based scheme was used along with the conditional random field. The window-based processing improved the image for the betterment of diagnosis accuracy. The noise removal task was tackled through the conditional random field, yet it suffered due to the iterative scaling problem. As a result, sufficient image enhancement was not accomplished. In Ref. [2], watershed segmentation and Markov random field were exploited to learn the characteristics of the lesion portion of the image and to segment the lesions from the uterine cervix image. The Markov random field was employed to mitigate the noise, and later, watershed segmentation was applied to carry out the segmentation process. While the Markov random field suffers from the inability to distinguish the local optimal points and the global optima, the watershed segmentation often results in oversegmentation. The oversegmentation that resulted from the image provided a wide heterogeneous environment.

Under a circumstance of heterogeneous noisy characteristics, the image undergoes multimodal format, where both the Markov random field and watershed segmentation fail to perform image enhancement and segmentation, respectively. In Ref. [18], NN was deployed to understand the degree of lesions in breast cancer. The NN supported image enhancement, unless the data interpolation deviated from the actual data. Moreover, the finite element modeling technique interpolation was based on the statistical nature of the image. The lack of practical measurements possibly degraded the performance of the NN.

3 Texture and Alignment Analysis on the US Images of the Pelvis

3.1 Texture Analysis

Assume a US image that contains a number of pixels as N_v and N_h in the vertical and horizontal directions, respectively. In every pixel, the gray level is quantized to a level of N_g . Let $I_v = \{1, 2, \dots, N_v\}$, $I_h = \{1, 2, \dots, N_h\}$, and $G = \{1, 2, \dots, N_g\}$ represent the vertical spatial domain, horizontal spatial domain, and set of N_g quantized gray levels, respectively. The resolution clique of an image that is ordered by its row-column designations is represented as a set of $I_v \times I_h$. Let I be an image that is assigned as a function with G as the gray level in all the resolution cells or as a coordinate pair in $I_v \times I_h$; $I_v \times I_h \rightarrow G$.

The texture features are taken from the four closely related measures, named as angular nearest-neighbor gray tone spatial dependence matrices, and these matrices represent the adjacent neighbor resolution cells. A resolution cell without the peripheral image and with eight nearest-neighbor resolution cells is also considered. In an image I , the information of the texture content is occupied in the average or the complete spatial relationship. Assume the information of the texture content to be indicated with a matrix of relative frequencies f_{ij} and that contains two neighbor resolution cells with distance d on the image, which has the gray tone i and j . This type of matrix with gray tone spatial dependence frequencies indicates two functions, namely the distance between the cells and the angular relationship between the nearby resolution cells. In a horizontal neighboring resolution cell set with distance 1 and gray tone image, the distance 1 horizontal gray tone spatial dependence matrix is estimated. In a 45° quantized angle interval, the unnormalized frequencies are given by

$$\begin{aligned} f(i, j, d, 0^\circ) = & \{((k, l), (m, n)) \in (I_v \times I_h) \\ & \times (I_v \times I_h) \mid k-m=0, |l-n|=d, \\ & I(k, l)=i, I(m, n)=j\} \vee \text{all the elements,} \end{aligned} \quad (1)$$

$$\begin{aligned} f(i, j, d, 45^\circ) = & \{((k, l), (m, n)) \in (I_v \times I_h) \\ & \times (I_v \times I_h) \mid (k-m=d, l-n=-d) \\ & \text{or } (k-m=-d, l-n=d), \\ & I(k, l)=i, I(m, n)=j\}, \end{aligned} \quad (2)$$

$$\begin{aligned} f(i, j, d, 90^\circ) = & \{((k, l), (m, n)) \in (I_v \times I_h) \\ & \times (I_v \times I_h) \mid |k-m|=d, \\ & l-n=0, I(k, l)=i, I(m, n)=j\}, \end{aligned} \quad (3)$$

$$\begin{aligned} f(i, j, d, 135^\circ) = & \{((k, l), (m, n)) \in (I_v \times I_h) \\ & \times (I_v \times I_h) \mid (k-m=d, l-n=d) \\ & \text{or } (k-m=-d, l-n=-d), \\ & I(k, l)=i, I(m, n)=j\}, \end{aligned} \quad (4)$$

where Eqs. (1)–(4) are performed for all the elements that are available in the set. The matrices are symmetric in nature and, so, $f(i, j; d, a) = f(j, i; d, a)$. Further, ρ is suggested as the distance metric and, so, $\rho((k, l), (m, n)) = \max\{|k-m|, |l-n|\}$.

Assume a 4×4 image with number of gray tones as 4, which ranges from 0 to 3. In a gray tone spatial dependence matrix, the (2, 1) position's element with distance 1 horizontal f_h matrix refers to the number of times both the gray tones with values 1 and 2 appear nearby to each other horizontally. This number is estimated using the number of pairs in the resolution cells in R_h , so that the first as well as the second resolution cell consist of gray tone 2 and 1, respectively. All the four distance 1 gray tone spatial dependence matrices and the frequency normalizations are estimated. In each row, there are about $2(N_h - 1)$ neighboring resolution cell pairs, N_v rows, and a total of $2N_v(N_h - 1)$ adjacent horizontal neighbor pairs. This shows the nearest horizontal neighbor relationship ($d=1, a=0^\circ$).

A total of $2(N_v - 1)(N_h - 1)$ adjacent right diagonal neighbor pairs are noted, when the neighboring resolution cell pair for each row, except the first row, is $2(N_h - 1)45^\circ$ and there are N_v rows. The existing relationship is the nearest right diagonal neighbor ($d=1, a=45^\circ$). With respect to symmetry, there are $2N_h(N_v - 1)$ and $2(N_h - 1)(N_v - 1)$ nearest vertical and nearest left diagonal neighbor pairs, respectively. The gray tone spatial dependence matrix is normalized by dividing the elements in the matrix with R , if the neighbor resolution

pair R is applied on the matrix. During image processing, the number of resolution cells n is directly proportional to the number of operations in the proposed method. Specifically, if Hadamard or Fourier transform is used to extract the texture information, then the number of operations will be in the order of $n \log n$. It is necessary to provide two lines of image data at a time to calculate the gray tone spatial dependence matrix entries.

In this paper, entropy that is mathematically depicted below and the autocorrelation of the gray level co-occurrence matrix are considered to understand the texture characteristics of the image. To demonstrate the significance of selecting the two features, we select five US images with pelvic lesions, as given in Figure 1. For the selected images, entropy and autocorrelation are determined and illustrated through the quantile-quantile plots of Figure 2A and C. As the quantile-quantile plots show the mutual distribution of the images, it is capable of distinguishing the texture means of the images. This is well visualized through Figure 2B and C, where each image has its own plot, rather than overlapping with the other plots. This shows that the selected entropy and autocorrelation represent the US images effectively.

(1) Entropy: In texture analysis, entropy refers to the spatial disorder measure [12, 32]:

$$Entropy = -\sum_{i,j} p(i, j) \log(p(i, j)). \quad (5)$$

In random distribution, entropy tends to be so high due to the chaos. However, the entropy value sets to zero for a solid tone image. This feature of entropy is helpful in predicting the type of textures that are statistically more chaotic in nature.

3.2 Alignment Analysis

In order to ensure the alignment of the US image, an alignment feature of the image, termed as skewness, is extracted. Skewness is defined in relation to the third as well as the second moments around the mean, and it is represented as

Figure 1: Sample US Images of the Pelvic Region with Pelvic Lesions.

Figure 2: Texture Analysis.

(A) Skewness; (B) entropy; (C) autocorrelation for five images.

$$m_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2, \quad (6)$$

$$m_3 = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3. \quad (7)$$

The traditional Fisher-Pearson coefficient of skewness is given as

$$g_1 = \frac{m_3}{m_2^{3/2}} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left[\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right]^{3/2}}. \quad (8)$$

g_1 may occur in negative values and, so, the Pearson statistic is known to be $\sqrt{\beta_1}$. For the case of g_1 , when used as a test for deviation from normality, a table is introduced by Pearson and Hartley in the year 1970 [29]. This Pearson formula exists as a modification of the sample size, and an adjusted Fisher-Pearson standardized moment coefficient has been introduced, which is given as

$$G_1 = \frac{n}{(n-1)(n-2)} \sum_{i=1}^n \left(\frac{x_i - \bar{x}}{s} \right)^3. \quad (9)$$

The selection of skewness to represent the alignment of the US, with respect to its texture features, is substantiated using Figure 2A that is plotted for the five images of Figure 1. Figure 2A depicts the power spectral density of the skewness parameter, with respect to its frequency of occurrence. The skewness of those images show high variations with increase in the normalized frequency. Each image exhibits a good skewness variation and ensures that the alignment of the image is recognized.

Hence, the extracted features are subjected to train the DBN for allowing the decomposing band of filter to be determined.

4 DBN for Enhancing US Images

4.1 Weightage-Based Quality Assessment

While the training attributes of the US images are extracted through texture analysis, the target of the image is set through weightage-based quality assessment. The weightage-based quality assessment includes both the reference-based quality assessment and the non-reference-based quality assessment. The reference-based quality assessment considers the peak signal-to-noise ratio (PSNR) measure, and the non-reference-based quality assessment considers the edge-based quality metric, termed as the ESSIM [6]. The PSNR measure considers the mean squared error (MSE) between the enhanced image and the original image (before noise contamination), as per Eq. (8).

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2, \quad (10)$$

$$\text{PSNR} = 10 \cdot \log_{10} \left(\frac{\text{MAX}_I^2}{\text{MSE}} \right), \quad (11)$$

$$\text{PSNR} = 20 \cdot \log_{10} \left(\frac{\text{MAX}_I}{\sqrt{\text{MSE}}} \right), \quad (12)$$

where MAX_I represents the maximum image pixel. The MAX_I value is 255, if the pixels are indicated with 8 bits per sample. However, if the pixels are indicated with linear pulse-code modulation (PCM) and B bits per sample, then $MAX_I = 2^B - 1$.

- (1) ESSIM: The ESSIM [7] uses the information of the edge to compare the information between the original image block and the distorted image block at the same time, and to replace the structure comparison $s(x, y)$ with the edge-based structure comparison $e(x, y)$. The edge information can be obtained using local gradients, Sobel operator, or simple edge detection algorithm. Hence, the ESSIM calculation is performed in three steps sequentially, namely edge map calculation, determination of the edge direction vector, and edge comparison.

Let the pixel be $p_{i,j}$ and the edge vector for each pixel be given as $\vec{D}_{i,j} = \{dx_{i,j}, dy_{i,j}\}$, where $dy_{i,j}$ and $dx_{i,j}$ represent the horizontal edge mask and the vertical edge mask, respectively. In terms of direction and amplitude, the edge vector representation is made and it is given as

$$Amp_{i,j} = |dx_{i,j}| + |dy_{i,j}|. \quad (13)$$

The angle, corresponding to the pixel edge direction, is given as

$$Ang_{i,j} = \frac{180^\circ}{\pi} \times \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right), \quad (14)$$

where $\frac{dy_{i,j}}{dx_{i,j}}$ indicates the pixel's edge direction. An image consists of pixels with edge vector that includes edge direction as well as amplitude, which together form the edge map of the image.

The comparison of edge information among the original and the distorted image blocks is done using the edge histogram. The steps involved in obtaining the edge direction histogram are (i) estimating the edge direction as well as the amplitude of each pixel using Eqs. (13) and (14); (ii) determining the direction of each pixel that is related to one of the eight discrete directions quantitatively; and (iii) adding the amplitude of the edge of a pixel with the direction of the same pixel.

Let D_y and D_x indicate the block edge direction vector of the distorted image and the original image, respectively. Using the correlation coefficient of D_y and D_x , the edge comparison $e(x, y)$ can be estimated and it is given as

$$e(x, y) = \frac{\sigma'_{xy} + C_3}{\sigma'_x \sigma'_y + C_3}, \quad (15)$$

where C_3 and σ'_{xy} represent a small constant that prevents the denominator from becoming zero and the covariance of vector D_y and D_x , respectively; σ'_y and σ'_x represent the standard deviation of D_y and D_x , respectively. The ESSIM is written as

$$ESSIM(x, y) = [l(x, y)]^\alpha \cdot [c(x, y)]^\beta \cdot [e(x, y)]^\gamma. \quad (16)$$

The mean of all the subimages of ESSIM is used to calculate the similarity of the full image, and it is written as

$$MESSIM(X, Y) = \frac{1}{M} \sum_{j=1}^M ESSIM(x_j, y_j). \quad (17)$$

- (2) Cumulative quality assessment: The cumulative quality assessment is the proposed weightage-based quality assessment that includes both the PSNR and ESSIM [i.e. MESSIM of Eq. (17)], as given in Eq. (18):

$$Cq = w_1 PSNR + w_2 ESSIM, \quad (18)$$

where Cq is the cumulative quality assessment that is determined using the enhanced image of the training database.

4.2 Construction of the Training Library

Let T_{Mn} be the training library, which is to be generated for carrying out learning in the DBN [1, 15, 20], where m refers to the volume of the training records and n refers to the number of features (Figure 3).

$$T_{Mn} = [\text{Entropy autocorrelation skewness}]. \quad (19)$$

The construction of the training library is shown in Figure 1.

$$w^* = \arg_w^{\max} (F(w) | T_{Mn}), \quad (20)$$

$$F(w) = w_1 \text{MESSIM} + w_2 20 \log_{10} \left(\frac{\text{MAX}_I}{\sqrt{\text{MSE}}} \right), \quad (21)$$

where w_2, w_1 refer to the weightage; w and $F(w)$ refer to the network weight and the objective function, respectively. Equation (21) is the cumulative quality assessment metric that is equivalent to Eq. (18).

4.3 Deep Belief Learning

(1) DBN Model: Consider the input as x and the hidden variables in layer i as g^i with a joint distribution of

$$P(x, g^1, g^2, \dots, g^l) = P(x | g^1) P(g^1 | g^2) \dots P(g^{l-2} | g^{l-1}) P(g^{l-1} | g^l), \quad (22)$$

where the conditional layers $P(g^i | g^{i+1})$ represent the factorized conditional distributions, wherein probability and sampling are not difficult. If the hidden layer g^i has n_i as the binary random vector with elements g_j^i , then

$$P(g^i | g^{i+1}) = \prod_{j=1}^{n_i} P(g_j^i | g^{i+1}): P(g_j^i = 1 | g^{i+1}), \quad (23)$$

$$P(g^i | g^{i+1}) = \text{sign} \left(b_j^i + \sum_{k=1}^{n_{i+1}} W_{kj}^i g_k^{i+1} \right), \quad (24)$$

Figure 3: Construction of the Training Library.

where $\text{sign}(\bullet)$ is a non-linear function, usually referred to as $\text{sign}(t) = \frac{1}{1+e^{-t}}$; b_j^i and W^i represent the biases for the j^{th} unit of the i^{th} layer and the weight matrix for the i^{th} layer, respectively. Equation (23) is followed by the generative model of the first layer $P(x|g^1)$, even though $g^0 = x$.

- (2) Restricted Boltzmann machine (RBM): The RBM is $P(g^{l+1}|g^l)$, which lies between two layers l and $l-1$. Consider v and h as the input layer activation and the hidden layer activation for a generic RBM. The joint distribution is written as

$$P(v, h) = \frac{1}{Z} e^{h'Wv + b'v + c'h}, \quad (25)$$

where c , b , Z , and W represent the hidden unit bias vector, visible unit bias vector, normalization constant, and weight matrix, respectively. If the argument of the exponential is notated in minus, then it is called the energy function and it is given as

$$\text{energy}(v, h) = -h'Wv - b'v - c'h. \quad (26)$$

In the above equation, the RBM parameters are indicated with $\theta = (W, b, c)$ and the layer-to-layer conditional distributions are denoted as $P(v|h)$ and $Q(h|v)$. Similar to Eq. (15), factorization occurs in the layer-to-layer conditionals that are associated with the RBM. Therefore,

$$P(v_k = 1|h) = \text{sigm}(b_k + \sum_j W_{jk} v_j), \quad (27)$$

$$Q(h_j = 1|v) = \text{sigm}(c_j + \sum_k W_{jk} v_k). \quad (28)$$

- (3) Gradient learning in RBM: Consider the Gibbs Markov chain for the pair of variables – hidden and visible – to determine the RBM's estimator of the log likelihood gradient. Gibbs sampling is performed, and then sampling of h given v and v given h is done.

In the Markov chain, the t^{th} v sample is indicated as v_t , and it starts at $t=0$ with v_0 as the input observation. Hence, from joint $P(v, h)$, the sample obtained is (v_k, h_k) for $k \rightarrow \infty$. With the RBM model, the log likelihood of a value v_0 is given as

$$\log P(v_0) = \log \sum_h P(v_0, h), \quad (29)$$

$$\log P(v_0) = \log \sum_h e^{-\text{energy}(v_0, h)} - \log \sum_{v, h} e^{-\text{energy}(v, h)}. \quad (30)$$

In the above equation, the gradient corresponding to $\theta = (W, b, c)$ is given as

$$\begin{aligned} \frac{\partial \log P(v_0)}{\partial \theta} &= - \sum_{h_0} Q(h_0|v_0) \frac{\partial \text{energy}(v_0, h_0)}{\partial \theta} \\ &+ \sum_{v_k, h_k} P(v_k, h_k) \frac{\partial \text{energy}(v_k, h_k)}{\partial \theta}, \end{aligned} \quad (31)$$

for $k \rightarrow \infty$.

The unbiased sample is given as $-\frac{\partial \text{energy}(v_0, h_0)}{\partial \theta} + E_{h_k} \left[\frac{\partial \text{energy}(v_k, h_k)}{\partial \theta} | v_k \right]$, where h_0 and (v_k, h_k) represent the sample from $Q(h_0, v_0)$ and Markov chains, respectively.

4.4 DBN-Aided Enhancement

Given a test US image with pelvic lesions, along with weightage vectors and the noise variance, the DBN estimates the decomposing model of the homomorphic wavelet filter through which the enhancement is to be done. The weightage vectors are defined based on the significance to be given for the quality assessment metrics.

Though homomorphic wavelet filter enhances the image, DBN plays the primary role in selecting the decomposing mode for the wavelet filter. As a result, enhancement is obvious over the other decomposing modes.

5 Experimental Results

5.1 Experiments

The experimental investigation is carried out in the MATLAB platform, and the performance of the proposed as well as the conventional methods is studied. The pelvic images of three patients – case 1 with 4 images, case 2 with 10 images, and case 3 with 2 images – are collected from the database (<http://www.ultrasound-cases.info/case-list.aspx?cat=26>). As the technique is related to supervised learning enhancement, the data are trained by 50% of the images. The trained data are subjected to analysis of visual quality. The results of the developed method are compared with NN-based enhancement methods. The performance assessment is done with two types of metrics, namely reference-based quality metric (PSNR) and non-reference-based quality metric [known as the second-order derivative metric (SDME)] [35]. The results that are related to the metric measures are analyzed and discussed further.

5.2 Quality of Enhancement

The quality of the selected five images has been studied with varied variance – 10%, 20%, 30%, 40%, and 50% – by comparing the noisy image and the enhanced image for the proposed method, and they are shown

Figure 4: US Images Contaminated by 10% Speckle Noise (A–E) and the Enhanced Images (F–J).

Figure 5: US Images Contaminated by 20% Speckle Noise (A–E) and the Enhanced Images (F–J).

in Figures 4–8, respectively. To study the quality of enhancement, the noisy images are contaminated by 10%, 20%, 30%, 40%, and 50% of noises.

While analyzing the enhanced image, it is found that the enhanced images exhibit improved visualization for the proposed method than the existing method.

5.3 Reference-Based Quality Assessment

The performance of the developed method, with respect to the image data collected, is estimated using the PSNR metric for the reference-based quality assessment and it is tabulated in Tables 2–4, with respect to the three sets of data. From the results obtained, it is noticed that there is no significant difference in the

Figure 6: US Images Contaminated by 30% Speckle Noise (A–E) and the Enhanced Images (F–J).

Figure 7: US Images Contaminated by 40% Speckle Noise (A–E) and the Enhanced Images (F–J).

Figure 8: US Images Contaminated by 50% Speckle Noise (A–E) and the Enhanced Images (F–J).

Table 2: PSNR Analysis between the Proposed Enhancement and the NN-Based Enhancement for Patient 1.

Variance	Image 1		Image 2	
	DBN	NN	DBN	NN
0.01	60.8	60.8	60.7	60.7
0.02	61	61	60.8	60.8
0.03	61.1	61.1	60.9	60.8
0.04	61.2	61.2	60.9	60.9
0.05	61.1	61.1	61	61

Table 3: PSNR Analysis between the Proposed Enhancement and the NN-based Enhancement for Patient 3.

Variance	Image 1	
	DBN	NN
0.01	61.7	61.7
0.02	61.9	61.8
0.03	62	62
0.04	62.1	62
0.05	62.2	62.2

PSNR of image 2 with noise variance of 0.01 and 0.02. Specifically, the proposed method shows better quality enhancement.

5.4 Non-Reference-Based Quality Assessment

The performance of both the proposed and the existing methods, with respect to the image data collected, is estimated with the SDME metric with varied variance for the non-reference-based quality assessment, and it is tabulated in Tables 5–7, with respect to the three sets of data. The proposed method

Table 4: PSNR Analysis between the Proposed Enhancement and the NN-Based Enhancement for Patient 2.

Variance	Image 1		Image 2		Image 3		Image 4		Image 5	
	DBN	NN	DBN	NN	DBN	NN	DBN	NN	DBN	NN
0.01	61	61	60.6	60.5	60.7	60.7	60.7	60.7	60.8	60.7
0.02	61.2	61.2	60.7	60.6	60.8	60.8	60.9	60.9	60.9	60.9
0.03	61.3	61.3	60.7	60.7	60.9	60.9	60.9	60.9	61	61
0.04	61.4	61.3	60.8	60.8	61	61	61	61	61.1	61.1
0.05	61.5	61.4	60.8	60.8	61.1	61	61.1	61.1	61.1	61.1

Table 5: Edge-Based Quality Analysis between the Proposed Enhancement and the NN-Based Enhancement for Patient 1.

Variance	Image 1		Image 2	
	DBN	NN	DBN	NN
0.01	−0.7	−0.7	−0.61	−0.61
0.02	−0.71	−0.71	−0.59	−0.59
0.03	−0.71	−0.71	−0.59	−0.59
0.04	−0.73	−0.73	−0.58	−0.58
0.05	−0.71	−0.71	−0.62	−0.62

Table 6: Edge-Based Quality Analysis between the Proposed Enhancement and the NN-Based Enhancement for Patient 3.

Variance	Image 1	
	DBN	NN
0.01	−0.6	−0.59
0.02	−0.59	−0.58
0.03	−0.58	−0.59
0.04	−0.61	−0.59
0.05	−0.58	−0.58

Table 7: Edge-Based Quality Analysis between the Proposed Enhancement and the NN-Based Enhancement for Patient 2.

Variance	Image 1		Image 2		Image 3		Image 4		Image 5	
	DBN	NN	DBN	NN	DBN	NN	DBN	NN	DBN	NN
0.01	−0.77	−0.78	−0.64	−0.66	−0.33	−0.34	−0.65	−0.65	−0.64	−0.66
0.02	−0.79	−0.77	−0.62	−0.64	−0.33	−0.34	−0.66	−0.66	−0.64	−0.67
0.03	−0.78	−0.78	−0.62	−0.64	−0.33	−0.34	−0.65	−0.65	−0.62	−0.64
0.04	−0.79	−0.78	−0.62	−0.64	−0.32	−0.32	−0.64	−0.64	−0.62	−0.62
0.05	−0.76	−0.76	−0.63	−0.64	−0.33	−0.32	−0.62	−0.62	−0.63	−0.64

shows increased performance than the NN-based method in all sets of images. The non-reference-based quality assessment is more practical because there will not be an original image in real-time while investigating the quality of the enhanced image. Under such circumstance, it is essential to quantify the quality of enhancement on the acquired images. Even under non-reference-based quality assessment, the proposed algorithm performs better, which asserts the practical significance of the proposed enhancement algorithm.

6 Conclusion and Future Work

This paper has addressed the challenges that are posed, while enhancing the image from noise as well as the misalignment problem. A methodology based on supervised learning has been proposed to counteract these challenges. The adopted methodology has been developed in MATLAB, and the performance has been investigated using renowned metrics such as PSNR and SDME. Patient-wise US images have been categorized as datasets, and experimentation has been carried out. Comparative analysis has been made against the conventional enhancement mechanisms, like NNs, to evaluate the performance of the proposed enhancement methodology. From the observations, it has been confirmed that the proposed method shows better image quality enhancement. As the benchmark images for pelvic lesions are found to be less available, the experimentations are limited to few openly available images. Similar experiments will be extended to a huge number of images that are to be acquired from local diagnostic centers.

Bibliography

- [1] M. Abdel-Zaher and A. M. Eldeib, Breast cancer classification using deep belief networks, *Expert Syst. Appl.* **46** (2016), 139–144.
- [2] A. Alush, H. Greenspan and J. Goldberger, Automated and interactive lesion detection and segmentation in uterine cervix images, *IEEE Trans. Med. Imaging* **29** (2010), 488–501.

- [3] M. E. Badawy, D. G. E. Elkholy, M. F. Sherif and M. A. E. Hefedah, Magnetic resonance imaging for diagnosis of pelvic lesions associated with female infertility, *Mid. East Fertil. Soc. J.* **20** (2015), 165–175.
- [4] C. Balas, A novel optical imaging method for the early detection, quantitative grading, and mapping of cancerous and precancerous lesions of cervix, *IEEE Trans. Biomed. Eng.* **48** (2001), 96–104.
- [5] K. A. Brocker, C. D. Alt, G. Gebauer, C. Sohn and P. Hallscheidt, Magnetic resonance imaging of cervical carcinoma using an endorectal surface coil, *Eur. J. Radiol.* **83** (2014), 1030–1035.
- [6] G. H. Chen, C. L. Yang, L. M. Po and S. L. Xie, Edge-based structural similarity for image quality assessment, in: *2006 IEEE International Conference on Acoustics, Speech and Signal Processing, 2006, ICASSP 2006 Proceedings*, vol. 2, France, 2006.
- [7] G. H. Chen, C. L. Yang, L. M. Po and S. L. Xie, Edge-based structural similarity for image quality assessment, in: *Proceedings of the IEEE International Conference on Acoustics Speech and Signal Processing*, 2006.
- [8] R. Conforti, A. M. Porto, R. Capasso, M. Cirillo, G. Fontanella, A. Salzano, M. Fabrazzo and S. Cappabianca, Magnetic resonance imaging of a transient splenic lesion of the corpus callosum resolved within a week, *Radiography* **22** (2016), 97–99.
- [9] A. Das, A. Kar and D. Bhattacharyya, Detection of abnormal regions of precancerous lesions in digitised uterine cervix images, in: *Proceedings of the International Electrical Engineering Congress*, 2014.
- [10] E. Epstein, A. Testa, A. Gaurilcikas, A. D. Legge, L. Ameye, V. Atstupenaite, A. L. Valentini, B. Gui, N. O. Wallengren, S. Pudaric, Cizauskas A, Måsbäck A, Zannoni GF, Kannisto P, Zikan M, Pinkavova I, Burgetova A, Dundr P, Nemejcova K, Cibula D, Fischerova D, Early-stage cervical cancer: tumor delineation by magnetic resonance imaging and ultrasound – a European multicenter trial, *Gynecol. Oncol.* **128** (2013), 449–453.
- [11] H. J. Gallowitsch, E. Kresnik, J. Gasser, G. Kumrig, I. Igerc, P. Mikosch and P. Lind, F-18 fluorodeoxyglucose positron emission tomography in the diagnosis of tumour recurrence and metastases in the follow-up of patients with breast carcinoma; a comparison to conventional imaging, *Invest. Radiol.* **38** (2003), 250–256.
- [12] R. M. Haralick, K. Shanmugam and I. Dinstein, Textural features of image classification, *IEEE* **3** (1973), 610–621.
- [13] R. Harmouche, N. K. Subbanna, D. L. Collins, D. L. Arnold and T. Arbel, Probabilistic multiple sclerosis lesion classification based on modelling regional intensity variability and local neighbourhood information, *IEEE Trans. Biomed. Eng.* **62** (2015), 1281–1292.
- [14] M. Jafar, S. Giles, V. Morgan, M. Schmidt, M. Leach and N. M. D. Souza, Evaluation of distortion correction of diffusion-weighted MR images of human cervix, in: *9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 514–517, 2012.
- [15] H. Jang, S. M. Plis, V. D. Calhoun and J. H. Lee, Task-specific feature extraction and classification of fMRI volumes using a deep neural network initialized with a deep belief network: evaluation using sensorimotor tasks, *NeuroImage* **145** (2017), 314–328.
- [16] Q. Ji, J. Engel and E. Craine, Texture analysis for classification of cervix lesions, *IEEE Trans. Med. Imaging* **19** (2000), 1144–1149.
- [17] N. Kurimoto, M. Murayama, S. Yoshioka and T. Nishisaka, Analysis of the internal structure of peripheral pulmonary lesions using endobronchial ultrasonography, *Chest* **122** (2002), 1887–1894.
- [18] J. H. Lee and C. H. Won, The tactile sensation imaging system for embedded lesion characterization, *IEEE J. Biomed. Health Inform.* **17** (2013), 452–458.
- [19] W. Li and A. Poirson, Detection and characterization of abnormal vascular patterns in automated cervical image analysis, *Adv. Vis. Comput.* **4292** (2006), 627–636.
- [20] H. Li, X. Li, M. Ramanathan and A. Zhang, Identifying informative risk factors and predicting bone disease progression via deep belief networks, *Methods* **69** (2014), 257–265.
- [21] A. Mahajan, R. Engineer, S. Chopra, U. Mahanshetty, S. L. Juvekar, S. K. Shrivastava, N. Deseekar and M. H. Thakur, Role of 3T multiparametric-MRI with BOLD hypoxia imaging for diagnosis and post therapy response evaluation of postoperative recurrent cervical cancers, *Eur. J. Radiol. Open* **3** (2016), 22–30.
- [22] H. G. C. Mesa, N. R. Erez and R. H. Jimenez, Aceto-white temporal pattern classification using k-NN to identify precancerous cervical lesion in colposcopic images, *Comput. Biol. Med.* **39** (2009), 778–784.
- [23] A. Milbourne, S. Y. Park, J. L. Benedet, D. Miller, T. Ehlen, H. Rhodes, A. Malpica, J. Matisic, D. Van Niekirk and E. N. Atkinson, Results of a pilot study of multispectral digital colposcopy for the in vivo detection of cervical intraepithelial neoplasia, *Gynecol. Oncol.* **99** (2005), 67–75.
- [24] M. L. Palmeri, H. Feltovich, A. D. Homyk, L. C. Carlson and T. J. Hall, Evaluating the feasibility of acoustic radiation force impulse shear wave elasticity imaging of the uterine cervix with an intracavity array: a simulation study, *IEEE Trans. Ultrason. Ferroelect. Freq. Control* **60** (2013), 2053–2064.
- [25] G. Papoutsoglou, T. M. Giakoumakis and C. Balas, Dynamic contrast enhanced optical imaging of cervix, in vivo: a paradigm for mapping neoplasia-related parameters, in: *35th Annual International Conference of the IEEE EMBS*, Osaka, Japan, July 2013.
- [26] S. Y. Park, A study on diagnostic image analysis for the detection of precancerous lesions using multispectral digital images, PhD Thesis, University of Texas, Austin, 2007 (thesis submitted to the University of Texas).
- [27] S. Y. Park, M. Follen, A. Milbourne, H. Rhodes, A. Malpica, N. MacKinnon, C. MacAulay, M. K. Markey and R. R. Kortum, Automated image analysis of digital colposcopy for the detection of cervical neoplasia, *J. Biomed. Optics* **13** (2008), 014029.

- [28] S. Y. Park, D. Sargent, R. Lieberman and U. Gustafsson, Domain-specific image analysis for cervical neoplasia detection based on conditional random fields, *IEEE Trans. Med. Imaging* **30** (2011), 867–878.
- [29] E. S. Pearson and H. O. Hartley, *Biometrika Tables for Statisticians*, 3rd ed., Cambridge University Press, Cambridge, 207, 1970.
- [30] A. A. Roma and E. D. Kelly, Reliability of PAX8 in clinical practice to accurately determine primary site of origin in female pelvic or abdominal lesions?, *Ann. Diagn. Pathol.* **18** (2014), 227–231.
- [31] C. Romagnolo, A. E. Leon, A. S. C. Fabricio, M. Tadorelli and J. Polesel, HE4, CA125 and risk of ovarian malignancy algorithm (ROMA) as diagnostic tools for ovarian cancer in patients with a pelvic mass: an Italian multicenter study, *Gynaecol. Oncol.* **141** (2016), 303–311.
- [32] L. K. Soh and C. Tsatsoulis, Texture analysis of SAR sea ice imagery using grey level co-occurrence matrices, *IEEE Trans. Geosci. Remote Sens.* **37** (1999), 780–795.
- [33] M. S. Umerani, A. Abbas, S. K. Bakhshi, U. M. Qasim and S. Sharif, Evolving brain lesions in the follow-up CT scans 12 hours after traumatic brain injury, *J. Acute Dis.* **5** (2016), 150–153.
- [34] T. C. Wright, Cervical cancer screening using visualization techniques, *J. Nat. Cancer Inst. Monogr.* **31** (2003), 66–71.
- [35] J. Xia, K. Panetta and S. Agaian, Color image enhancement algorithms based on the DCT domain, in: *Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 1496–1501, 2011.
- [36] J. Xin, Q. Ma, Q. Guo, H. Sun, S. Zhang, C. Liu and W. Zhai, PET/MRI with diagnostic MR sequences vs PET/CT in the detection of abdominal and pelvic cancer, *Eur. J. Radiol.* **85** (2016), 751–759.
- [37] A. Yavariabdi, A. Bartoli, C. Samir, M. Artigues and M. Canis, Mapping and characterizing endometrial implants by registering 2D transvaginal ultrasound to 3D pelvic magnetic resonance images, *Comput. Med. Imaging Graphics* **45** (2015), 11–25.
- [38] M. Zanotto, Visual description of skin lesions, PhD Thesis, University of Edinburgh, 2010 (thesis submitted to the University of Edinburgh).
- [39] T. Zhan, Y. Zhan, Z. Liu, L. Xiao and Z. Wei, Automatic method for white matter lesion segmentation based on T1-fluid-attenuated inversion recovery images, *IET Comput. Vis.* **9** (2015), 447–455.