SURVEY ON AUTOMATED COLPOSCOPY IMAGE CLASSIFICATION

[1]Monisha B, [2]Akshaya rani R, [3]Kiran Manie R, [4] Ranichandra V
[1], [2], [3]-Student,[4] –Assistant Professor
[1]Computer Science and Engineering,
[1]Easwari Engineering College, Chennai, India

Abstract : ***Uterine cervical cancer is the second most common cancer in women worldwide, with nearly 500,000 new cases and over 270,000 deaths annually. Effective diagnosis in early stages can give women a better chance of full healing and survival. Digital imaging technologies and neural networks allow us to assist the physician with an automated Computer-Aided Diagnosis(CAD) system. In Colposcopy, epithelium that turns white after the application of acetic acid is called aceto-white epithelium. Aceto-white epithelium is one of the major diagnostic features observed in detecting cancer and precancerous regions. Automatic extraction of aceto-white regions from the cervical images has been a challenging task due to specular reflection and most importantly large intra-patient variation. The proposed structure aims to ease the diagnosis of cervix cancer though an automated smart system, which allows the physician to upload the colposcopic images through a web application, track the patients Electronic Medical Records(EMR), detect the abnormality in the images and finally grade the severity level of CIN (Cervical Intra-epithelical Neoplasia) using neural networks. The proposed system aims to enhance the functioning and processing of cervical cancer screening, thus trying to make the process more reliable.***

IndexTerms - Uterine cervical cancer, neural networks, aceto-white epithelium, colposcopy.

I. INTRODUCTION

Cervical cancer has marked its existence in every 3 out of 10 women in the world. It is quite unusual as there are no explicit symptoms to this variety of cancer in the early stages. Women who are sexually active are suggested to undergo the preliminary tests every three years to ensure that they are out of danger. Another unusual fact is that, the existence of abnormal cells might not always infer to them being cancerous cells. So, the diagnosis should be efficient enough to as to not invoke any false alarms.

The diagnosis of cervical cancer involves three stages namely: pap smear, colposcopy and a biopsy. Pap smear is the initial microscopic test where a chunk of cells is scraped from the cervical surface and checked under the microscope for abnormalities. If any sort of abnormality is detected, the second test is conducted which is colposcopy. Colposcopy is the procedure where acetic acid to applied directly to the surface of the skin and then captured using the telescopic camera that is a part of the colposcopic device. These images are very crucial as these help the medical experts to decide whether or not to move to the final stage of diagnosis; biopsy. The medical experts analyze the images and conclude whether or not the abnormal cells are cancerous and also the level of severity.

Artificial neural networks (ANNs) or connectionist systems are computing systems that have been designed based on the biological neural networks that constitute actual animal brains. Such systems learn tasks by considering examples, generally without task-specific programming. This happens to be one of the major differences between Machine Learning and ANN; Machine learning involves feature engineering where the patterns are explicitly given to the machine to learn them whereas in Artificial Neural Networks, the patterns are analyzed and figured out automatically.

An ANN is based on a collection of connected units or nodes called artificial neurons. Each connection between the artificial neurons transmit a signal from one to another. The artificial neuron that receives the signal processes it and then signals the artificial neurons connected to it. The original goal of the ANN approach was to solve problems in the same way that a human brain would. ANNs have been used for a variety of tasks, including computer vision, speech recognition, machine translation, social network filtering, playing board and video games and also medical diagnosis. This application of the ANN can be applied to a particular variety of cancer called cervix cancer. It can be used to assist in the screening process by helping with the medical image processing

II. LITERATURE SURVEY

There have been a lot of researches and ideas proposed on medical image processing. Some of them are:

Amir Alush et al., [1] proposed an idea for automated lesion detection in cervix images. The proposed approach involved three components. The first component involved using the watershed shed algorithm to establish the boundaries by converting the image into an edge map. The next one corresponds to a feature-free boundary representation method based on a visual dictionary. The third one is to view the watershed map as an MRF where each watershed superpixel corresponds to a binary random variable. This was a major step forward towards automated boundary detection. The proposed algorithm helps with the detection of legion regions. However, the approach did not include appropriate boundary information. It used no added region-based input. So the predictions were not accurate.

M. Falagario et al., [2] involved pre-processing the digital images followed by feature extraction and classification of the images to their corresponding grades of severity. The steps are as follows:

a. Region of Interest(ROI) detection using k-means algorithm.
b. Conversion from RGB to HSI scale to support human eye vision.
c. Thresholding using his color histogram to specify the boundaries.

Finally, separation of regions using region filling.The system was expected to provide a fast and a reliable diagnosis. It is also expected to be used as a tutorial for the training phases of students applying to gynecology. However, the flaw was that it uses traditional methods. Each process requires a separate module which makes the process time consuming and less efficient.

Charence Wong et al., [3] focused focuses on the application of deep learning to a variety of streams including bioinformatics, medical imaging, pervasive sensing, medical informatics and public health. This paper explains the advantage of deep learning over other machine

learning models and algorithms. It emphasizes that deep learning involves the use of many hidden layers which has proved to be an architectural advantage. It focuses on the different deep learning architectures such as:
- Deep Neural Network which is used for classification and regression problems.
- Deep Autoencoder, who's main purpose is feature extraction.
- Convolutional Neural Network are best suited to process 2D data such as images.

Deep learning proves to be an efficient advancement in the field of machine learning. The paper states that it has provided a positive revival of NNs and connectionism. Issues were that, constant involvement and research, exclusively in health informatics might slow down the development of new machine learning algorithms.

Claudia L. Werner et al., [4] talked about the essentials of colposcopy. It elaborates on the process of colposcopic usage, the role of importance it plays in the diagnosis of cervical cancer, its benefits and the terminologies used. It states that Colposcopy helps in the evaluation of abnormal or inconclusive screening tests. It helps in the identification of cervical precancers that are possible to treat and also the management of abnormalities that cannot be progressed. The main purpose of the colposcope device is to capture images of the cervix aiding in the recognition of lesions, checking their visibility and spread that helps establish the severity. This research explains the foundations of colposcopy. It specifies the effectiveness of establishing the aceto-white region features for aiding in classifying the grades of severity.

Hector Gabriel Acosta-Mesa et al., [5] talked about three methods for the classification of aceto white region. It focuses on a comparative analysis between k-Nearest Neighbors, Naive Bayes and C4.5 algorithms. The paper elaborates that the digital image initially undergoes preprocessing that focuses on achieving anatomical correspondence. This is followed by the process of time series extraction. Then, the three methods on classification are explained and compared to determine the most effective one.
- k-Nearest Neighbors: It is an instance based learning method where each instance is assigned to class; one among the k nearest neighbors, based on the distance.
- Naive Bayes: It is a probabilistic approach that uses Bayes theorem. This method gives a probability value that the new member is assigned to a particular class, based on observations from the training set.
- C4.5: This is a decision tree approach where each time series value is specified as an attribute.

A comparison between the three methods was conducted and the performance was assessed by giving different continuous and discrete values. The evaluation was based on three factors namely sensitivity, specificity and accuracy. And the complete study was based on the dynamics of acetowhitening. However, it involved quite a number of false positives and false negatives.

Angela Hui-Chia Liu et al., [6] focused on the evaluation of a process named Endocervical Curettage(ECC) that is used for the diagnosis of high grade abnormalities.It resulted in a particular level of sensitivity. However, the issues involved not having a proper criteria for ECC.

Marjolein Hermens et al., [7] aimed at explaining the effect of Multimodal Hyperspectroscopy(MHS) which is one of the digital colposcopy technique. It is known to have shown a high sensitivity rate. The goal was to provide a prospective evaluation of the MHS clinical values for the detection of high -grade cervical intraepithelial neoplasia. The process involved performing a MHS scan before the conventional colposcopic examination. The device is set and the cervical tube is inserted into the vagina using a live video feed to focus and look at the field of view. Then, the measurements are taken automatically using the software. The necessary detailed are checked and the cervical tube is removed. The results are using to grade the severity. After the completion of the MHS scan, the conventional diagnosis mechanism is conducted. The results are compared. It resulted in high sensitivity helping detect high-grade abnormalities. But, it caused a specificity value resulting in limited clinical value.

Maria Fasero et al., [8] talked about the improved results obtained by using Dynamic Spectral Imaging(DSI) when compared to conventional Colposcopic tests. The dynamic spectral imaging system is an new colposcopic device that was designed to improve the sensitivity. It is a digital device that allows the medical experts to perform the conventional colposcopy and also measure the dynamic color changes that appears on the cervix after applying the specific acids. The process involves measuring the dynamic color changes and then calculate the color-coded dynamic spectral imaging map and also display it. This calculation is done based on the intensity and spread of the aceto-white regions. In comparison to the conventional method, it is known to increase the sensitivity to an 87.8% along with a drop in specificity an 85.6% .

Lei Wang et al., [9] put forward an algorithm for the automatic detection of CIN. It proposes using a segmentation algorithm for the aceto-white regions. In this study, the method involves establishing the cervix region using fuzzy c-means clustering algorithm in the pre-acetic acid test image(the image obtained before applying acetic acid to the skin). Then, from the post-acetic acid test image, the image registration is obtained, that is, the lesion region whitens for up to almost 60 seconds after the application of acetic acid. These two images are correlated to finally obtain the aceto-white region using a level set algorithm. The proposed method using the ratio image and level set algorithm together caused the development of a new segmentation algorithm that resulted in finer sensitivity in comparison to the traditional methods.

Masamichi Kashimura et al., [10] emphasized about importance of considering new parameters to help with better diagnosis. Consideration of new diagnostic character helps tune the process yielding refined results. The human papillomavirus (HPV) test is considered a very sensitive method for the detection of cervical neoplasia. The combination of cytology and HPV test is considered to be the most reliable cervical cancer screening. A double negative result for both cytology and HPV test may prolong the screening interval because the future morbidity of high-grade cervical lesion is extremely low. The findings were divided into minor or major, single or a combination of two or three findings. The new criteria are fully acceptable.

III. RESEARCH TRENDS

In this survey, data has been collected from different research papers that have been published in colposcopy diagnosis and the area of cervical cancer diagnosis. Here,we have discussed 10 standard research papers that were published in the timeline 2010-2017.

The above details are depicted in the table below.

S.NO	Author and Title	Techniques/Algorithm	Result	Issue
1	Amir Alush, Hayit Greenspan, Jacob Goldberger, "Automated and	Watershed Algorithm	Detection of lesion regions	It did not include proper boundary extraction.

	Interactive Lesion Detection and Segmentation in Uterine Cervix Images"			
2	M.Falagario, C.Guaragnella, M.Traversi, "CaDdy-Colposcopy Learning machine for Computer Aided Diagnosis	K-means algorithm,Specular Reflection removal.	Tutorial for training phase of students.	Uses traditional approach, time consuming.
3	Daniele Rav`ı, Charence Wong, Fani Deligianni, Melissa Berthelot, Javier Andreu-Perez, Benny Lo, "Deep Learning for Health Informatics"	Convolutional neural network is best suited for 2D data such as images.	Efficient advancement in field of machine learning.	High computational cost.
4	Claudia L. Werner,MD, Cara Mathews, MD, Michelle J. Khan, et al., ASCCP Colposcopy Standards: Role of Colposcopy, Benefits, Potential Harms, and Terminology for Colposcopic Practice	Explains foundations of colposcopy.	Focus on aceto-white region features for aiding in classifying severity.	Helps in identification of cervix precancers.
5	Hector Gabriel Acosta-Mesa, Karina Gutierrez-Fragoso, Nicandro Cruz-Ramirez, Rodolfo Hernandez-Jimenez, Optimization of Classification Strategies of Acetowhite Temporal Patterns towards Improving Diagnostic Performance of Colposcopy	K-nearest neighbour, Naive Bayes.	Provides comparison between three methods specified in the algorithm.	Quite number of false positives and false negatives
6	Angela Hui-Chia Liu, MD, Joan Walker,MD, et al.,Diagnosis of Cerical Precancers by Endocervical Curettage at Colposcopy Women With Abnormal Cervical Cytology	Endocervical Curettage	Increasing the sensitivity	Criteria for performing ECC
7	Leon F.A.G. Massuger, MD, PHD, Marjolein Hermens, Bsc, et al., "Multimodal Hyperspectroscopic Imaging for Detection of High-Grade Cervical Intraepithelial Neoplasia"	Multimodal Hyperspectroscopy	High sensitivity	It caused specificity value resulting in limited clinical value.
8	Maria Fasero, Pluvio J. Coronado, Colposcopy combined with dynamic spectral imaging.A prospective clinical study	Uses dynamic spectral imaging.	sensitivity 87.8% specificity 85.6% .	Drop in specificity
9	Lei Wang, Ling Li, Jun Li, "Acetowhite region segmentation in uterine cervix images using a registered ratio image"	Fuzzy c-means clustering algorithm,Level set algorithm	Finer sensitivity	It does not stress in specificity
10	Masamichi Kashimura, Naoyuki Toki, Toshinori Kawagoe,"New diagnostic of colposcopy for uterine cervix neoplasia"	Suggests new parameters.	New criteria helps in better diagnosis.	Has not been put into practice yet.

IV. CONCLUSION

This survey paper involves tremendous information. It has put forth the various techniques and algorithms used in the field of colposcopic diagnosis and also in digital image processing, that has evolved over the past years. New algorithms were developed with the aim for producing improved results in comparison to the former methodologies. Although so many techniques have been designed, a very high level of accuracy and sensitivity had still not yet been achieved.

V. FUTURE WORK

An ideology for extra improved results in the field of cervical cancer diagnosis could be done by automating the entire processing. Although there exists several automation algorithms for each single module like segmentation of the image or classification of the image into the corresponding level of severity, each of the algorithms work separately. Correlation of these methods involve tedious work level as it has to ensure the proper functioning of the algorithms, their performance in sync and much work. So, it would result in great efficiency to design a model that automates the entire digital image processing and classification. Deep convolutional neural networks can be used for this purpose. This is mainly used for 2D data sets and each of the processes such as segmentation, feature extraction and classification can be made to happen in each of its hidden layers and produce one overall output predicting the grade.

REFERENCES

[1] Amir Alush, Hayit Greenspan, Jacob Goldberger, "Automated and Interactive Lesion Detection and Segmentation in Uterine Cervix Images", *IEEE transaction on medical imaging*, vol. 29, no. 2, February 2010.

[2] M.Falagario, C.Guaragnella, M.Traversi, "CaDdy-Colp oscopy Learning machine for Computer Aided Diagnosis", *IEEE Third International Conference on consumer electronics*, 2013.

[3] Daniele Rav`ı, Charence Wong, Fani Deligianni, Melissa Berthelot, Javier Andreu-Perez, Benny Lo, "Deep Learning for Health Informatics", *IEEE Journal of Health Informatics, IEEE*, 12, pp:1, 2017.

[4] Claudia L. Werner,MD, Cara Mathews, MD, Michelle J. Khan, et al., "ASCCP Colposcopy Standards: Role of Colposcopy, Benefits, Potential Harms, and Terminology for Colposcopic Practice", *Journal of lower genital tract disease*, vol. 21, no. 4, October 2017.

[5] Hector Gabriel Acosta-Mesa, Karina Gutierrez-Fragoso, Nicandro Cruz-Ramirez, Rodolfo Hernandez-Jimenez, "Optimization of Classification Strategies of Acetowhite Temporal Patterns towards Improving Diagnostic Performance of Colposcopy", *Computational and mathematical methods in medicine*, July 2017.

[6] Angela Hui-Chia Lui, MD, PHD, Joan Walker, MD et al., "Diagnosis of Cervical Precancers by Endocervical Curettage at Colposcopy Women With Abnormal Cervical Cytology", *Obstetrics and gynecology*, vol. 130, no. 6, December 2017.

[7] Leon F.A.G. Massuger, MD, PHD, Marjolein Hermens, Bsc, et al., "Multimodal Hyperspectroscopic Imaging for Detection of High-Grade Cervical Intraepithelial Neoplasia", *Journal of lower genital tract disease*, vol. 21, no. 3, July 2017.

[8] Maria Fasero, Pluvio J. Coronado, "Colposcopy combined with dynamic spectral imaging.A prospective clinical study", *European journal of obstetrics and gynecology and reproductive biology*, vol. 196, January 2016.

[9] Lei Wang, Ling Li, Jun Li, "Acetowhite region segmentation in uterine cervix images using a registered ratio image", *Computers in biology and medicine*, vol. 93, February 2018.

[10] Masamichi Kashimura, Naoyuki Toki, Toshinori Kawagoe,"New diagnostic of colposcopy for uterine cervix neoplasia", *Journal of obstetrics and gynaecology research*, vol. 43, no. 2, February 2017.