CHAPTER 5

A VERSATILE DETECTION OF CERVICAL CANCER WITH i-WFCM AND DEEP LEARNING BASED RBM CLASSIFICATION (CERVI-CYTO-RBM)

5.1 Introduction

Cervical cancer is the third most universal cancer discovered among women. Approximately in 2018, there were 5,70,000 cervical cancer cases and 3,11,000 deaths. However, the assumed age-based cervical cancer is 13.1 in 1,00,000 global women, which differs across nations (Buskwofie, David-West, & Clare, 2020). However, the determined infection with a greater risk of HPV is the main reason for cervical cancer and pre-cancer growth. In preceding days, Cytology was utilized for major cervical screening; however, testing of HPV at present has substituted cytology in several nations. Since testing of HPV has a greater sensitivity rather than cytology for great-rate cervical pre-cancer CIN (Cervical Intraepithelial Neoplasia 2+) prediction. At this moment, cytology triage is used to eradicate over-referral to colposcopy and HPV-affected female over treatment. Followed by histopathology and tissue biopsy, colposcopy will be conducted in order to screen the cervix after a test of abnormal Pap (Shaikh, Daniel, & Lyng, 2023).

In existing methods, the techniques for diagnosis and screening of pre-cancer and cervical cancer are limited to early detection of cervical cancer. Therefore, to predict the effective features, i-WFCM has been incorporated in the segmentation process. The image is segmented by proposing improved weighted FCM algorithm, which is affected through outliers, noise, or a few other artefacts. Hence, the proposed algorithm takes minimal computational time to recognize clusters and is employed in image segmentation. Subsequently, the active features are mined from the fragmented area, and the features are utilized via RBM regarding DL in order to categorize cancer. Finally, the execution of the introduced cervical cancer prediction method can be evaluated based on accuracy, precision, recall and F-Measure. As a result, the overall performance of the proposed model attains better outcome compared to traditional algorithms such as RVDLNN, DLNN, AlexNet, and ResNet50V2 methods for predicting cervical cancer.

5.2 Methodology

The research methodology indicates that the utilization of DL is progressing and proceeds with medical images, especially in cervical cancer diagnosis. Whereas the present study employs enhanced cervical cancer segmentation process by using improved WFCM in order to attain enhanced performance. The workflow of proposed research methodology for cervical cancer prediction is illustrated in Figure 5.1.

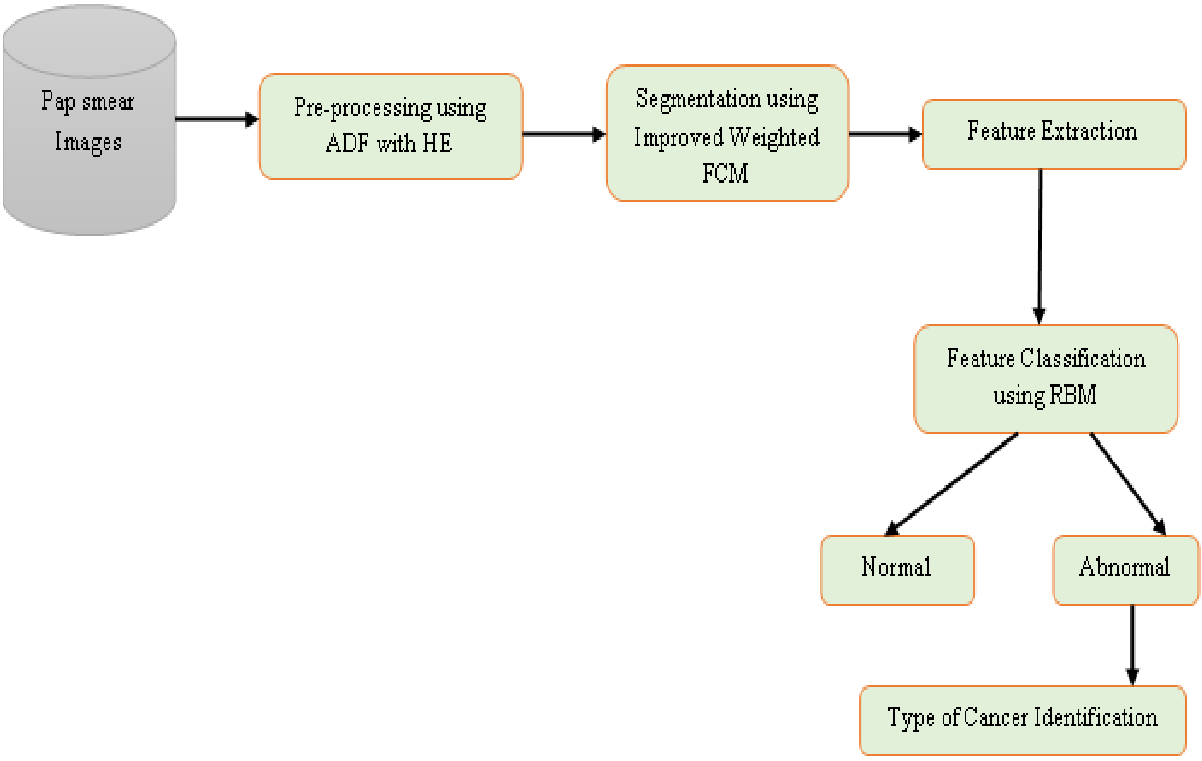

Figure 5.1 Overall Flow of the Proposed System

From Figure 5.1, it is understandable that the proposed system uses Pap smear images for cervical cancer prediction. Therefore, Pap smear images are processed into the proposed system for pre-processing. In the pre-processing phase, the proposed system uses ADF (Anisotropic Diffusion Filter) with HE (Histogram Equalization) technique. ADF adaptively eliminates noise and maintains the image edges, whereas HE processes images to regulate the image contrast by altering the histogram's intensity distribution. Then, the pre-processed image is entered into the segmentation phase, where i-WFCM is used for segmenting the pre-processed images. The working of i-WFCM algorithm is to map the input data by non-linear transformation using FCM in a greater-dimensional characteristic. Subsequently, the features are extracted from the segmented images. This process is done to extract more information from images. Then, feature classification is accomplished through utilizing RBM, which influences symmetric, a non-directed relation with the exclusion of several intra-layer relations among the hidden as well as visible nodes. Using the classification process, the images are classified into five classes, in the case of multi-cell images and seven classes, in the case of single cell images.

5.2.1 Data Selection

In the proposed study, the SIPaKMeD and Herlev datasets are used. Dataset SIPaKMeD is used to classify multi-cells and Herlev is used to classify single cells.

5.2.1.1 SIPaKMeD Dataset

The SIPaKMeD dataset comprises of 4049 isolated cells images, which were manually extracted from 966 group cell Pap smear slide images. However, skilled cytopathologists divided these cells into five categories based on cellular morphology and advent. A CCD camera captures the cell images and enhanced into an optical microscope (OLYMPUS BX53F). Moreover, the cell dispersal in class distribution is discussed in Table 5.1.

Table 5.1 Cell Distribution of SIPAKMED

Classes	**No. of Imageries**
Koilocytotic	238
Metaplastic	271
Superficial/Intermediary	126
Dyskeratotic	223
Parabasal	108
Aggregate	**966**

Among the five categories, Parabasal and Superficial/Intermediate cells are considered normal cells. In contrast, Metaplastic cells are regarded as benign cells, and the others are considered abnormal cells.

Superficial-Intermediate cells*:* Generally, it is a group of cells that are identified in the Pap test. The physical characteristics involves flat with either polygonal, oval, or round in shape. However, due to severe lesions, these kinds of cells exhibit the features of morphological alterations.

Parabasal cells*:* The cells that are in shape of small epithelial and immature squamous cell which is present in a typical vaginal smear. Also, Parabasal cells have the identical morphological properties with the cells found in metaplastic, which is hard to differentiate between them.

Koilocytotic cells*:* Koilocytotic cells are usually present in squamous cells that are mature and least noticeable in metaplastic-type koilocytic cells. For HPV infection, these cells are pathognomonic, and the nucleus of koilocytes usually depicts different degrees of degeneration based on the various stages and types of infection.

Metaplastic cells*:* They are large or small parabasal-kind cells with essential cellular borders, frequently displaying eccentric nuclei as well as rarely consist of an intracellular vacuole. However, their presence in the Pap test confers the greater prediction rates of pre-cancerous lesions.

Dyskeratotic cells*:* These are squamous cells that are gone through premature irregular keratinization in separate cells or frequently in three-dimensional clusters. Hence, these cells are generally three-dimensional clusters, thick, and it is hard to vary either the cytoplasm margins or the nucleus.

5.2.1.2. Herlev Dataset

Herlev dataset includes 917 single cell images which are classified into seven classes. Among the seven classes, Superficial, Intermediate and Columnar classes come under the normal cell category. Light, Moderate, and Severe come under precancerous cell category. Carcinoma_in_situ comes under the cancerous cell category.

Superficial: Superficial cells are found in the outermost layer of the cervix. It is large in size with small nucleus.

Intermediate: These types of cells are found in middle layer of the cervix.

Columnar: These cells with columnar shape are found in the columnar epithelium of the cervix.

Light: Light_dysplastic cells are slightly abnormal with small changes in shape and size of the nucleus.

Moderate: This class contains cells that are abnormal with more large and irregular nuclei.

Severe: These cells are considered precancerous with abnormal nuclei.

Carcinoma: The cells which are abnormal with change in shape and size of the nuclei comes in this class.

5.2.2 Pre-Processing

Pre-processing is a technique that performs certain procedures on an image, with the aim of attaining an improved imagery or to extract some required information from an imagery. However, the utilization of image pre-processing techniques is to enhance image quality or extract the information from an image. This technique is applicable to various fields such as medical imaging, scientific research, entertainment, and remote sensing applications.

In the proposed study, the functions in image are at the minimum level of abstraction, which is considered pre-processing, that provides output as well as input images together are intensity images. Typically, an intensity image is considered as an imagery function values matrix (brightness). However, the typical images are the same as the actual information, which is taken through the sensor. Pre-processing techniques are categorized for performing transformations in images such as translation, rotation, and scaling. Pre-processing aims to enhance the image quality, which suppresses the biases or augment a few image characteristics that are essential for further processing. Efficiently, adaptive histogram equalization is utilized to generate improved image, which is depicted in Figure 5.2.

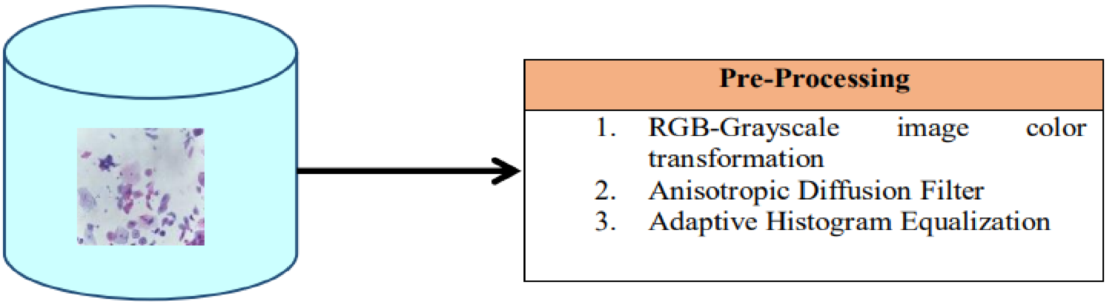

Figure 5.2 Pre-processing

5.2.2.1 Anisotropic Diffusion Filter (ADF)

The performance of image processing algorithms is degraded due to noise present in an image and eliminates the clarity, which affects its applications for analysis or display.

Moreover, high-frequency noise available in magnetic resonance images can be eliminated through the process of filtering. In the proposed study, ADF is used for pre-processing method. ADF is introduced to eliminate the noise and to maintain the image edges. Therefore, ADF is the first step of pre-processing to transform the RGB image to Grayscale of a Pap smear image. ADF is used to improve image quality by eradicating the image's noise. However, anisotropic diffusion is a known Perona–Malik diffusion that aims to eliminate image noise without eliminating the important part of the image contents. These contents include lines, edges, or fewer significant data for medical image explanation. Regarding a similar technique to anisotropic diffusion, imagery generates a parameterized family of consecutively many blurred images.

Partial Differential Equations (PDE) is utilized to design the ADF, which simplifies the process of image diffusion. Hence, bound imageries are generated using a filter in the diffusion process, the superposition and the input image are utilized for output image generation. As a result, PDE is often used for image edge prediction, noise removal, and detail conservation, which is the anisotropic diffusion equation's source. Whereas the edge data of an image can be conserved through denoising based on the equation that modifies the diffusion coefficient with respect to the image features. As a result, the diffusion model is expressed as shown in Eq. (5.1).

$$\partial I/\partial t = div(c(a,b,t)\nabla I) = \nabla c.\nabla I + c(a,b,t)\Delta I \tag{5.1}$$

Here, Laplacian is represented by Δ, gradient is denoted by ∇, divergence operator is denoted by div(...) and diffusion coefficient is denoted by c(a, b, t) Δ

For t>0, the output image is presented as I (., t) with greater t generating dimmer imageries.
c (a, b, t) regulates the diffusion value, which is generally selected as an image gradient operation to maintain image edges.

$$c(\|\nabla I\|)=e^{-}(\frac{\|\nabla I\|}{K})^2 \tag{5.2}$$

Function of ADF in the proposed study

The functions of ADF in the proposed study are described below.

- The pre-processing technique transforms the colour of given cervical cancer imagery to grayscale, using the ADF that aims at reducing the noise from imagery to achieve the best quality outcome.

- The use of ADF generates the exact transformation, which results in better accuracy in the proposed study.

5.2.2.2 Histogram Equalization (HE)

It is an image processing technique that intensifies the contrast in images. This can be done by effectively spreading out the frequent intensity values by protracting the intensity ranges in particular image.

Advantages

The process can improve the view of medical imaging and the detail present in images that are either under or overexposed. The major advantage is due to the straightforward approach for being adaptable to perform an invertible operator and the input image. Image contrast enhancement can be executed by utilizing a simple HE process.

Limitations

This technique is most familiar, as it is fast processing and is easy to implement. But, HE possess some of the limitations, which are described below,

- It augments noise to the output image, while progressing the contrast of its background and then establishes a distorted signal.

- It can generate undesirable impacts (such as visible image gradient) when imageries with low color depth are applied.

The HE has above drawbacks, but when combined with ADF, it generates clear images by filtering the noise in the proposed system.

5.2.2.3 Anisotropic Diffusion Filter – Histogram Equalization

In the proposed study, ADF-HE is used in a pre-processing stage to reduce noise and maintain the image edges, and HE adjusts the image contrast by modifying the intensity of the histogram distribution. A supervised process is used for segmenting region of image by initiating a seed or contour. An adaptive anisotropic diffusion is an estimated statistics that are utilized to extract the images from region of interest, which can be further applied for adaptive choosing of conduction operation, which leads to a better smoothness of an image and improves the clarity of the image in the proposed system.

5.2.2.4 Contrast Limited Adaptive Histogram Equalization (CLAHE)

In the proposed system, CLAHE is an enhanced algorithm of AHE technique. Where, imagery is fragmented by the CLAHE algorithm into contextual areas. Every histogram of the contextual area is produced, and then it is clipped at a prefixed rate. Meanwhile, the histograms are reordered with clipped quantity, and the common limitation of over-enhancement and edge-shadowing in AHE technique is minimized in enhanced CLAHE algorithm. Hence, its efficacy is improved by enhancing low-contrast medical imageries. This technique also generates the hidden features of the images more transparent by redistributing the grey values. However, every pixel's intensity level is associated with its adjacent pixels' intensity level to identify its status. Hence, the pixel is at a new intensity value, which is proportional to its available range's rank. Finally, the contextual or resident region in this technique prefers boost contrast rather than the whole image.

5.2.3 Segmentation: improved-Weighted Fuzzy C-Means

In unsupervised learning, clustering is an approach that classifies the data objects into various sets based on the resemblances and variations across them. In the statistical machine learning and data mining, it is one of the significant tasks. The clustering algorithm is further categorized into fuzzy clustering and hard clustering techniques. These clustering approaches have been effectively utilized in diverse areas of information retrieval, protein structure clustering, anomaly identification, electricity usage, pattern mining, image processing and pattern acknowledgement of gene expression. The algorithm utilized in this study is FCM.

5.2.3.1 Fuzzy C-Means

FCM facilitates to offer clusters with moderately uniform sizes, even if the cluster mass of input data is diverse. Several studies have detected that the FCM has a greater uniform impact than the k means; it diverges with deviations in fuzzier value. The FCM may generate biased outcomes due to uneven distribution of datasets across clusters which are considered as a huge drawback.

5.2.3.2 Weighted Fuzzy C-Means

The advancement of the conventional FCM is Weighted FCM (WFCM), where weights are allotted to every data point in accordance with its relative significance in the dataset. These weights are utilized to regulate the contribution of every data point to the procedure of clustering. It makes the clustering more effective and robust. WFCM has the capacity to manage imbalanced datasets. The weight allocated to data points act as a filter and minimizes the impact of noisy or outlier datasets on the clustering procedure. One of the frequently used techniques for analyzing images is image segmentation in digital image processing. This method involves breaking down imagery into smaller, distinct areas regarding the unique features of its pixels. Whereas the image segmentation aims to isolate the image foreground from the background or group pixels into distinct areas according to the similarities in shape or color.

In general, noise reduction is considered one of the main applications of image segmentation. By segmenting the image, the noise can be isolated and removed, that results in a cleaner and more visually appealing image. This is especially useful in medical imaging, where accurate and clear images are necessary for diagnosis and treatment. Another significant application of image segmentation is to identify objects in satellite images. Image segmentation is also widely used in the field of object detection and recognition. This is especially useful in surveillance and security systems, where the recognition and detection of objects and people are crucial for ensuring safety and security. One of the most innovative applications of image segmentation is in automating traffic management systems. However, it has a few limitations including the weights in FCM are decided in accordance with the cluster centres and Euclidean distance among data points that can handle noisy data and outliers. It may affect the precision of clustering outcomes. To overcome these challenges, a further enhancement of WFCM, regarded as improved WFCM (i-WFCM), has been proposed.

5.2.3.3 improved - Weighted Fuzzy C-Means Algorithm

This approach involves native information that is useful in deciding the weights of data points. It considers the weights of neighborhood information for every dataset, taking into account its resemblance and space to other datasets within the identical cluster. The i-WFCM has revealed higher performance relative to conventional FCM and WFCM in running time, robustness, and accuracy.

The objective task of FCM is described as Eq. (5.3),

$$Y_i = \sum_{i=1}^{n} \sum_{j=1}^{c} b_{ij}^{l} \parallel x_i - c_j \parallel^2 \quad (5.3)$$

Here, i depicts the level of fuzziness, and it is an exact value higher than one, b_{ij} is the connection level of the i^{th} data in the j^{th} cluster.$c_j$represents the cluster center, as well as,

x_i depicts data points. Further, $\|.\|$ addresses the Euclidean distance. c shows the cluster quantity, on the other hand, represents the amount of data points. The fortitude to the center of the cluster is not precise. For recognizing the optimal weighting factor and enhancing the cluster center, i-WFCM algorithm is introduced. Meanwhile, the traditional clustering algorithm runs in the actual data space. However, the i-WFCM clustering remains effective in the initial data space, meaning that the prototypes are situated within the data space rather than following the typical approach used in FCM. This makes i-WFCM especially advantageous for handling incomplete data, as it displays a higher level of resilience towards outliers and noise compared to FCM.

The i-WFCM is an introduced method for performing FCM (Fuzzy C Means) in a greater-dimensional feature space. It can be achieved through mapping the input data into the higher-dimensional space using a nonlinear transform. The goal of this algorithm is to improve the accuracy and effectiveness of traditional FCM by integrating the concept of kernel weighting. The proposed i-WFCM permits a more precise representation of the data and ultimately leads to better clustering results.

After mapping the data into higher-dimensional space, the i-WFCM algorithm diminishes an objective function using a technique known as Kernel Weighting. It is shown in Eq. (5.4).

$$Y_{\omega} = \sum_{i=1}^{k} \sum_{j=1}^{n} b_{ij}^{i} \parallel \omega(X_i) - \omega(V_i) \parallel^2 \quad (5.4)$$

This technique assigns weights to the data points based on their proximity to cluster centers. This means that the data points that are closer to the center of a cluster will have a higher weight, while the points that are further away will have a lower weight. This approach boosts the clustering accuracy by contributing more weight to the data points that are more relevant to each cluster. Whereas the minimized objective function through the i-WFCM algorithm contains the combination of two terms: the FCM objective function and a kernel weighting term. The FCM objective function is a measure of how well the data points are clustered, while the kernel weighting term ensures that clusters are well-separated and distinct. By curtailing this objective function, the i-WFCM algorithm is able to achieve better clustering results compared to traditional FCM.

In equation (5.4), b_{ij} represents the association of X_i in cluster I, $\omega(V_i)$ is the middle of cluster I in the feature space, whereas the mapping from the input space $X$to the feature space F is represented by ω.

Algorithm
Algorithm of i-WFCM 1. Choose first prototype category {Vi} c 2. Modify total memberships b_{ij} using weighted average for middle of the cluster using Eq(5.5) 3. $u_{ik} = \frac{(1-\mathrm{B(xk,vi)})^{-\frac{1}{(m-1)}}}{\sum_{j=1}^{c}(1-\mathrm{b((xk,xj\ )}^{-\frac{1}{(m-1)}}}$ (5.5) 4. Get the clusters' prototype in terms of weighted average by the succeeding equation Eq(5.6). 5. $v_i = \frac{\sum_{k=1}^{n} u_{ik}^{m} B\ (x_k, v_i) x_k}{\sum_{k=1}^{n} u_{ik}^{m} B\ (x_k, v_i)}$ (5.6) 6. Repeat phase 2-3 till discontinuing standard is encountered 7. The final principle is$\lvert V_{new} - V_{old} \rvert \leq \varepsilon$ 8. The Euclidean value is$\lVert . \rVert$. Whereas, the vector of cluster centres is denoted by V, and ε is a least quantity that is fixable by the user (here ε=0.01)

5.2.4 Feature Extraction

Feature extraction is the process of selecting and converting the most suitable and informative features or attributes from a dataset. These features can be in the form of numerical, categorical, or textual data, and provide valuable insights and patterns that assist in decision-making and predicting outcomes. The primary objective of feature extraction is to decrease dataset dimensionality while retaining as much information as possible. This is particularly beneficial for handling datasets containing a high number of dimensions. By extracting only the most essential features, the data becomes more manageable, and models can produce better outcomes.

One of the main benefits of feature extraction is its potential to minimize the presence of noise and irrelevant information in a dataset. This is accomplished through choosing the features that only have a significant impact on the target variable. By excluding unnecessary data, models become more precise and trustworthy. Moreover, feature extraction can also extract hidden patterns and relationships within the data. It is particularly advantageous in the fields of image and speech recognition, where extracted features aim at exposing critical characteristics that are not easily noticeable to the human eye.

Feature extraction is a significant stage during the categorization of cervical cancer from Pap smear imageries. It incorporates algorithms to extract appropriate data from the images that helps in characterizing the cancer and feeding the classifier. These algorithms range from fundamental to advanced approaches and are constantly developing with advancements in technology and research. Some of the most commonly used features are geometrical features such as eccentricity, perimeter, the area, concavity, diameter, and asymmetry. These features are extracted using image processing techniques and are utilized to identify the texture, size, and the shape of cervical cancer cells.

In addition to geometrical features, texture analysis has also gained crucial recognition in recent years for being an effective instrument in diagnosing cervical cancer. This involves the practice of advanced techniques such as GLCM (Gray-Level Co-occurrence Matrix) and Haralick features to gather abstract data about the textural patterns available in the Pap smear images. The presence of abnormal textures designates the existence of cancerous cells, supporting the early and accurate diagnosis of cervical cancer. These techniques involve training a neural network on a large dataset containing Pap smear image to automatically extract relevant features and classify the images as normal or abnormal. This approach has shown promising outcomes and has the ability to further intensify the efficacy and precision of cervical cancer diagnosis.

5.2.5 Classification through Restricted Boltzman Machine – Deep Belief Network

RBM is a unique simulated neural network employed for unsubstantiated learning processes such as pattern identification, feature learning, and dimensionality reduction. It is propagative stochastic dissemination over its inputs and can be utilized for several tasks in ML. It is a neural network with two layers involving hidden and visible data. The noticeable unit determines the input data on the other hand; unseen units learn attributes from the input data. The two layers are associated through weighted relations, and the neurons in every layer are not correlated together. The RBM training incorporates two major stages backward pass and forward pass. The input data are loaded in noticeable units and the stimulations of the hidden units are assessed in the forward pass. The hidden units' activations are employed to update the weights of relations among the visible and hidden units in the backward pass.

Advantages of Restricted Boltzmann Machine

- Incorporation of RBM in recognizing cervical cancer has a major benefit in its capacity to manage extensive and intricate datasets. In addition, it is effective at managing complexity and extracting insights, RBM is an excellent tool for detecting cervical cancer.

- RBM has the ability to take out and detect important features from the data automatically. When it comes to cervical cancer, RBM can detect crucial factors such as age, HPV infection, and other potential risk factors that could influence the disease.

- RBM has demonstrated great precision in identifying cervical cancer. It is a dependable tool for early detection of cervical cancer symptoms, resulting in improved treatment results.

- In the domain of medical diagnosis, it is frequent to encounter missing data. RBM can effectively handle incomplete data and still produce precise forecasts. This is achieved through the utilization of a method known as imputation.

- RBM is more affordable than the other machine learning algorithms. It uses minimal computational power, making it appealing to medical facilities with scarce resources.

Owing to its significant features, RBM can also be utilized in amalgamation with other neural network structures, such as convolutional neural networks and deep belief networks, to accomplish even more enhanced outcomes. These are all the factors that make RBM a powerful instrument for feature classification. Therefore, the present study adopts RBM for feature classification.

DBN is a type of DNN (Deep Neural Network) consisting of multiple unnoticed layers where every pair of associated layers within a DBN is regarded as RBM. The weights of these associations among the hidden and visible nodes are learned through Hinton's contrastive divergence algorithm. This algorithm is specifically structured for DBN as it considers the absence of intra-layer relations. By utilizing this algorithm, the network can efficiently learn the association among the varied layers and make a precise depiction of input data.

In an RBM, a function of energy is described in accordance with the noticeable unseen organization of Gaussian neurons as shown in Eq. (5.7).

$$E(v,h) = \sum \frac{(v_i - a_i)^2}{2\sigma_i^2} - \sum_i \sum_j \frac{v_i}{\sigma_i^2} h_j w_{ij} - b'h \qquad (5.7)$$

Here w is the noticeable unseen weight matrix, further, a as well as b are the bias vectors correspondingly. In RBM configuration, the joint probability, p (v,h) can be measured by using Eq. (5.8).

$$P(v,h) = \frac{exp^{-E(v,h)}}{k} \qquad (5.8)$$

Here, the divider function is K which can be represented in Eq. (5.9).

$$K = \sum_{v,h} exp^{-E(v,h)} \qquad (5.9)$$

Every vector in the RBM can be allocation in the probability as shown in Eq. (5.10).

$$P(v) = \frac{1}{K} \sum_h exp^{-E(v,h)} \qquad (5.10)$$

The conditional probabilities can be transcribed in accordance with the sigmoid function as shown in Eq. (5.11), (5.12) and (5.13).

$$P\left(h_j = 1 \middle| v\right) = \sigma\ (\sum_i \frac{v_i}{\sigma_i^2} w_{ij} + b_j \qquad (5.11)$$

Where $\sigma(X) = \frac{1}{1+ e^{-X}}$ (5.12)

And

$$p(v_i = v|h) = \epsilon(v \middle| \sum_j h_j w_{ij} + a_i, \sigma_I^2 \qquad (5.13)$$

5.3 Summary

After segmenting the cervix region, the subsequent task is to categorize cervical cells as regular or irregular by employing deep learning-driven RBM classification. RBM is a neural network trained with unsupervised learning to extract important data features. These characteristics are subsequently utilized for categorizing the data into various groupings. The performance of RBM-DBN has been demonstrated to surpass conventional machine learning techniques while analyzing different medical images. In case of cervical cancer identification, RBM-DBN is trained using a vast collection of images of cervical cells to understand the differences between healthy and unhealthy cells. Subsequently, the trained RBM-DBN is utilized to categorize fresh cervical cell images as either healthy or unhealthy.

This technique has showcased utmost performance in precisely identifying abnormal cells, even in contradiction of visual distinctions between regular and irregular cells are slight. The utilization of i-WFCM along RBM-DBN for classification is known as a flexible method for identifying cervical cancer. The i-WFCM is a precise and effective technique for segmenting the cervix area, while RBM-DBN classification can precisely categorize cervical cells as either normal or abnormal. This method could enhance the precision of cervical cancer diagnosis and diminish the strain on healthcare systems. Furthermore, the flexibility is profound in this technique due to its capability to identify not just cervical cancer, but also other cervical irregularities namely, precancerous lesions. This is critical because it enables prompt identification and action, which could help prevent the onset of cervical cancer.