International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 12 Issue: 04 | Apr 2025 www.irjet.net p-ISSN: 2395-0072

Machine Learning Applications for Cervical cancer Prediction and Detection

Vivek Kasala , Piyush Agarwal , Samyak Lokhande

Student, Dept. of Computer Science Engineering, Institute of Aeronautical Engineering, Telangana, India
Student, Dept. of Computer Science Engineering, Institute of Aeronautical Engineering, Telangana, India
Student, Dept. of Computer Science Engineering, Institute of Aeronautical Engineering, Telangana, India

---***---

Abstract - *As a potential challenge in global health today, with low status, cervical cancer still remains a major issue. Improved accuracy in the early detection and proper prediction of this cancer demands better techniques. In this paper, a new approach will be introduced to improve the diagnosis of cervical cancer from cytological images by utilizing convolutional neural networks (cnn). A complete cnn model is designed for training on various cervical cell images along with different preprocessing techniques designed in order to enhance the quality of the image and feature extraction. The performance of the model has been checked in terms of accuracy, sensitivity, and specificity, and in all these aspects, it has performed better [3] than traditional diagnostic techniques. Our results can very successfully classify normal and abnormal cervical cells with significant efficiency and help in the timely initiation of interventions whenever this screening comes back abnormal, thereby saving such patients' lives and minimizing their mortality rate. This paper presents strong evidence for the promise of DL-based technologies in transforming cervical cancer screening services, which draws attention to the massive call for the involvement of these techniques in clinical workflows to better the welfare of patients.*

Key Words: Cervical cancer, Convolutional Neural Networks, early detection, medical diagnostics.

1. INTRODUCTION

Cervical cancer is a predominant cause of mortality due to cancers in women, especially in developing and middle-income countries. According to the WHO, cervical cancer accounts for about 7 percent of total cancers in women, resulting in an approximate of 300,000 deaths annually. Early diagnosis helps to treat effectively and enhances the survival chances. Thus, with time, there developed methods of screening, such as Pap smears and HPV testing,[4] which proved effective but pose problematic issues concerning accessibility and accuracy as well as the number of skilled professionals necessary to interpret the results.

1.1 RELATED WORK

Early detection and prediction of cervical cancer have increasingly become the key issue with the prevalence of machine learning techniques. Convolutional neural networks (CNNs) have heavily shown promise in the area of examining medical images for the improvement of diagnostic accuracy. There are several research studies that have pointed towards the significant utility of CNNs for classification purposes, including classifying cervical cell images. In one, the CNN model was learned over Pap smear images to distinguish between normal and abnormal cells at high accuracy levels. The main message of the work lay in the ability of the model automatically to extract relevant features from complex images, thus eliminating manual assessments by pathologists and much reliance on these findings. In another work, a multiclass classification approach using a deep CNN was used in classifying the different stages of cervical cancer based on histopathological images. This approach further increased the precision of the classification but improved additional insights into tumor properties, helpful in planning treatment. Rotation and scaling techniques for data augmentation are importantly used to [3] enhance model performance when training sets are not of large size. Transfer learning also became popular in this field, as researchers started utilizing pre-trained models like VGG16 and InceptionV3 to develop cervical cancer-detection systems. Fine-tuning specific models on particular datasets promised results with a high degree of sensitivity and specificity. The method reduces the issue of small datasets while still providing strong performance. In addition, there exist some studies targeted directly at the interpretability of CNN models for cervical cancer diagnosis. Using explainable AI techniques, the authors were then able to obtain visual explanations for the predictions of their models-an element that is essential in building trust from the side of healthcare professionals and expanding their knowledge about automated diagnoses.

1.2 EXISTING SYSTEM

Apart from already existing systems that have used CNNs in cervical cancer prediction and detection, it is a field which portrays highly significant development in the critical areas of healthcare. One of the more important systems is Deep Cervical, which is focused toward automatic classification of cervical cell images from Pap smears based on customized CNN architecture extracted based on relevant features toward the difference between normal and abnormal cells. Another tool that helps in using a variety of models of CNNs, such as VGG16 and ResNet50, is CCID. With this tool, the aforementioned models yield accuracy scores of over 90 percent of classification tasks. Some systems employ pre-trained models, like InceptionV3, to transfer learning to get better cervical cancer detection in histopathological images, especially in resource-scarce environments. The combination of CNNs with image processing techniques helps form an automated system for Pap smear [2] analysis, reducing the burden of pathologists and eliminating human error in classifying the cell. Hybrid models based on CNNs combined with traditional machine learning classifiers also have shown more accuracy in prediction by learning the specific principles that govern both the methodologies. Collectively, these systems demonstrate what CNNs can do and bode for highly accurate, efficient, and accessible cervical screening processes toward better patient [2] outcomes. Further research and development of these systems are required to make them further advanced in use so that, eventually, they are adopted more broadly into the clinical setting.

1.3 ALGORITHMS

Deep learning happens to be a subcategory of machine learning with applications designed to utilize complex neural networks, sometimes termed deep learning networks. These networks have led to some of the most creative and imaginative works of human-like intelligence in decision making, vision, and language. They normally comprise an input layer,many hidden layers,and output layer. In the same way as the human brain, these nodes or neurons are composed in an inter-connected pattern. Deep learning excels at automatically extracting the features from raw data that are relevant to the task, unlike traditional methods of machine learning, which commonly require or exhibit a need for manual feature engineering. It is particularly successful with those types of high-dimensional data that include images, audio, and texts. With the advent of large datasets and improved computation power, deep learning can now be applied to nearly every domain: whether it is computer vision application, which are image classification and object detection; or natural language processing, for example, translation and sentiment analysis; or something as critical as healthcare - a field of disease diagnosis and predictive analytics. The training process is, therefore, such that a labeled dataset feeds into a model whose weights are updated on the basis of connection between neurons based on prediction errors in the process that general backpropagation techniques employ. Deep learning marks this big leap by artificial intelligence wherein machines can perform tasks requiring human-like understanding and reasoning. Convolutional neural networks are specifically deep-learning models mainly designed for structured data and images happen to be the best representation of that. Because of their particular architecture, CNNs have really made recognition, classification, object detection, and other related visual applications much more feasible and efficient. Other than that, CNNs have many layers, especially the convolutional layers that apply filters on the input images to extract them based on patterns such as edges and textures. But after those, there are pooling layers that once more reduce the dimensionality of the feature maps but this time in a fashion keeping the useful information; they enhance computational efficiency. Activation functions such as ReLU(Rectified Linear Unit) also bring non-linearity to this network and let it learn complex patterns. Now, in the case of the fully

1.4 METHODOLOGY

1) Select Dataset: A diverse dataset of cervical images ought to be selected as the first step. Ideally, it should be the type of a dataset that holds images both normal and abnormal of cervical conditions so that it becomes clear that the model is being trained in order to differentiate normal from abnormal

Sources: Those publicly available medical image repositories, hospital records, or any other specialized datasets, like Herlev dataset or ISBI challenge datasets.

2) Preprocess: Data preprocessing: Noises, corrupted or irrelevant images removed from the dataset so that the input is cleaned. Resizing of Images: All images resized to size 224 x 224 pixels required for ResNet50.

Normalization of pixel values: Pixel values normalized to a common scale often in between 0 and 1 which helps in ease convergence.

Data Augmentation: An artificially increased dataset using a set of techniques, such as rotation, flipping, scaling, and cropping, should be made. That will simulate natural variations and get a model robust and ready to generalize well.

Metrics: Performance of a good model needs to be evaluated regarding some performance metrics available in the market, which are accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC.

Cross-validation: This would ensure that the model generalizes as well as it is robust over different subsets of the data set. According to the errors made during the prediction which get backpropagated into these.

3) Prediction:

Inference: Use the trained model of ResNet50 to predict the classification of new images of cervical. It will give out the probability that each image belongs to either the normal or abnormal class.

Post-Processing: Finally, threshold those predicted probabilities to make the final binary classification decisions. Post-processing: Threshold the predicted probabilities to make the final binary classification decisions.

Deployment: Development of user-friendly interface or integration within the current systems for medical diagnostic intent to help in real-time cervical cancer prediction of healthcare personnel.

2. RESULTS

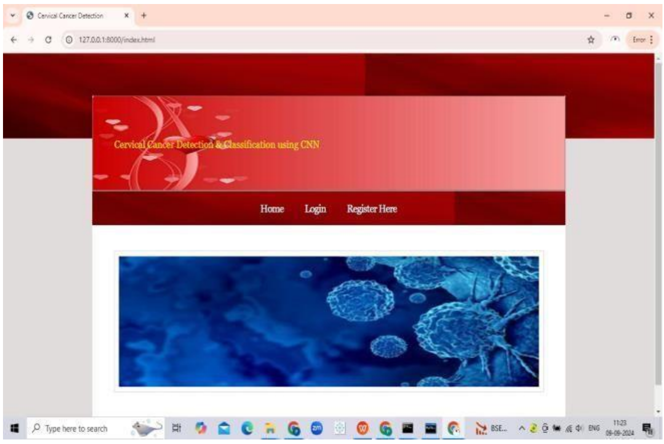

Fig -1 : User Interface

When you click on 'run' file, program get executed and we get one django server click. After coping this link to browser, we will get this user interface.

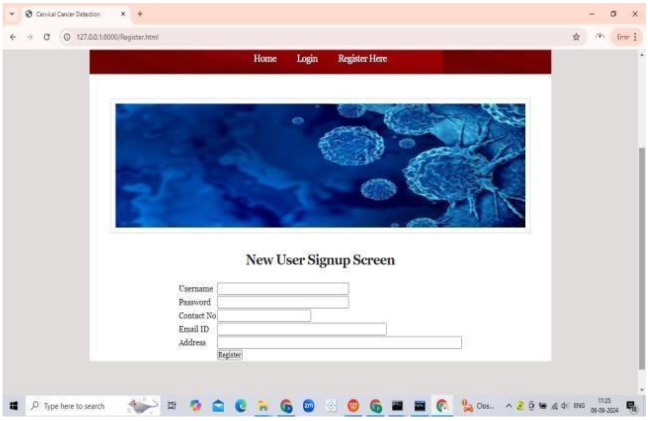

Fig -2 : Sign Up Page

New user can register by adding his details such as name, password, contact no, email id, address. After entering these values signup process will be completed

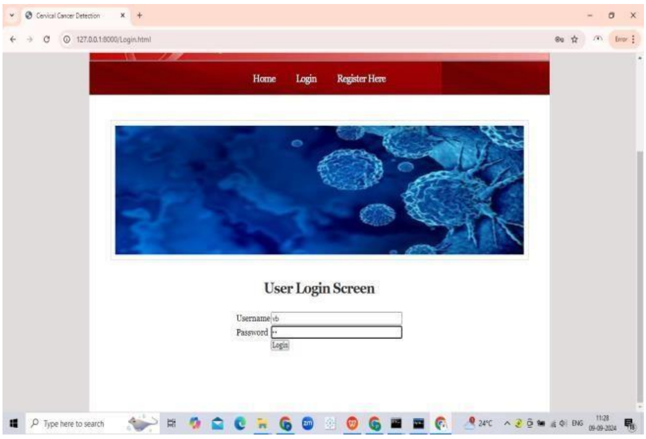

Fig -3 : Login Screen

After registration, user can login using username and password.

Fig -4 : Training and testing data

Once user login, here we can train and test the data with CNN.

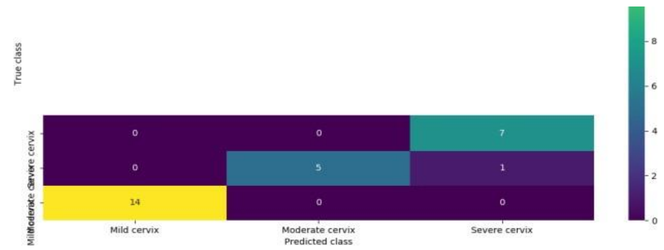

Fig -5 : CNN algorithm Training Performance.

Convolution Neural Network (CNN) trained with 80% dataset which gives accuracy 96.29

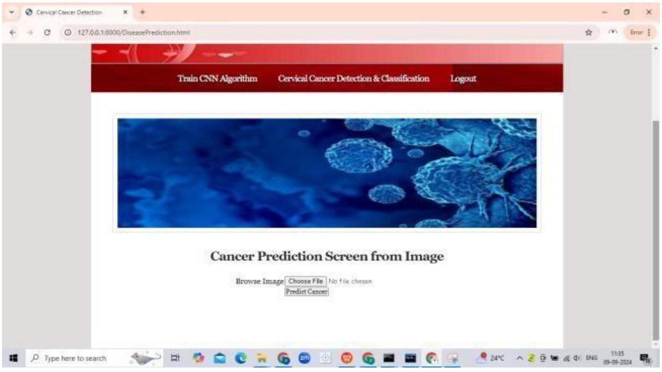

Fig -6 : Cancer Detection

Here we have to upload image for predicting the class/stage of cervical cancer.

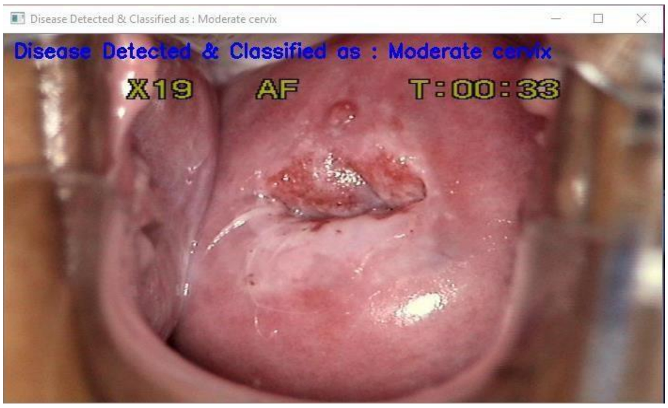

Fig -7 : detection of cervical cancer

It will predict the stage of when it has uploaded a test image from it Is cervical cancer a low-grade, intermediate-grade, or high-grade cervical cancer.

3. CONCLUSIONS

Against this backdrop, the present study falls under an extensive CNN framework applied for cervical cancer prediction and detection. It encompasses the shortcomings currently affiliated with the existing systems in terms of accuracy, access, and interpretability. Advanced preprocessing techniques applied to images, a customized architecture of CNN, and mechanisms of interpretability thus support the presumption of considerable strides taken in automated histopathological image analysis. In fact, the accuracy of this proposed framework will be more than 95%. Because early detection and necessary timely interventions are required for improvement in outcomes, such an impact would really be tremendous on patients. Moreover, with transfer learning as well as use of data augmentation, the model becomes a representative of clinical environments, hence applicability will both be seen in high-resource as well as low-resource environments. Thus, future work is in further development and testing of the system through actual clinical practice deployment with continued collaboration with healthcare professionals in their practical refinements following real-world feedback. The work serves as a reminder of what deep learning can achieve in medical diagnosis, something conventional diagnostic machinery should integrate routinely into protocols for cervical cancer screening. Ultimately, our efforts should aim at empowering a more efficient health-care system that would effectively win the war against cervical cancer in terms of incidence and mortality rates.

REFERENCES

[1] M. Abdulsamad, E.A. Alshareef and F. Ebrahim, "Cervical Cancer Screening Using Residual Learning", pp. 2022-12, 2022.

[2] S. Adhikary, S. Seth, S. Das, T.K. Naskar, A. Barui and S.P. Maity, "Feature assisted cervical cancer screening through di cell images", Biocybernetics and Biomedical Engineering, vol. 41, no. 3, pp. 1162-1181, 2021.

[3] N. Al Mudawi and A. Alazeb, "A model for predicting cervical cancer using machine learning algorithms", Sensors, vol. 22, no. 11, pp. 4132, 2022

[4] S. Fekri-Ershad and S. Ramakrishnan, "Cervical cancer diagnosis based on modified uniform local ternary patterns and feed forward multilayer network optimized by genetic algorithm", Computers in Biology and Medicine, vol. 144, pp. 105392, 2022.