Journal of Ambient Intelligence and Humanized Computing
https://doi.org/10.1007/s12652-020-02191-9

ORIGINAL RESEARCH

Computer aided innovation method for detection and classification of cervical cancer using ANFIS classifier

Sukumar Ponnusamy[1] · Ravi Samikannu[2] · Sampath Kumar Venkatachary[3] · Sharmila Sukumar[4] · Rohini Ravi[5]

Received: 18 February 2020 / Accepted: 6 June 2020
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

Abstract
Early detection of cervical tumour is very important to minimise deaths due to cervical cancer. Further it provides a deep insight into the anatomical information of the normal and abnormal cervix and helps in planning for a good treatment well in advance. Numerous techniques are used to detect malignancy through image segmentation. One such segmentation technique is discussed here. The proposed technique uses Artificial Neural Network Fuzzy Inference system (ANFIS) and watershed segmentation techniques for image classification and processing and compares the results with known techniques. A comprehensive set of fuzzy rules was used in the experiment to classify abnormal images to the corresponding malignancy. The experiment shows that the proposed technique is feasible and provides greater accuracy in detection of tumour types.

Keywords Cervical cancer · Image classification · Tumour · Fuzzy rules · Neural network

1 Introduction

Cervical malignancy is the abnormal growth of cells inthe cervix and can invade other tissues and organs in the body. It is one of the main causes of death in middle-aged ladies across the globe especially in the developing countries which stands at about 80%. However, it is important to understand that, as cervical cancer is slow-progressing, its precancerous sores due to slow progression, provide a great opportunity for prevention and is totally preventable provided the sores are detected early and treated in early stages. As per the World Health Organisation (WHO) 2002 report on cervical cancer, cervical malignancy is the most dangerous disease which affect most middle aged women. Approximating 12% of all cancers in women which affects middle aged women. Accordingly, WHO reports a detection rate of 471,000 new cases in 2001 and an estimated 288,000 die from it every year and was noticed especially between ages 15–44 predominantly.

Cervical cancer is predominant 4th place in all forms of genital cancer in among all the forms of female cancers. Cervical cancer is the 6th leading reason of deaths for middle aged women in the United States. Of the total number is reported or diagnosed cancer cases in developing countries, 75% account for cervical malignancy (Jang 1993). This is primarily due to the fact that most developing countries have no effective detective or screening and preventive

✉ Sukumar Ponnusamy
sukumarwin@gmail.com

Ravi Samikannu
ravis@biust.ac.bw

Sampath Kumar Venkatachary
sampathkumaris123@gmail.com

Sharmila Sukumar
kssharmi@gmail.com

Rohini Ravi
rohinir84@gmail.com

[1] Department of Electronics and Communication Engineering, Nandha Engineering College, Erode, Tamil Nadu, India

[2] Faculty of Electrical Computer and Telecommunications Engineering, Botswana International University of Science and Technology, Palapye, Botswana

[3] Information Technology, Grant Thornton, Gaborone, Botswana

[4] Department of Civil Engineering, Kongu Engineering College, Erode, Tamilnadu, India

[5] Department of Computer Science and Engineering, Vivekanandha College of Engineering for Women, Elayampalayam, Tiruchengode, India

Published online: 15 June 2020

mechanisms either due to cost factors or lack of acceptability. Cervical cancer normally begins with changes in cervix due to abnormalities.The Human Papilloma Virus (HPV) is the primary factor which causes cervical cancer. The HPV virus is transmitted because of unprotective sex, early sexual contact, multiple sex partners while other factors like oral contraceptives to prevent birth which tend to increase the risk and thereby lead to greater exposure. In most cases, the root cause can be attributed to males who carry the virus and transmit it, and in most cases many still are unaware of the risks it poses.

Cervical Intraepithelial Neoplasia (CIN) also known as cervical dysplasia (Roland et al. 2013), the Squamous Intraepithelial Lesion (SIL) is the abnormal growth of cells on the surface areas and are referred to as precancerous transformation for the cells of the cervix. These precancerous cells later become malignant unless detected and treated at an early stage. These squamous cells on the surface of the cervix generally comprise of two different types of cells squamous epithelial cells and columnar epithelial cells which form the outer and the inner part of the cervix (ectocervix and endocervix) respectively.

Numerous methods are used for early detection and prevention of cervical cancer. Some of these include the Papanicolaou test (Pap Smear), biopsy of cervical tissue, vaccination against HPV etc. Of these pap smear test is the widely used method for detection. Some of the widespread tools include colposcopy, Loop Electrosurgical Exclusion Procedure (LEEP) etc. Most screening and image processing techniques used are mostly based on optical technologies and researchers have for decades tried to improve the optical image contrast of the cervix and precancerous lesions (Sarwar et al. 2016). Simple techniques with the use of acetic acid and Lugols iodine are popular methods to highlight suspicious regions in the cervix. However, it is to be noted that as the image specifics are low, they routinely are biopsied (Mitchell et al. 1998). As some of these technologies are expensive to afford in developing countries, there is an increased research focus on visual screening technologies. Image processing using visual screen technologies is a popular method to analyse the images. In image processing technologies an analogous photographic method is put to use. Figures 1 and 2 shows the images of normal and malignant cervix image. The small 35 mm camera used to photograph the uterine lining of cervix. A small ring flash is used additionally with the camera to provide the better illumination in the region.

Depend upon the individual patient's profile Cervix varies in size and shape. Due to the different factors in the process of the acquisition the viewing angle substantically varies depend on the images. For example, for obtaining a cervigram image, a complex high-resolution camera with a good flash and proper illumination is used, thus generating



Fig. 1 Normal cervical image



Fig. 2 Abnormal or malignant cervical image

several artefacts during the acquisition process. The inclusion of a strong flash strongly causes shading artefacts that result in inhomogeneous appearance on the tissues, while also resulting in specular reflection that plays an interference role during the automated analysis. The second part is the variability present within the cervigram image which is generally not constant. The third challenge is the variability of the content within the cervix image. This is primarily due the absence of tissue types present in the cervigram image. Lastly, the dynamic colour range which is narrow and the absence of non-existent boundary among the tissue region makes it difficult to analyse the image and carryout data analysis. With this background, this paper focuses on using a combination of image process algorithms, keeping in mind the expense part, while ensuring imaging modality and automating analysis.

The paper is organized as follows: Sect. 1 presents the introduction. Section 2 reviews various related works carried out in relation to detection mechanisms of cervical cancer.

Section 3 presents the proposed methods. Section 4 presents the data selection and discusses the image processing in detail. Section 5 discusses results and concludes the paper.

2 Digital imaging of cervix and automated analysis

Digital imaging technologies have been extensively researched to analyse cervical cancer detection. The image acquisition process performs a complex task with the use of a strong camera flash to provide the required illumination of the cervix. Within and across the tissues the inhomogeneous appearance is shaded with a dye is the process of cancer detection. The use of Multispectral Digital Colposcope (MDC) using two-stage analysis with a success rate of 79% and specificity of 88% serves as an alternative method for early detection. Two stages are involved in this process. In the first stage, the similar optical patterns are clustered together, and later classification algorithms are used to identify the regions containing neoplastic tissues. The use of acetic acid in the experiment forces changes in the intensity of the reflected light which changes from green to red and the same is used to differentiate High Grade Squamous Intraepithelial Lesion (HGSIL) and malignant cells from normal low-grade squamous intraepithelial tissue (Thekkek and Richards-Kortum 2008). The data-driven image processing technologies for detection greatly increase the possibilities of early detection. In an experiment, Song et al. successfully obtained 74% sensitivity and 90% specificity in differentiating HGSIL with Low Grade Squamous Intraepithelial Lesion (LGSIL). While in comparison it was noticed that pap smear tests resulted in a marginal success rate of only 37% with a specificity of 96% while the HPV tests provided better results with 57% and 93% specificity. The modified algorithm based on multimodal entity coreference also provided better results in the multiple tests based on gradient-based approach which resulted in 83.21% and 94.79% sensitivity and specificity (Park et al. 2008). 2D imaging analysis, a cost-effective solution proposed by Park et al. provided more insight into the cancerous lesions based on domain-specific diagnostic features and was validated based on 48 patients clinical data demonstrating the performance colposcopic annotations against histopathology (Song et al. 2015).

Cervigrams lesions are difficult to analyse due to the complex patterns. The complexity of training physicians in pathology recognition patterns further complicated the imaging process. This led to a modified pattern recognition algorithm. The complexity of the pattern analysis was further modified, and pattern recognition provided more accurate analysis for accurately recognising cervical textures from colposcopic images as perceived by a human (Park et al. 2011). This complexity and the extreme data volume in pattern recognition alsorequiresextremely complex algorithms like Gaussian Mixture Model (GMM) modelling (Ji et al. 2000). Though GMM modelling did address some of the issues and aided in understanding better on cervical cancer, the demonstration of genetic factors had hardly been studied.The demonstration on the importance of hereditary genetic factors influencing cervical cancer also provides an insight into analysing cervical cancer patterns. Horng et al. (2004) employ a Bayesian network method to analyse and compare the hereditary genetic component with four different decision tree algorithms to understand single-nucleotide polymorphism (SNP) and microsatellites (Varadharajan et al. 2018). The results of PART provided an accuracy of 59.73% accuracy with a specificity of 79.55% and an F-measure of 60.11% (Gordon et al. 2004). Greenspan et al. (2008) highlight the importance of multistage analysis in cervical malignancy and propose a model for multistage segmenting the anatomical research within the boundaries of the cervigrams. They extract the cervix region along with image of cervix boundary and eliminate the specular reflection in the preprocessing step. They conclude that the segmentation results on evaluation dice 88% and average sensitivity of 90% (Kumar et al. 2012). In understanding more on the cervical boundary, class specific region and boundaries with the aim to detect the lesions in the uterine cervix images studied with the help of automatic extraction and segmentation (Alush et al. 2010). They apply watershed segmentation using a modified Markov Random Field (MRF) algorithm.

The Markov randomused for field demonstrate the watershed segmentation map. The watershed aread is compared with the binary random factors. The anlayze shows the area is a piece of the sort tissue or not. The neighbourhood factors on the curves of the watershed map show whether the arc is a piece of the object boundary. The elements depend on directed learning of a visual word appropriation. The last sore area segmentation is gotten utilizing a loopy belief propagation applied to the watershed arc-level MRF. Trial results on genuine information show cutting edge segmentation results on this difficult errand that, if fundamental, can be intuitively improved. The results that were obtained were at 65% average dice with a 73% sensitivity and a false-positive rate of 6%. Against this backdrop, this paper aims to enhance further the scope of a digital image using ANFIS.

3 Proposed methodology and analysis

The proposed technique utilizes the image of the cervix to recognize the malignant growth affected areas by a trained classifier dependent on the features extracted from the cervical image (Herrero et al. 1997). Figure 3 delineates the-proposed methodology. The Guanacaste database contains

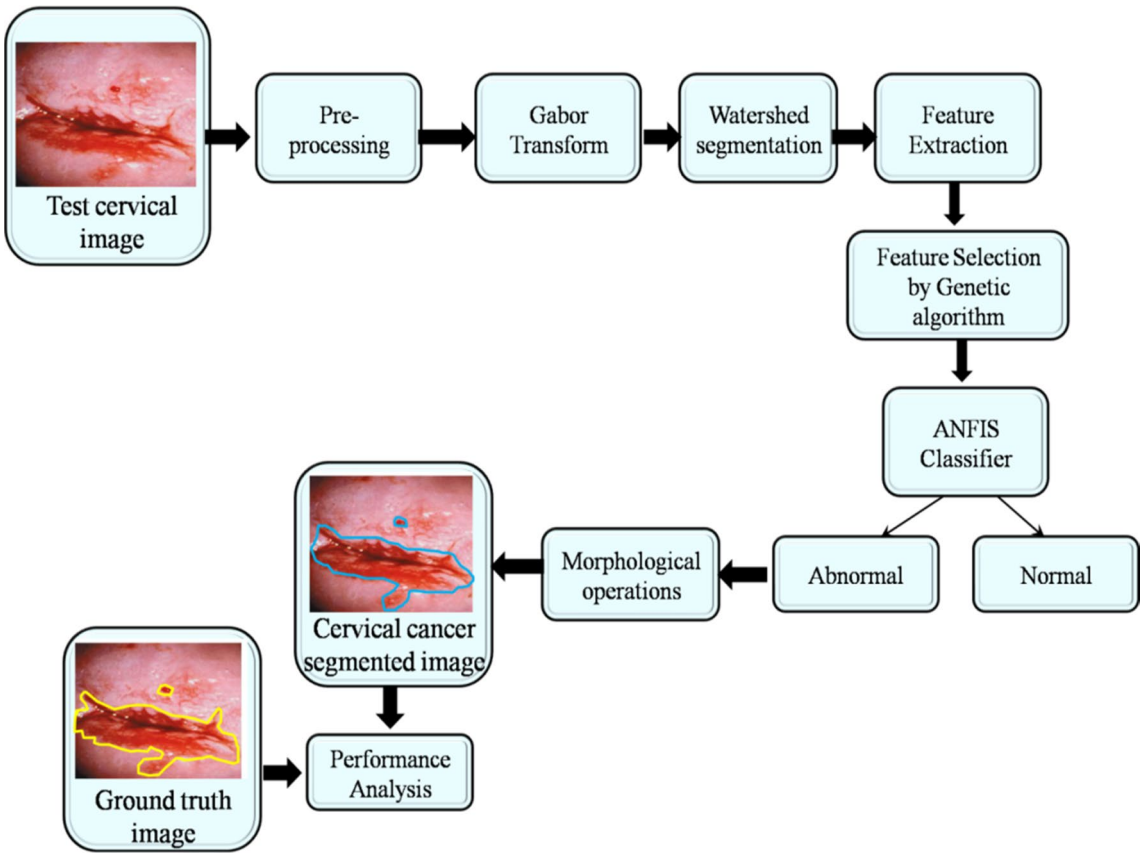


Fig. 3 Proposed flow of cervical cancer detection system

different cervical images of various patients at different stages for the preparation phase. At the underlying stage the images are preprocessed and then it is upgraded. The noise factors are removed in the preprocessing stage and the image quality is improved in a better level. The Gabor transform and the watershed algorithms applied in the enhanced image. The transform and algorithms extract the features of the image. Extracted images given to the ANFIS classfierand used in training mode for both normal, abnormal cervix images. The performance evaluation is obtained by comparing the actual images with abnormal and normal images (Nallathambiet al. 2017).

3.1 Dataset selection and pre-processing

The dataset with a size of 280 randomly selected samples included high quality 2891 × 1973 pixel cervix images for the study. The sample comprised women who participated in the National Cancer Institute's (NCI) Guanacaste study. The data set sample were classified as patients with malignant growth, patients without cervical sore at the time of enlistment; however, later the sores developed and healthy women with intact cervix. Out of the 280 randomly selected samples, an unequal number of cases had to be picked for analysis due to the availability of a limited number of cancer cases in the entire dataset. As the data needed to arranged in binary classification, i.e., arranging patient cases into one of the accompanying two classifications: < CIN2 (mild cases) and CIN 2/3+ (severe cases), so that in the target classification there exists same number of patient cases in these two classes: 140 cases in each (CIN2, and CIN 2/3+).

The selected images are converted to grayscale image using a matrix of numerical values of 0 and 255 and then pre-processed for reducing noise to enhancing the image quality. Zero represents black pixels, and 255 represents white pixels. The background from the processed image is then extracted to improve finer details for analysis. The image is then resized and reduced to 256 × 256 pixels.

3.2 Gabor transform

Subsequent to the image processing, a Gabor filter is utilized for enhancing the quality of the image and smoothening the edges. The frequency portrayals of Gabor channels are like that of the human visual grouping and appropriately suite the texture representation. In the spatial space, a 2D Gabor filter channel is a Gaussian kernel function modulated or adjusted by a sinusoidal plane wave. In Gabor filer channels,

all filer channels can be created from one mother wavelet by dilation and henceforth, they are similar. The Gabor filters, therefore, permit designing numerous dilation and rotations exhibiting the desirable characteristics directly related to Gabor wavelets. These characteristics of the spatial locality are optimally localised in the frequency domains and successfully used in image processing. The Gabor Kernels used are defined as follows:

$$\varphi_{p,q} = \frac{k_{p,q}^2}{\sigma^2} exp\left(-\frac{k_{p,q}^2 z^2}{2\sigma^2}\right) \times \left[exp\left(ik_{p,q}^2 z^2\right) - exp\left(-\frac{\vartheta^2}{2}\right)\right] \tag{1}$$

The Gabor kernels orientation and scale is given by σ and υ. The wave vector is $k_{p,q}$. The Gabor filter used to remive the noise from the cervical image and it is transformed into frequency domain image.

3.3 Watershed segmentation

Watershed segmentation is a popular morphological device that is known to reduce the variabilities in the negative images when applied to greyscale images. It is also known to reduce computational costs substantially. The watershed change is applied and the cervical image is segmented into superpixels. The change can be viewed as the flooding of a topographic surface (the data input), which loosens up from its minima while waters from different sources are kept from consolidating. The indisputable result is a ton of portions of superpixels, secluded by peak lines, or watershed twists. The superpixels have sensible, reasonable highlights, and their breaking points are inverse or opposite to the image inclinations. The advantage of the watershed calculation is that the huge points of confinement of dangerous districts are caught in the watershed edge map and these edges don't cover with different tissues.

3.4 Feature extraction

The goal of feature extraction is to extricate unique features from the selected images to separate it from the input pattern while diminishing the first initial dataset by removing certain features. The features separated thus, act as input to the classifier, portraying the important properties of the image. The features isolated, therefore, present the qualities of the info example to the classifier in the wake of depicting the significant properties of the picture into a component space.

3.5 Gray Level Co-occurrence Matrices (GLCM) features

Haralick extracted the features and properties from GLCM. The matrix symmetric is obtained by adding the transpose the GLCM with the original GLCM. The sum of the GLCM elements are calculated and the element in the matrix is divided by the total sum. The normalized matrix can be calculated by dividing the matrix by the total sum.

A notable procedure to extricate texture characterisation is based on the statistical method using gray level co-occurrence matrices (GLCMs) method. The GLCM contains the second-order statistical data of spatial relationship the pixels of an image. The GLCM functions based on the characterisation of the texture by computing how often the pixels with a specific value. The values of i and j shows the grey level quality values of the image.

In four directions 0°, 45°, 90° and 135° the GLCM can be obtained. Throught the GLCM the different four properties: (1) energy, (2) correlation, (3) homogeneity, (4) contrast can be obtained. The other name of the energy parameter is uniformity. Textural uniformity is measured by energy, (example pixel pair's repetitions), and the image page considered is homogeneous. The smoothness of the image is measured by energy. The region is smooth then P_{ij} is uniformly distributed, and the lower is the value of the Angular Second Moment (ASM):

$$Energy = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} {P_{ij}}^2 \tag{2}$$

The ijth entry of the normalized co-occurrence matrix is given by P_{IJ}. The video number of gravy levels is given in N_g.The correlation used to find the grey tone linear dependencies in the image. Direction under investigation is same like the vector displacement. When the correlation values reaches 1 achieves linear relationship with respect to the gravy levels of the pixel pairs. The high correlation can be measured at low and high energy situations:

$$Correlation = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} i.j.{P_{ij}}^2 - \mu_x \mu_y}{\sigma_x \sigma_y} \tag{3}$$

Means and standard deviations are $\mu_x, \mu_y, \sigma_x, \sigma_y$ and the marginal probabilities $P_x(i), P_y(i)$ are obtained by adding the row and columns of the matrix P_{ij}. Larger values for smaller grey tone differences in pair elements can be measured by Homogeneity parameter. It takes high values for low contrast images:

$$Homogeneity = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i-j)^2} P_{ij} \tag{4}$$

The contrast parameter is used to measure the spatial frequency of the image. The continuous pixels can be obtained by using the difference between the lowest and highest values:

$$Contrast = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i-j)^2 P_{ij} \tag{5}$$

Energy is an element that estimates the smoothness of the image. Correlation relationship is a proportion of linear dependencies of grey tones in the image. The homogeneity values in the image measured by inverse difference moment. Inverse difference moment other wise called as homogeneity. Grey levels can determine the size of the GLCM.

3.5.1 Gray level based features

The difference between the grey level intensity value in the sample and its surrounding pixel forms the feature known as a grey level based feature. In a cervical cancer profile, the malignant region is always darker when compared to the adjacent surroundings. The grey-level features in the surroundings, therefore, forms an excellent area for segmentation. A set of grey-level-based descriptors are derived from homogenized images I_H considering only a small pixel region centred on the described pixel (p, q). $S^w_{p,q}$ stands for the set of coordinates in a 3×3 sized square window centered at the point (p, q). The grey level based feature descriptors each represent a feature image and are represented as follows:

$$\left.\begin{aligned} &\text{a) } F_1(p,q) = I_H(p,q) - \min\{I_H(s,t)\} \\ &\text{b) } F_2(p,q) = \text{map}\{I_H(s,t)\} - I_H(p,q) \\ &F_3(p,q) = I_H(p,q) - \text{mean}\{I_H(s,t)\} \\ &\text{c) } F_4(p,q) = \text{std}\{I_H(s,t)\} \\ &\text{d) } F_5(p,q) = I_H(p,q) \end{aligned}\right\}$$

3.5.2 Future set

From the cervical image the different performance parameters mean. The performance parameters like mean, standard deviation and skewness are estimate from the cervical image, which helps in identifying the malignant areas and therefore, can be used as an input for training the classifiers. The features are defined as follows:

The average intensity values of the pixels are denoted by mean:

$$\text{Mean } S_M = \bar{b} = \sum_{b=0}^{L-1} bp(b) \tag{6}$$

The deviation (or) the variance between the pixels in the input image is represented by the standard deviation:

$$\text{Standard deviation } S_D = \sigma_b = \left[\sum_{b=0}^{L-1} \left(b - \bar{b}\right)^2 p(b)\right]^{1/2} \tag{7}$$

Skeweness is a measure of the asymmetry of the probability distribution of a real valued random variable. The skewenss value can be positive or negative:

$$\text{Skewness } S_S = \frac{1}{\sigma_b^3} \sum_{b=0}^{L-1} \left(b - \bar{b}\right)^3 p(b) \tag{8}$$

3.6 Genetic algorithm and feature selection

Genetic algorithm is a type of inductive learning platform and was first introduced as a concept and adaptation by (Holland 1975) based on the theory of natural selection by Charles Darwin and is largely based on a neural-genetic model similar to neurons in humans. The process starts with the natural selection of the fittest from the population and the crossover or mutations begin in the form of offspring inhering the parent characteristics and thus will be added next generation and so on. The process iterates itself and by the end of the iterative process,the fittest individual results. The genetic algorithm generally is a process of five stages, fitness function, feature selection, cross over, offspring, mutation and termination. The evaluation function or the fitness function determines the performance of the classifiers, and therefore it is essential to ensure a good or appropriate fitness function. Feature selection is key to a genetic algorithm. The primary reason is that a proper selection will ensure that there is minimum storage requirement, reducing the training sets thereby resulting in reduced computation time while also ensuring good facilitation of information perception understanding and perception. While it is essential to ensure proper feature selection, the cross over phase is the most significant as each parent from the population are crossed at a certain point. This crossover results in the offspring inheriting the characteristics of the parents with the best gene pool. The mutation is largely due to random changing of the individual components of the selected individual and acts as a perturbation operator while ensuring new genes are inserted in the formation of new population.

3.7 ANFIS classifier

ANFIS is a network-based fuzzy inference system that is based on fuzzy logic. The network is structured into two parts premise and consequence parts, that is the first layers are the input layers where the membership functions are determined (fuzzification layer), the second layer generates the firing strengths based on the parametric sets while the third normalizes the firing strengths and fourth uses

the normalized values as input for further functions. As the system is bi-directional adaptive network, it facilitates backpropagation. This means that the output is dependent on the parameters of the nodes and the learning rule specify how these parameters are tweaked to minimiseerrors. As explained, the features extracted from the cervical test images are then classified and trained with the ANFIS algorithm using MATLAB simulator. The images are analysed through the tool and then classified as normal and abnormal images.

3.8 Morphological operation based tumor segmentation

The morphological tool is used for enhancing and sharpening the regions while filling the gaps in the image. The morphological function basically includes dilation, erosion, opening and closing of the images. The process used here is only dilation and erosion. In the erosion process, every pixel that is fully or partly touching the background is changed into a background pixel while the breaking down the object to smaller bits and pieces of multiple objects. This mathematically can be expressed as follows:

Erosion:

$$(G \ominus H)(x,y) = {}_{(r,a)\in H}{}^{inf}\{x+r, y+s) - H(r,s)\} \tag{9}$$

The image is denoted by G(x,y). The function is denoted by H(x,y) the H is mentioned as structuring element. The four fundamental binary morphological operations is denoted by the symbols $\oplus$, $\ominus$, $\circ$, and $\bullet$.

While erosion breaks down, the dilation process merges the multiple small objects into one and can be represented as follows:

Dilation:

$$(G \oplus H)(x,y) = {}_{(r,a)\in H}{}^{sup}\{x-r, y-s) + H(r,s)\} \tag{10}$$

The inf{ } and sup { } denotes the infirmum and supremum operation.

Erosion and dilation when combined results in opening, by which objects that are adjacent are spaced, and objects that are joined are detached, and the holes within the objects are enlarged. An opening can be expressed as follows:

Opening:

$$G \circ H = (G \ominus H) \oplus H \tag{11}$$

Closing:

$$G \bullet H = (G \oplus H) \ominus H \tag{12}$$

On applying the morphological tool on the greyscale cervical image, the tool separates the tumour infected portion on the cervical image and encircles the region for further diagnosis. Finally, the tumour portion of the image, with the highest intensity becomes visible.

4 Data selection and image processing

In the current paper, two classification techniques (training and testing) were adapted using a sample of 280 cervical images. The sample, taken up for the study included 140 normal and 140 abnormal images. These images were used as a training data set. The data were then clustered by applying subtractive clustering using fuzzy inference component of ANFIS. Subsequently, the member functions and the rules (if–then) were generated from the ANFIS, both before and after the training.

Though there are several algorithms that can be used with feedforward networks, the Fuzzy Inference System (FIS) optimization tool was used in the study. The selection was basically due to the facilitation of least squares with backpropagation gradient descent algorithm and generation of multiple sets of solutions. The selected data or the test case images were assigned weights manually varying between 0 and 1 to integrate data and ensure image similarity and ensure accuracy. Of the same sample, 80 images were tried for testing using 30 normal and 50 abnormal images. The MATLAB simulator tool was then applied to understand and extract the features from the cervical images in the testing phase. The property features included normal, abnormal or malignant areas, entropy and homogeneity and matched for the best solution.

Figure 4 provides a detailed set of Guanacaste database images. Figure 4a is an input image for the experiment. Figure 4b is the image obtained from the radiologistwhich is highlighted in yellow, Fig. 4c is the results of cervical cancer segmented image by the proposed method, while Fig. 4d is the comparative image obtained using the conventional method.

The first process is to extract features such as GLCM from the selected images and then apply an algorithm for compiling the information and comparing the same with the radiologically obtained image.

Similarly, the selected image Fig. 4a is subjected to grey-level feature extraction. The selected extraction is then processed by applying the proposed algorithm. The results obtained thus are tabulated in Table 1.

The resultant images after applying through the proposed algorithm are shown in Fig. 4c. The images were then analysed and evaluated using the following parameters.

The performance is evaluated in Table 2, and the graphical representation is provided in figure:

$$\text{Sensitivity } \left[\text{Se} = \text{TP}/(\text{TP}+\text{FN})\right]$$

$$\text{Specificity } \left[\text{Sp} = \text{TN}/(\text{TN}+\text{FP})\right]$$

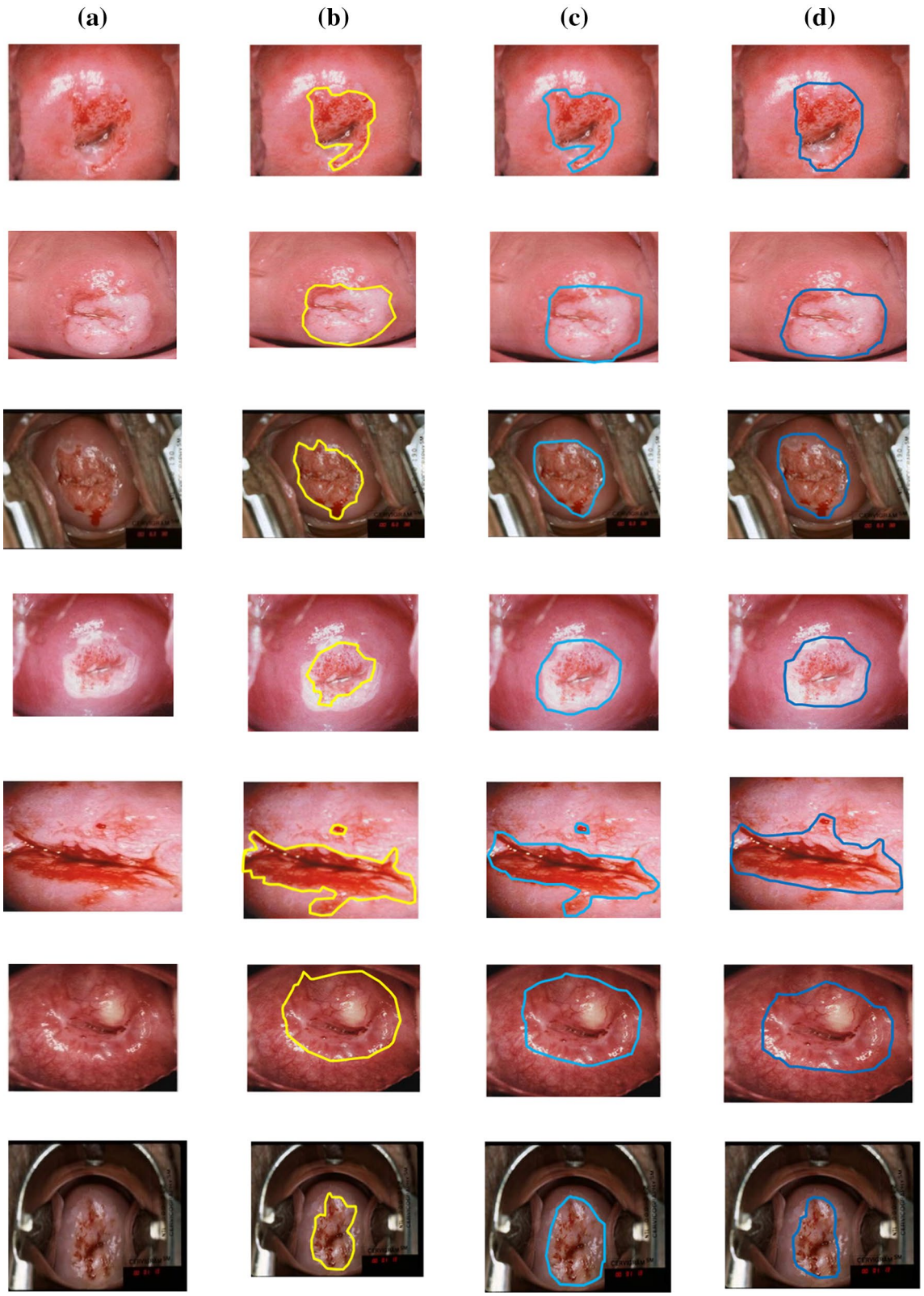


Fig. 4 The source imagse is shown in **a** column, source images shown in column (**b**), proposed method segmented cancer image in column (**c**), the conventional method segmented image in column (**d**)

Table 1 GLCM features

Gray level co-occurrence matrix features	Usual region	Unusual region
Energy	$1.55*10^{-5}$	$1.625*10^{-5}$
Entropy	1.1079	1.105
Autocorrelation	$1.618*10^{4}$	$1.594*10^{4}$
Contrast	$1.12*10^{4}$	$1.303*10^{4}$

$$\text{Accuracy } \left[\text{Acc} = (\text{TP} + \text{TN})/(\text{TP} + \text{FN} + \text{TN} + \text{FP})\right]$$

where TP is true positives and is the classification result of the abnormality inthe form of cancer pixels, TN is true negatives and is the classification result of a negative result. In absence of cancer pixels that are identified and classified by the system wrongly. FP is false positives and is the classification result of correctly. Identified non-cancer pixels. FN is false negatives and is the classification result of wrongly identifiedusual pixels. Se non affected cancer region and Sp cancerous region in the images. Acc denotes the ratio the spotted and categorized cervical regions.

The performacecomparision of the results for the proposed system is compared with other results and it is shown in Table 2.

The graph is plotted is shown in Fig. 5. The different parameters like sensitivity, specificity and accuracy is analyzed in the Table 2.

5 Results, discussion and conclusions

Cervical cancer is one of the most significant contributors to cancer-related deaths among women in a developing country. This paper discusses and compares the proposed method with two methods as demonstrated by Kim and Cheng (Chang et al. 2005; Kim and Huang 2013). Table 2 shows the improvement in the accuracy of the systems. As seen from Table 2, which analyses the performance, a sensitivity of over 92.1% is obtained as against the system proposed by Kim and Cheng,

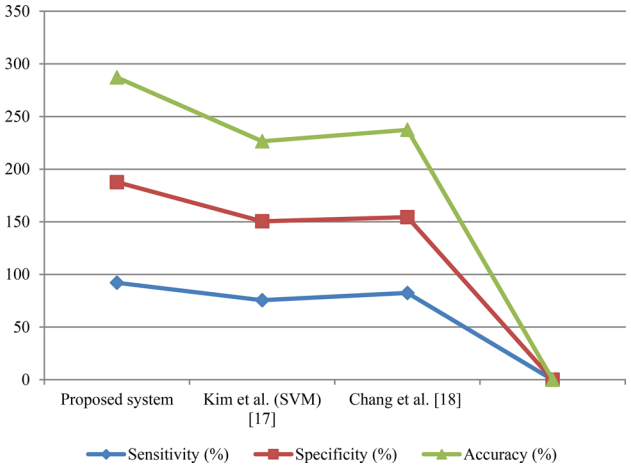


Fig. 5 Graphical representation and performance comparison of the proposed technique

while the specificity ratio was 96% as compared to 75% and 72% of Kim and cheng respectively.

The accuracy obtained by the proposed is an astounding 99.5% as against 76% and 83%. This indicates that the proposed method clearly aggregates the image similarities and significantly improves the system accuracy with fewer inputs. In conclusion, it can be said that the extracted features.The features extracted is trained by the ANFIS classifier. The 100% classification rate is achieved by the proposed method. The edges of the cancer region detection achieved by the morphological operations.The performance of the proposed system is verified using ground truth images.

Table 2 Performance comparisons of proposed method

Methodology	Sensitivity (%)	Specificity (%)	Accuracy (%)
Proposed system	92.1	95.6	99.5
Kim and Huang (2013) (SVM)	75.5	75	76
Chang et al. (2005)	82.39	72	83

References

Alush A, Greenspan H, Goldberger J (2010) Automated and interactive lesion detection and segmentation in uterine cervix images. IEEE Trans Med Imaging 29(2):488–501

Chang S, Mirabal Y, Atkinson E, Cox D, Malpica A, Follen M, Richards-Kortum R (2005) Combined reflectance and flourescense spectroscopy for in vivo detection of cervical pre-cancer. J Lower Genetal Tract Dis 10(2):24–31

Gordon S, Zimmerman G, Greenspan H (2004) Image segmentation of uterine cervix images for indexing in PACS. In: Proceedings IEEE symposium on computer-based medical systems, pp 1–6

Greenspan H, Gordon S, Zimmerman G, Lotenberg S, Jeronimo J, Antani S, Long R (2008) Automatic detection of anatomical landmarks in uterine cervix images. IEEE Trans Med Imaging 28(3):454–468

Herrero R, Schiffman MH, Bratti C, Hildesheim A, Balmaceda I, Sherman ME, Morales J (1997) Design and methods of a population-based natural history study of cervical neoplasia in a rural province of Costa Rica: the Guanacaste Project. Revista Panamericana de SaludPública 1(5):362–375

Holland JH (1975) Adaptation in natural and artificial systems. University of Michigan Press, Michigan

Horng JT, Hu KC, Wu LC, Huang HD, Lin FM, Huang SL, Chu TY (2004) Identifying the combination of genetic factors that determine susceptibility to cervical cancer. IEEE Trans Inf Technol Biomed 8(1):59–66

Jang JS (1993) ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans Syst Man Cybern 23(3):665–685

Ji Q, Engel J, Craine E (2000) Classifying cervix tissue patterns with texture analysis. Pattern Recogn 33(9):1561–1573

Kim E, Huang X (2013) A data driven approach to cervigram image analysis and classification. In: Color Medical Image Analysis, Lecture Notes in Computational Vision and Biomechanics book series, vol 6. pp 1–13

Kumar D, Hammed S, Ali HAVH (2012) Detecting diseased images by segmentation and classification based on semi-supervised learning. In: International conference on hybrid intelligent systems (HIS), pp 549–554

Mitchell MF, Schottenfeld D, Tortolero-Luna G, Cantor SB, Richards-Kortum R (1998) Colposcopy for the diagnosis of squamous intraepithelial lesions: a meta-analysis. Obstet Gynecol 91(4):626–631

Nallathambi B, Karthigaikumar P, Paul A (2017) High performance multi-operand adder for medical images. Intell Autom Soft Comput 23(2):345–349

Park SY, Follen M, Milbourne A, Rhodes H, Malpica A, MacKinnon NB, Richards-Kortum RR (2008) Automated image analysis of digital colposcopy for the detection of cervical neoplasia. J Biomed Opt 13(1):1–10

Park SY, Sargent D, Lieberman R, Gustafsson U (2011) Domain specific image analysis for cervical neoplasia detection based on conditional random fields. IEEE Trans Med Imaging 30(3):867–878

Roland KB, Benard VB, Greek A, Hawkins NA, Manninen D, Saraiya M (2013) Primary care provider practices and beliefs related to cervical cancer screening with the HPV test in Federally Qualified Health Centers. Prev Med 57(5):419–425

Sarwar A, Suri J, Ali M, Sharma V (2016) Novel benchmark database of digitized and calibrated cervical cells for artificial intelligence based screening of cervical cancer. J Amb Intell Humaniz Comput 7(4):593–606

Song D, Kim E, Huang X, Patruno J, Munzo-Avila H, Heflin J, Long RL, Antani S (2015) Multimodal entity coreference for cervical dysplasia diagnosis. IEEE Trans Med Imaging 34(1):229–245

Thekkek N, Richards-Kortum R (2008) Optical imaging for cervical cancer detection: solutions for a continuing global problem. Nat Rev Cancer 8(9):725–731

Hsu CH, Manogaran G, Panchatcharam P, Vivekanandan S. A (2018) A new approach for prediction of lung carcinoma using back propogation neural network with decision tree classifiers. In: 2018 IEEE 8th International Symposium on Cloud and Service Computing (SC2), pp 111–115

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.