Cervical Cancer Analysis

Aishwarya Kadam[1], Divya Bharti B Kareti[2], Priyanka Kurli[3], Shreya Deshpande[4], Pratibha Badiger[5]

[1,2,3,4] Student, Dept. of Information Science Engineering, SDMCET, Dharwad, Karnataka, India
[5]Professor, Dept. of Information Science Engineering, SDMCET, Dharwad, Karnataka, India

---

Abstract – *Gynecological cancers are among the most common cancers in women and hence are an important public health issue. Due to the lack of cancer awareness programs, variable pathology, and dearth of proper screening facilities in developing countries such as India, most women report at advanced stages, adversely affecting the prognosis and clinical outcomes. Ovarian cancer has emerged as one of the most common malignancies affecting women in India. Cervical cancer remains the second most common cancer in women after breast cancer. The causes of cervical cancer are HPV (Human Papilloma Virus), smoking, oral contraceptives, multiple pregnancies. It can be prevented in adult women with early detection tests, such as the HPV test or PAP test, followed by treatment. Early detection tests can identify pre-cancerous lesions in the cervix, which can then be treated before the lesion develops into cervical cancer. In our proposed work, we used Machine Learning and Deep Learning algorithms to find a model capable of diagnosing cervical cancer with high accuracy and sensitivity.*

Key Words: Machine Learning (ML), Convolution Neural Network (CNN), Deep Learning, Cervical Cancer.

1. INTRODUCTION

Cervix is the lower, narrow end of the uterus that forms a canal between the uterus and vagina. Cervical cancer is caused by a sexually transmitted virus called Human Papilloma Virus (HPV) which accounts for 99.7% of all cervical cancer cases. In India, cervical cancer contributes to approximately 6–29% of all cancers in women and every year 122,844 women are diagnosed with cervical cancer and 67,477 die from this disease. In the early stages, there are no exact symptoms and side effects of the disease, but normal PAP smear screening is performed. Among other screening methods used to recognize malignant growth cells, the PAP smear test is excellent.

The data collection of different types of cancer cells are taken and are preprocessed and trained to get a trained model which when used for the detection compares the accuracy of each class type in the dataset, the one with the highest class type accuracy is responded as a result.

1.1 Literature Survey

An automated computer based technique has been a reliable method [1] as Cervical cancer is more common in women and worldwide it is most feared disease. Due to abnormal growth in the cervix cells, cervical cancer occurs and slowly it also spreads to the other organs of human body. Cervical cancer is caused by number reasons like human papillomavirus, using birth control pills, cigarette smoking, etc. In the initial stage, cervical cancer will not show any signs. However, if it is identified in earlier stage, it will be cured successfully. Nowadays, number of computer vision based approaches has been introduced to identify the cervical cancer disease and its stages. Cervical cancer arises due to uncontrolled development of cervical cells, they will not die instead they continue to divide. Literature reports that HPV virus, smoking, and weak immune system, etc are the causes of cervical cancer. Nowadays, death rate due to the cervical cancer is reduced significantly by detecting the cervical cancer in its early stage using the pap smear test. Screening process undergone for cervical cancer manually has higher issues of producing false negative rates in Pap smear test. Hence an alternate method came into existence called automated computer based technique to increase accuracy for testing cervical smears.

Image processing techniques are proposed here to detect the cervical cancer early [2] as Cervical cancer is most common malignancy in female. It arising from the cervix. There are different treatment like primary surgery, primary radio therapy, chemo therapy and combination therapy. There are many techniques to diagnosis cervical cancer the important ones are Pap smear test, LBC test, HPV test, Biopsy and different screening techniques. The automatic screening of the cervical malignant growth cells has been created by means of morphological image processing techniques. Cervical most cancers are screened manually with the aid of the usage of the Pap smear test and LCB check which does not deliver correct classification effects in classifying the normal and uncommon cervical cells inside the cervix region of the uterus. The manually screened technique suffers from excessive faux fee because of human errors and also value effective to be executed by means of the usage of the professional cytologist. in this paper, several methods are proposed for the automatic detection of cervical cancer the usage of image processing techniques. The automatic techniques are achieved to supply correct outcomes and to make effective type of ordinary and atypical cells.

The paper [3] proposed here is Cervical Cancer Diagnosis using CervixNet - A Deep Learning Approach. Cervical cancer is caused due to the Human Papilloma Virus (HPV) which leads to abnormal growth of cells in the cervix region. Regular testing for HPV in women has helped reduce the death rate in developed countries Motivated by the Deep Learning solutions in Biomedical imaging, we propose a novel CervixNet methodology which performs image enhancement on cervigrams followed by Segmenting the Region of Interest (RoI) and then classifying the RoI to determine the appropriate treatment. For the classification task, a novel Hierarchical Convolutional Mixture of Experts (HCME) algorithm is proposed. HCME is capable of tackling the problem of overfitting, given that small datasets are an inherent problem in the field of biomedical imaging. Our proposed methodology has outperformed all the existing methodologies on publicly available Intel and Mobile-ODT Kaggle dataset giving an Accuracy of 96.77% and kappa score of 0.951. Hence, the results obtained validate our approach to provide first level screening at a low cost. The proposes methodology CervixNet includes a) Image Enhancement, b) RoI Extraction and c) HCME: Proposed RoI Classification Algorithm.

Machine Learning technique [4] here highlights the novel idea of using bilateral filter for pre-processing of the images. In this further they have carried it in the following steps ,they have gathered the DATASET from HOSPITAL DENMARK from the Pap-Smear test then the images are gone through the Pre-processing with filtering using Bilateral filter on basis of filtration next they have segmentation process in which they have extracted the part that is used for the further detection(or study) next is feature extraction where the useful data from microscopically observed is taken for the classifier algorithms then feature selection is done finally classification is done and the results are prepared and submitted for further treatment process.

Deep learning approach [5] is used and proposed that the colposcopy image is an essential aid in early cancer diagnosis. The assessment and identification of people with irregular cytology who need further care or follow-up depend on the transition zone colposcopic examination (TZ). A new deep learning architecture name CYENET is proposed for classifying the cervical cancer type from colposcopic images. The image dataset is balanced using the oversampling technique for improving the classification results. Two models are presented in this paper. One is using a transfer learning approach with VGG19 architecture. The other is a dedicated new model called CYENET for cervical cancer type classification using the ODT colposcopy image dataset. Both the models are evaluated using classification accuracy, sensitivity, specificity, Cohen's Kappa score, and F1-measure. The classification accuracy for VGG19 was 73.3%. Relatively satisfied results are obtained for VGG (TL).

1.2 Proposed System

A model for a data system that is far-reaching, including the origins of knowledge and utilitarian sections. In terms of requirements, this model highlights a calculated framework that perceives screening systems and their data structures to be staged periodically after some time, screening modalities, accessible information, level of computerization, equipment and programming, and levels of human and financial resources. The primary CNN principle is to automatically segment and create pixel masks for each image object. A smaller VGG-like network that is inspired to distinguish between the impacts of segmentation by the VGG network family. A deep neural network is a convolutional neural network (ConvNet). It consists of several layers in sequence, including convolutional, non-linear and pooling layers, followed by one or more layers that are convolutional and totally connected.

Input to a convolutional network takes the raw pixel values of an image. In the output layer, some neurons were poised. Output layer target class corresponds to each neuron, In the Cervix type classification scheme, i.e., 3 types are Type 1/ Type 2/ Type 3, Output layer of the convolutional neural network consists of three neurons, each corresponding to one type of cervix. The ConvNet's weights (W), using a method of back propagation from the classification layer, are designed to remove errors. As shown in Figure 2.1, in the segmentation training phase, transfer learning is applied to Mask CNN weights trained using the dataset. The figure 1.2 schematic block diagram of proposed methodology.

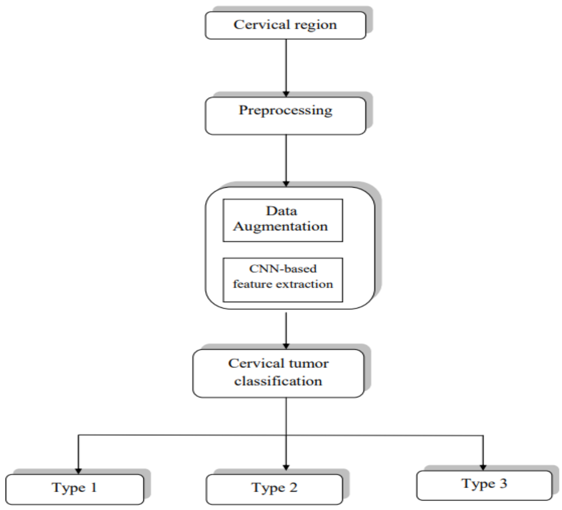

Fig 1: Block diagram of proposed system

2. IMPLEMENTATION

2.1 Image Dataset

The dataset selected has three subgroups of which total images are of about 6734 that are obtained from the Intel and MobileODT. Initially the images taken were divided into two class that is "Sharp" and "Not Sharp" . On a total of 6734 images, 1981 images were divided to as "Not Sharp" and 4753 to as "Sharp". To further information of the classes, "Sharp" means semantically focused.

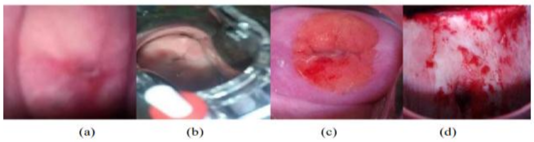

Fig 2: Instances of MobileODT (EVA) images. (a) & (b) are marked as "Not Sharp", (c) & (d) are named as "Sharp"

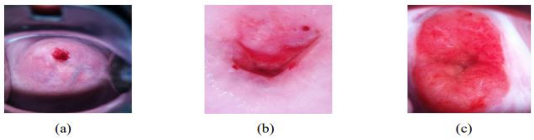

Fig 3: (a) Squamous (Type 1), (b) Adenocarcinoma (Type 2) (c) Adenosquamous Carcinoma (Type 3)

2.2 Image Preprocessing

The image area are selected by Selecting the RoI of that image which contains all the biomedical characters in it . Then the characteristics from the images are taken and the smaller sections are further used for appropriate Classification of interested region.

The images in the dataset are of the JPG array format. Each image has a image structure of 2043*1536 pixels. Further we have assigned all the images to 2043*1362 pixels to 6 squares which does the fixing with the dimension of 681*681 pixels then the square fixes the images by resizing it to 227*227 pixels.

2.3 Data Augmentation

This is a form of creating a accurate model for evaluation. It involves selection of metric for models aptitude then the model evaluation lastly create a template or a floor for execution so that the information gathered is divided into training and testing data.

Here the image clarity is obtained by demonstrating the Gaussian High Passchannel for edge framing and then identifies the interested area leaving the background.

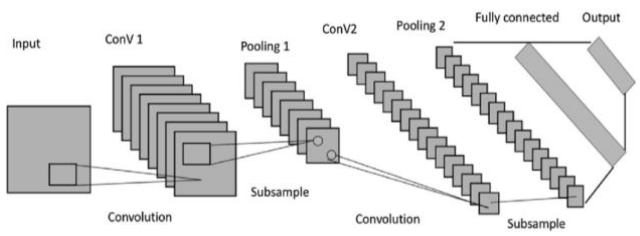

Fig 4: The Architecture of CNN

2.4 Classifier

One of the easiest approaches of solving the data problem is Convolutional Neural Networks (CNN).

It is a common method used which includes TensorFlow-Keras and many other common Frameworks it can be just done through Numpy library as well. Filters are used for feature extraction.

Each selected filters must have different functions in order to give the correct prediction class.

Pooling layers are added to reduce the parameters and convolutions are used to remove the features. Further the output layer is fully interconnected where the input is taken from the other layers and are flattened and are sent to the output to know number of classes. The output generated is compared with the error and the loss function is specified to compute the Mean Square Loss.

2.5 VGG 19(Visual Geometry Group)

This is 19 layers deep CNN+. As an input to ii a fixed size of (244,244) RGB image is given which forms the matrix (244,244,3). For spatial resolution of image, spatial padding and Max Pooling with stride 2 is performed over 2*2 pixel window which initially the kernel had 3*3 scale with a stride size of 1 pixel is now retained. Three fully connected layers were introduced of which first two were 4096 in size and then a 1000-way layer with 1000 channels. The classification of ImageNet Scale Visual Recognition Challenge (ILSVRC) is performed for evaluation and final layer is a softmax feature.

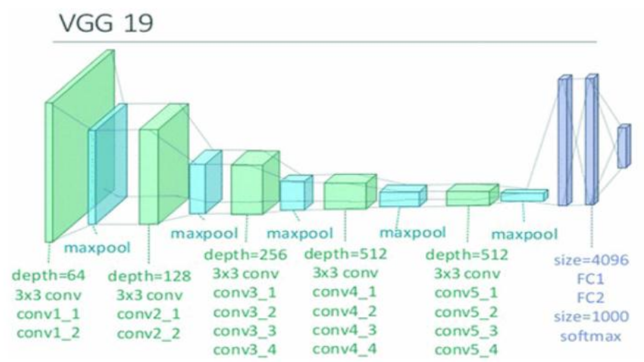

Fig 5: VGG 19 Architecture

3. CONCLUSIONS

Cervical Cancer is one of the dreadful diseases and widely spread nowadays and screening often entails lengthy clinical process, ML can provide high approach for accelerating the diagnosing process. Hence it is more convenient to get appropriate results through computer vision technology. As there is no human interference hence the results obtained are more accurate. Here the Machines are trained in such a way that for any user input the optimized results are obtained. It is more efficient and has less physical efforts.

REFERENCES

[1] Venkatesan Chandran, M. G. Sumithra, Alagar Karthick, Tony George, M. Deivakani, Balan Elakkiya, Umashankar Subramaniam and S. Manoharan, "Diagnosis of Cervical Cancer based on Ensemble Deep Learning Network using Colposcopy Images", Volume 2021, Article ID 5584004, BioMed Research International, May 04 2021, DoI: 10.1155/2021/5584004

[2] Aditya Arora, Anurag Tripathi, Anupama Bhan, "Classification of Cervical Cancer Detection Using Machine Learning Algorithms", IEEE,20 January 2021, DoI:10.1109/ICICT50816.2021.9358570

[3] Leel Sukel M K, Leel Sukesh M K, Naveen R, Theja N, "Survey on Detection Of Cervical Cancer",Volume: 06, Issue: 04 ,IRJET, Apr 2019

[4] V. Pushpalatha1, S. Sathiamoorthy and M. Kamarasan, "Cervical Cancer Detection: A Literature Survey", ISSN: 2249 - 6297, Vol. 7, The Research Publications, No. 2, 2018

[5] Vidya Kudva, Keerthana Prasad and Shyamala Guruvare ,"Automation of Detection of Cervical Cancer Using Convolutional Neural Networks", Biomedical Engineering, 46(2):135-145, 2018

[6] Rohan Gorantla, Rajeev Kumar Singh, Rohan Pandey and Mayank Jain, "Cervical Cancer Diagnosis using CervixNet - A Deep Learning Approach", 2019 IEEE 19th International on Bioinformatics and Engineering(BIBE), 2019.

[7] Akshitha Shetty, Vrushika Shah, "Survey of Cevical Cancer Prediction using Machine Learning: A Comparative approach", 18 October 2018 IEEE, DOI: 10.1109/ICCCNT.2018.8494169.

[8] Harmanpreet Kaur, Reecha Sharma, Lakhvinder Kaur, "Automated Cervical Cancer Image Analysis Using Deep Learning Techniques from Pap-Smear Images: A Literature Review", 15 November 2021 IEEE Xplore, DOI: 10.1109/ICRTO51393.2021.9596102.

[9] Enamul Karim, Nafis Neehal, "An Empirical Study of Cervical Cancer Diagnosis using Ensemble Methods", 19 December 2019 IEEE Xplore, DOI: 10.1109/ICASERT.2019.8904464.

[10] Sharma H., Zerbe N., Klempert,. et al., 2017 Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology. Comput. Med. Imaging Graph. 10.1016/j.compmedimag.2017.06.001

[11] Mavra Mehmood, Muhammad Rizwan,. et al. "Machine Learning Assisted Cervical Cancer Detection" 23 December 2021, Frontiersin ,DoI: 10.3389/fpubh.2021.788376

[12] M. Anousouya Devi, S, Ravi,. et al. "Classification of Cervical Cancer Using Artificial Neaural Networks", 2016 ScienceDirect, DoI: 10.1016/j.procs.2016.06.105

[13] Xiangyu Tan, Kexin Li,. Et al. "Automatic model for cervical cancer screening based on Convolutional neural networks: a retrospective, multicohort, multicenter study", 07 Janauary 2021. Springer, DoI: 10.1186/s12935-020-01742-6

[14] Sohely, M.D. Suimun Islam,. Et al. "Automated invasive cervical cancer disease detection at early stage through suitable machine learning model", 2021 Springer, DoI: 10.1007/s42452-021-04786-z

[15] Ye Rang Park, Young Jae Kim,. Et al. "Comparision of machine and deep learning for the classification of cervical cancer based on cervicography images", 09 August 2021, Nature, DoI: s41598-021-95748-3

[16] Penge Xue, Jiaxu Wang, .et al. "Deep Learning in image based breast and cervical cancer detection: a Systematic review and meta analysis", 15 February 2022, Nature, DoI: s41746-022-00559-z

[17] Rebecka Weegar ,Karin Sundstrom "Using machine learning for predicting cervical cancer from Swedish Electronic health records by mining hierarchical representations", 21 August 2020, DoI: 10.1371/journal.pone.0237911

[18] R. Rajpriya , Dr.M.S.Saravanan "Survey paper on cervical cancer detection through Artificial Intelligence techniques", 2022 IJCTT, DOI: 10.14445/22312803/IJCTT-V57P119

[19] D Suresh, N Jagadisha,. et al. "Detection of Brain Tumor using Image Processing", 2020 IEE, DoI: 10.1109/ICCMC48092.2020.ICCMC-000156

[20] B. Devkota, Abeer Alsadoon,. et al. "Image Segnentation for Early Stage Brain Tumor Detection using Mathematical Morphological Reconstruction" ,2018, DoI: doi.org/10.1016/j.procs.2017.12.017

[21] Neha Singh, Revati Hagir,. et al. "Automated Brain tumor detection using image preprocessing" , 24 April 2018, IJERT

[22] Aditya Garg, Roshan Lal,. et al. "Brain Tumor Detection using Image Processing Based on Anisotropic Filtration Techniques", 03 April 2020, DoI: 10.1007/978-981-15-2329-8_37

[23] Sweta Bhise, Deepmala Kale,. et al. "Breast Cancer Detection using Machine Learning Techniques", 07 july 2021, IJERT,

[24] Prerita, Ajay Rana,. et al. "Breast Cancer Detection using Machine Learning Algorithms", 15 November 2021, DoI: 10.1109/ICRITO51393.2021.9596295