Advanced Deep Learning Methods for Early Detection and Progressive Classification of Cervical Cancer

Keshav Handa
Department of Computational Intelligence
SRM Institute of Science and Technology
Kattankulathur,TamilNadu,India
kh1141@srmist.edu.in

Viabhav Sharma
Department of Computational Intelligence
SRM Institute of Science and Technology
Kattankulathur,TamilNadu,India
vs3534@srmist.edu.in

Dr. Meenakshi N
Department of Computational Intelligence
SRM Institute of Science and Technology
Kattankulathur,TamilNadu,India
meenaksn@srmist.edu.in

2024 4th International Conference on Artificial Intelligence and Signal Processing (AISP) | 979-8-3503-5065-4/24/$31.00 ©2024 IEEE | DOI: 10.1109/AISP61711.2024.10870716

***Abstract*—The proliferation of deep learning methodologies, especially convolutional neural networks (CNNs), has significantly improved the early detection and classification processes for cervical cancer, a predominant health concern for women globally. This research puts forth a sophisticated CNN model optimized to harness a comprehensive dataset enabling enhanced feature extraction from medical images, thereby elevating diagnostic precision. Incorporating transfer learning from adeptly trained models, this strategy not only augments the efficiency of our network but also accelerates the training process, thus ensuring a swifter adoption in clinical practice. The performance of the proposed model is meticulously assessed using accuracy and loss as the principal metrics. These metrics demonstrate a marked improvement in the model's ability to discern patterns indicative of cervical cancer, hence offering a superior alternative to conventional diagnostic methods. The encouraging outcomes of this study underscore the viability of deploying our CNN model as an integral component in clinical diagnostics, charting the course towards expedited and more precise patient evaluations. This investigation exemplifies the transformative impact of artificial intelligence on healthcare, signifying considerable progress in the prompt and accurate diagnosis of cervical cancer.**

***Index Terms*—Cervical cancer, convolutional neural networks, deep learning, medical image analysis, transfer learning, diagnostic precision, clinical diagnostics, artificial intelligence.**

I. INTRODUCTION

Cervical cancer remains a significant health concern, particularly in low-resource settings where traditional diagnostic methods such as Pap smears are still widely used. These methods, while effective, are often plagued by issues like subjectivity, high variability in sensitivity and specificity, and a heavy reliance on the expertise of the cytologist. Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have offered a promising alternative by automating the detection and classification processes. However, existing CNN models often face critical challenges, including sensitivity to image artifacts, the need for extensive preprocessing, and difficulties in generalizing across diverse datasets, which limit their applicability in real-world clinical settings. To address these gaps, our research introduces a robust CNN-based model tailored for the early detection and progressive classification of cervical cancer from cytology images. Unlike previous models that struggle with image quality variations and require significant manual intervention, our model employs advanced data augmentation techniques to enhance robustness and integrates dropout and batch normalization to prevent overfitting. Furthermore, by leveraging transfer learning from pre-trained networks like AlexNet and ResNet, as well as a custom-designed ManualNet, our approach reduces the dependency on large annotated datasets and accelerates the training process. These innovations allow our model to achieve superior accuracy in distinguishing between different stages of cervical cancer.

II. RELATED WORK

Recent advancements in the application of convolutional neural networks (CNNs) to medical imaging, particularly in the detection and classification of cervical cancer, highlight significant strides in AI-assisted diagnostics. Below is a detailed analysis of recent contributions in this area, as documented by several key studies: [1] This study evaluates the effectiveness of various machine learning algorithms against deep learning models in cervical cancer detection. The findings emphasize the superior performance of CNNs due to their ability to automatically learn complex features directly from medical images. Unlike the machine learning models reviewed in this paper, our CNN model directly extracts features from raw images without the need for manual feature selection, offering superior accuracy and automation. [2] This paper presents a novel approach by integrating Conditional Random Fields (CRFs) with CNNs to improve the segmentation of cervical images, which is critical for precise cancer localization and classification. While this study uses CRFs for better segmentation, our model simplifies the pipeline by integrating feature learning and classification in one step, reducing complexity and potentially increasing scalability. [3] The development of CervixNet, a specialized CNN for cervical cytology, showcases advancements in using deep learning for nuanced medical tasks. The study highlights the use of image augmentation to enrich training datasets, leading to improvements in model sensitivity and specificity. Similar to CervixNet, our model leverages deep learning for image analysis but extends its applicability by employing transfer learning, which was crucial for improving performance given a limited dataset. [4] Reviews the deployment of various deep learning models in cervical cancer screening. Both studies utilize transfer learning; however, our model incorporates more recent architectures and training techniques that enhance learning efficiency and diagnostic precision. [5] This research compares multiple CNN architectures, including AlexNet and GoogleNet, for their efficiency in diagnosing cervical cancer. Results favor ensemble models that consolidate predictions from several CNNs to enhance diagnostic accuracy. This paper evaluates multiple CNN architectures, while our research focuses on optimizing a single, robust model using advanced regularization and optimization strategies to better handle overfitting and improve generalization. [6] Explores the application of transfer learning, where models developed for general tasks are fine-tuned for specific medical imaging challenges, significantly reducing the need for large labeled datasets. Both models use pre-trained networks, but our approach updates the architecture to better accommodate the specific characteristics of cervical cytology images. [7] Investigates a combination of level set segmentation and Support Vector Machines for effective cervical cell classification, stressing the importance of accurate cell boundary delineation. Our model avoids the multi-stage process of segmentation and classification by integrating these tasks, offering a more streamlined and computationally efficient approach. [8] Focuses on a deep learning framework specifically designed to recognize various stages of cervical cancer, which is vital for appropriate treatment planning. While both models are aimed at stage detection, our CNN model is trained with a novel cost function that specifically addresses the imbalanced nature of stage-wise data in cervical cancer. [9] Provides a comprehensive overview of the state-of-the-art techniques employed in the automation of cervical cancer detection through deep learning, highlighting ongoing challenges and future directions. [10] Discusses an innovative method that employs independent level sets for segmentation, followed by multi-class Support Vector Machines for the classification of cervical cells, offering another layer of accuracy in detecting morphological changes associated with cancer progression. Unlike the hybrid approach of using level sets and SVMs, our model relies solely on deep learning, which simplifies the model training and deployment in real-world clinical settings. [11] utilized the ResNet101v2 architecture, demonstrating its efficacy in classifying cervical cells into seven distinct classes with a 71.4% accuracy using the Herlev dataset. This study underscored the potential of deep learning models in reducing human error in cytological diagnostics by enhancing the precision of pap smear tests. [12] Proposed a CNN model for cervical cell classification that not only categorizes cells into normal and abnormal but also differentiates various severity levels of dysplasia using a public health dataset. Their method achieved a commendable classification accuracy, thereby providing a robust tool for early cervical cancer detection. [13] Highlighted the importance of detailed segmentation in cervical cytology, particularly using CNNs for boundary detection which plays a crucial role in the precise classification of the severity of cell abnormalities. Their approach significantly improved the detection rates of high-grade squamous intraepithelial lesions (HSIL), which are critical precursors to invasive cervical cancer. [14] presented a comprehensive review of image-based deep learning approaches for cervical cancer screening, emphasizing the advancements in neural network architectures and their applications in medical imaging. Their evaluation of various models illustrated significant improvements in diagnostic accuracies and patient outcomes. [15] explored the inception recurrent residual convolutional neural network (IRRCNN) for cervical cancer prediction, which incorporates both residual and recurrent layers to enhance feature extraction and learning capabilities. This model was particularly noted for its high accuracy and low false positive rates in preliminary testing.

III. PROPOSED MODEL

We introduce a robust convolutional neural network (CNN) designed for early detection and classification of cervical cancer from cytology images, leveraging deep learning's capacity for enhanced feature extraction through transfer learning from pre-trained networks. This approach is crucial for discerning subtle image features critical for accurate medical diagnoses. Our CNN architecture includes multiple convolutional layers adept at identifying and localizing specific patterns and textures related to different cervical cancer stages. To ensure the model generalizes well on unseen data and to prevent overfitting, we incorporate regularization strategies such as dropout and batch normalization. These techniques stabilize the learning process and improve performance across varied datasets. Empirical evaluation is conducted using a dataset of 2000 training images and 1997 test images, encompassing five classes. This extensive dataset enables the model to learn diverse features, enhancing its diagnostic capabilities. We measure the model's effectiveness using accuracy and loss metrics, which preliminarily show high accuracy and suggest the model's potential to significantly advance cervical cancer diagnostics and clinical decision-making. The architecture of our proposed model "Fig. 1." serves as a systematic framework for the early detection and classification of cervical cancer, beginning with a robust dataset of cervical images. These images are first expanded through augmentation techniques, which simulate various visual perspectives, and then meticulously annotated to provide accurate labels for supervised learning. This preparatory stage ensures that our model has a diverse and rich dataset to learn from, increasing its ability to recognize a range of patterns in cervical cell images. Our framework divides the prepared dataset into training and testing sets, which are essential for both learning and evaluating the

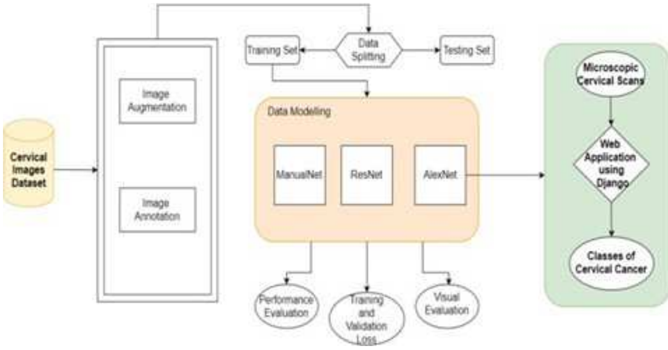

Fig. 1. Architecture of Proposed Model

model's performance. The training set enables the model to identify and learn distinguishing features of different cervical cancer stages, while the testing set is crucial for assessing the model's diagnostic capabilities on new, unseen data. The heart of the model lies within its data modeling phase, where different neural network architectures are employed. These include a custom-designed ManualNet and established architectures like ResNet and AlexNet. These architectures are instrumental in feature extraction, a key process that directly influences the accuracy of the classification task.

Using accuracy and loss metrics, we can rigorously assess the model's predictions against the actual classifications. The final stage integrates the model into a Django-based web application, offering an accessible interface for users to upload microscopic cervical scans and receive classification into one of the five defined categories. This end-to-end process not only demonstrates the practical application of the model but also showcases the potential impact of AI in enhancing cervical cancer diagnostics.

A. Data Acquisition and Curation

This stage requires the careful collection and annotation of a wide range of MRI images, which are categorized into different groups that correspond to the disease's stages (Stages 1 through 4) as well as a healthy baseline. In order to support the predictive ability of the model, the dataset—which should ideally consist of approximately 4000 annotated images—needs to be carefully partitioned in order to separate distinct sets for model training and performance assessment.

B. Implementation of ManualNet

The ManualNet we propose is a convolutional neural network (CNN) framework meticulously engineered for discerning the intricate variances inherent in cervical cytology imagery. This bespoke architecture, rooted in deep learning principles, harmonizes a series of convolutional layers—each followed by rectified linear unit (ReLU) activations—with strategically integrated pooling layers, which function to downscale the feature maps thereby condensing the data for more efficient processing. Interspersed within this structure are dropout layers, a quintessential regularization technique designed to curb the model's tendency to overfit by randomly omitting units during the training process, thus enhancing its ability to generalize when exposed to novel datasets. The ManualNet's architecture can be represented by mathematical formulas that encapsulate the operations within its layers. The convolutional layers perform discrete convolutions, which can be expressed as:

$$f_{ij}^k = \Phi\left(\sum_m \sum_n I_{(i+m)(j+n)} W_{mn}^k + b^k\right) \tag{1}$$

where f_{ij}^k is the activation of the k-th feature map at location (i, j), I is the input image, W is the kernel for the k-th feature map, b^k is the bias term for the k-th feature map, and Φ denotes the ReLU activation function.

Pooling layers, typically max-pooling, are utilized to reduce the spatial dimensions of the feature maps. This operation can be described by:

$$p_{ij}^k = \max_{(m,n)\in M_{ij}} f_{mn}^k \tag{2}$$

where p_{ij}^k is the output of the k-th feature map in the pooling layer, and M_{ij} represents the pooling region around location (i, j). During the dropout phase, the dropout function $D(\cdot)$ can be defined as:

$$y_i = D(x_i, p) = \begin{cases} 0 & \text{with probability } p \\ \frac{x_i}{1-p} & \text{with probability } (1-p) \end{cases} \tag{3}$$

where x_i is an input to a neuron, y_i is the output after dropout, and p is the dropout probability. The training process includes data augmentation techniques such as image rescaling, shearing, zooming, and horizontal flipping to introduce variability and improve the network's robustness.

For compilation, the model employs the RMSprop optimizer and categorical cross-entropy loss function to fine-tune its weights:

$$\text{Loss} = -\sum_{c=1}^{M} y_{o,c} \log(p_{o,c}) \tag{4}$$

where $y_{o,c}$ is a binary indicator of whether class label c is the correct classification for observation o, and $p_{o,c}$ is the predicted probability of observation o being of class c.

C. Implementation of ResNet

The ResNet architecture "Fig. 2." innovatively addresses vanishing gradients by introducing residual connections that allow direct gradient flow. This is critical for deep networks, particularly for intricate tasks like medical image classification. The architecture starts with an initial convolutional layer followed by layers of residual blocks, each consisting of batch-normalized and ReLU- activated convolutions. These residual blocks employ identity mapping to facilitate training deeper networks without performance degradation. They use the principle:

$$F(x) = H(x) - x \tag{5}$$

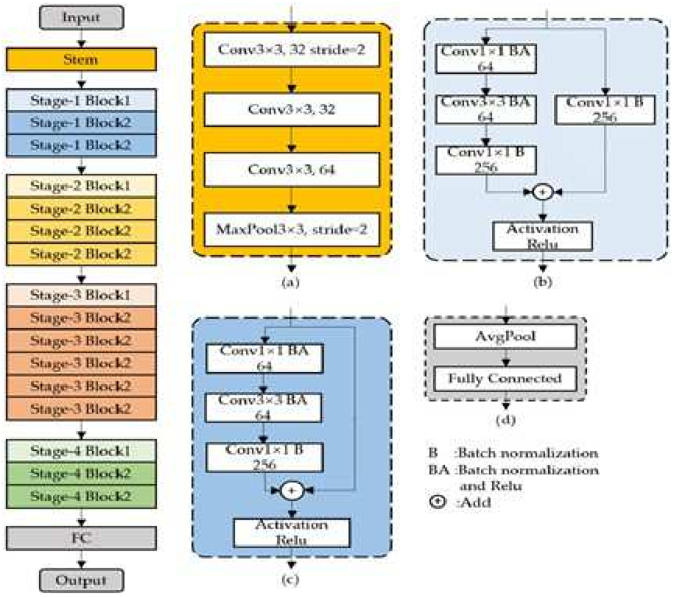

Fig. 2. Architecture of ResNet

where $F(x)$ is the residual mapping to be learned and $H(x)$ is the desired output. The ResNet structure is completed with a global average pooling layer and a fully connected layer with a softmax function, yielding a probability distribution over the classes.

D. Implementation of AlexNet

With its deep architecture, AlexNet ”Fig. 3,” a groundbreaking CNN for image recognition, dramatically reduces the dimensions of the input image while increasing the feature space. It starts with an input of 224x224x3, to which a large 11x11 convolutional filter is applied at a stride of four. This quickly compresses the spatial dimensions and increases the depth. The network is further deepened by subsequent convolutional layers that use smaller 5x5 and 3x3 filters. Intermediate max pooling layers lower dimensions and help with spatial translation invariance. To introduce non-linear transformations necessary for complex pattern learning, the layers make use of ReLU activations. During training, this model adjusts weights for classification across five possible outcomes by utilizing the Adam optimizer and categorical cross entropy. During training, dropout, applied at a rate of $p = 0.5$, keeps neurons from co-adapting and possibly overfitting. By randomly setting some of the neurons’ outputs to zero, this regulation method thins the network momentarily. Upon flattening the convolutional output, the architecture proceeds to pass it through densely connected layers, ultimately resulting in a softmax classifier. This classifier produces a probability distribution across classes, which represents the likelihood of each class given the input image.

E. Training the Model

Start training the Convolutional Neural Nets (CNNs) that we have chosen, as shown in our proposed architecture ”Fig. 1.”, with the aggregated dataset that is unique to our project. Refine the network weights iteratively by using the ADAM optimization technique. Throughout the procedure, consistently compute the categorical cross-entropy loss to measure the prediction error of the model and employ accuracy as a metric to appraise and track its predictive efficacy.

Fig. 3. Architecture of AlexNet

F. Testing the Model

Testing the CNN model to assess its generalization ability and diagnostic accuracy on unobserved data is a crucial step after the training phase. To do this, the test dataset that was not seen by the network during training must be fed into the model. Accuracy, precision, recall, and the area under the receiver operating characteristic (ROC) curve are among the performance metrics that are typically collected during this phase. These metrics provide important information about how robust and reliable the model is for use in practical settings. Through a thorough examination of the differences between expected results and observed labels, you can pinpoint possible areas for enhancement and confirm that the model is suitable for implementation in clinical environments.

G. Detection and classification

In assessing the predictive performance of our convolutional neural network (CNN) tailored for cervical cancer classification, the model’s deep learning layers scrutinize test images to assign one of five stages, indicating progressive disease or healthy tissue. This prediction phase is critical, not only to gauge the model’s immediate diagnostic accuracy but also to predict its long-term reliability in clinical settings. The model’s predictions are quantified through metrics such as accuracy and loss, providing a numerical testament to its capability to potentially enhance clinical outcomes and aid in the early and precise detection of cervical cancer, thereby improving treatment approaches.

H. Deployment

In the deployment phase of our cervical cancer classification project, the fully trained CNN model is seamlessly integrated into a Django-based web application framework. This application is specifically structured to receive user-uploaded cervical cytology images, which the model then processes to predict and classify into predefined cancer stages. Upon submission, images are instantly processed, leveraging the model's deep learning algorithms to provide real-time diagnostic predictions. The deployment is designed to be user-centric, focusing on providing a streamlined experience for clinicians and researchers to utilize the model's capabilities through a straightforward web interface, ultimately aiming to contribute to the advancement of medical diagnostics.

TABLE I
COMPARATIVE ANALYSIS

Proposed System	**Existing System (KGNMDA - Model)**
The proposed approach makes use of innovative advances in tech nology and deep learning algo rithms to determine the existence of cervical cancer directly.	The system employs the KGNMDA model to construct a Microbe-Disease Knowledge Graph (MDKG) from assorted biomedical databases. It leverages a graph neural network (GNN) to refine the prediction of microbe-disease associations.
Trained on 2000 images and tested on 1997 images across five classes.	Contains 5,628 cervical cancer-related differentially methylated items.
The method includes developing a web application and using advanced technology to identify stages, which is followed by recommendations for prevention and treatment.	The system predicts cervical cancer using a Graph Neural Network (GNN) to enhance microbe-disease association predictions from a Microbe-Disease Knowledge Graph (MDKG) sourced from various biomedical databases.
Gives 91.25% accuracy and can do comparative analysis on different CNN models.	Researchers manually validate the data accuracy using the HMDAD database, without any preprocessing of images.

IV. EXPERIMENTAL RESULTS

A. ManualNet

During the training process spanning 10 epochs and utilizing a batch size of 512, the custom-designed ManualNet architecture demonstrated a training accuracy peak of approximately 63%, as observed in the accuracy graph referred to as "Fig. 4.".

B. ResNet

The deep residual learning framework of the ResNet model is used in this project to improve the classification of cervical cancer stages from cytological image data. The model was subjected to a rigorous training regime over 100 epochs with a batch size of 512. This allowed the model to refine its feature detection capabilities, resulting in an accuracy of 88.5%.

C. AlexNet

Our selected model, AlexNet, trained over 100 epochs, attained a training accuracy of 91.25% and precision of 92.51%, with a validation accuracy of 86.04%. These metrics showcase its robustness in cervical cytology image classification, making it the preferred choice for our diagnostic study.

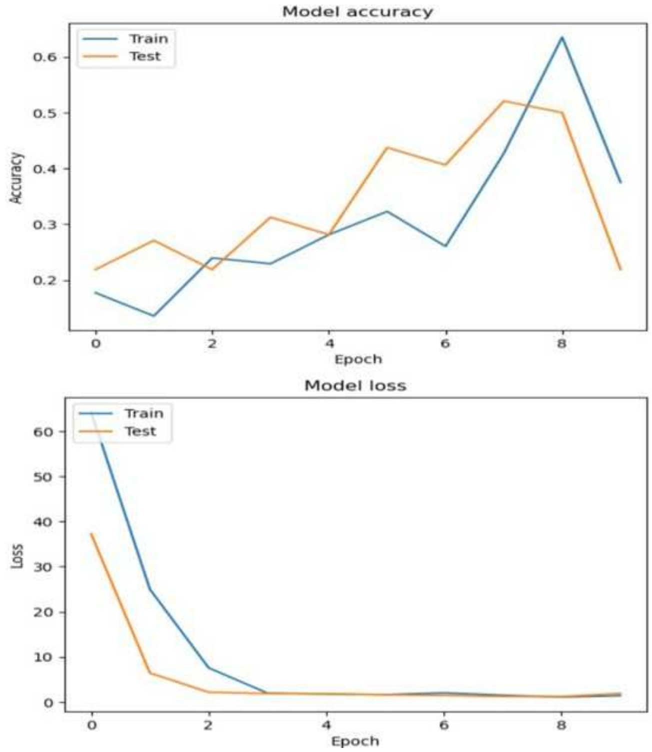

Fig. 4. Accuracy and Loss of ManualNet Model

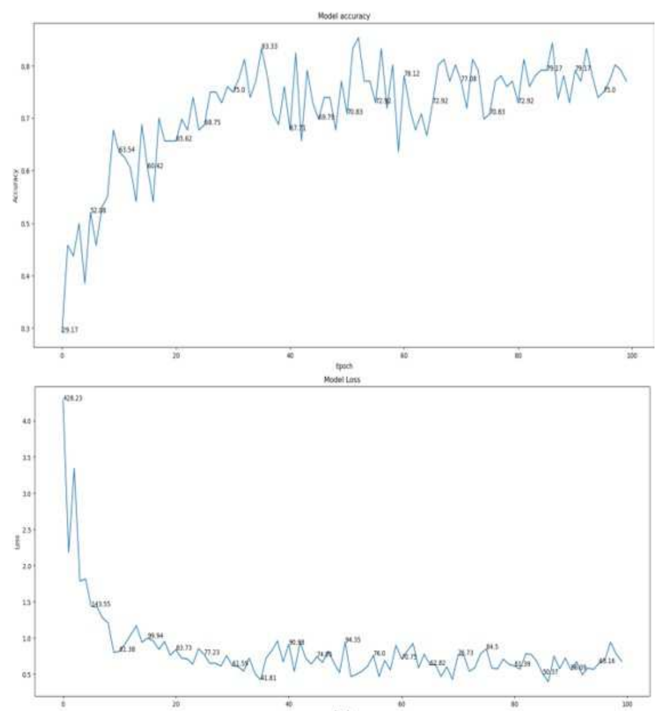

Fig. 5. Accuracy and Loss Graph of ResNet Model

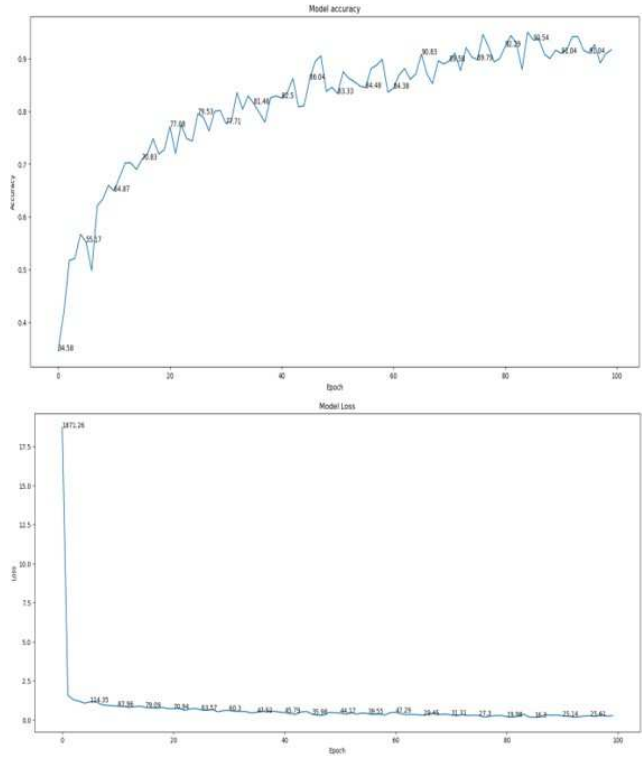

Fig. 6. Accuracy and Loss Graph of AlexNet Model

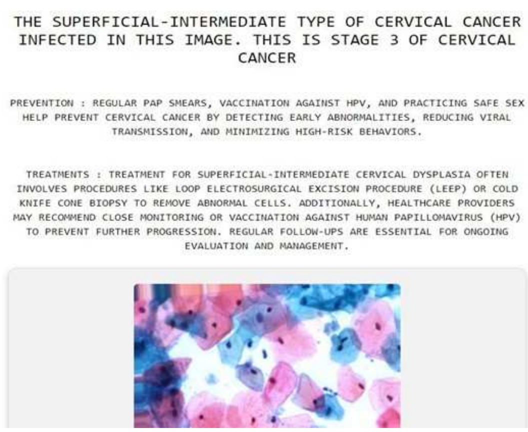

Fig. 7. Cervical Cancer Detection and Classification

V. CONCLUSION

In conclusion, AlexNet demonstrated strong effectiveness for cervical cancer detection due to its robust architecture and success in image classification tasks. Its deep convolutional layers enable extraction of complex features critical for distinguishing cancer stages, while ReLU activations prevent the vanishing gradient problem and dropout techniques reduce overfitting, enhancing generalization. Compared to other architectures like VGGNet and ResNet, AlexNet offered an optimal balance of computational efficiency and accuracy, making it particularly suited for large-scale analysis. While deeper models like VGGNet provide finer feature extraction, they involve higher computational costs, making AlexNet the more practical choice. These findings highlight AlexNet's potential as a reliable tool for early detection and classification of cervical cancer, with scope for further refinement.

REFERENCES

[1] A. Arora, A. Tripathi, and A. Bhan, "Classification of Cervical Cancer Detection using Machine Learning Algorithms," ICICT, 2021, doi: 10.1109/ICICT50816.2021.9358570.

[2] V. D. Soni and A. N. Soni, "Cervical cancer diagnosis using CNN with conditional random field," ICIRCA, 2021, pp. 1749-1754, doi: 10.1109/ICIRCA51532.2021.9544832.

[3] R. Gorantla et al., "Cervical Cancer Diagnosis using CervixNet - A Deep Learning Approach," BIBE, 2019, pp. 397-404, doi: 10.1109/BIBE.2019.00078.

[4] Khoulqi and N. Idrissi, "Deep learning-based Cervical Cancer Classification," ICTIH, 2022, pp. 30-33, doi: 10.1109/ICTIH57289.2022.10112111.

[5] S. N. Kong et al., "Cervical Cancer Identification using Deep Learning Approaches," ICCED, 2022, pp. 1-4, doi: 10.1109/ICCED56140.2022.10010544.

[6] H. K and V. Vetriselvi, "Deep Learning-based Classification of Cervical Cancer using Transfer Learning," ICESIC, 2022, pp. 134-139, doi: 10.1109/ICESIC53714.2022.9783560.

[7] D. Kashyap et al., "Cervical cancer detection using Independent Level sets and multi SVMs," TSP, 2016, pp. 523-528, doi: 10.1109/TSP.2016.7760935.

[8] P. Chatterjee, "Detection of Stages of Cervical Cancer using Deep Learning," ICAIS, 2023, pp. 488-493, doi: 10.1109/ICAIS56108.2023.10073730.

[9] O. E. Aina, S. A. Adeshina, and A. M. Aibinu, "Deep Learning for Cervical Cancer Detection and Diagnosis — A Survey," ICECCO, 2019.

[10] D. Kashyap et al., "Cervical cancer detection using Independent Level sets and multi SVMs," TSP, 2016, doi: 10.1109/TSP.2016.7760935.

[11] S. N. Kong et al., "Cervical Cancer Identification using Deep Learning Approaches," J. Med. Syst., vol. 34, no. 6, pp. 1043–1058, 2010.

[12] L. Zhang et al., "Deep Learning for Image-based Cervical Cancer Detection," IEEE J. Biomed. Health Inform., vol. 21, no. 6, pp. 1633–1643, 2017.

[13] H. A. Phoulady et al., "A deep learning approach for the analysis of masses in mammograms," Med. Image Anal., vol. 45, pp. 169–185, 2018.

[14] B. Taha et al., "Deep Learning Algorithms for Detection of Critical Findings in Head CT Scans: A Review," IEEE Access, vol. 7, pp. 115754–115768, 2019.

[15] M. Z. Alom et al., "The History Began from AlexNet: A Survey on Deep Learning Approaches," arXiv preprint arXiv:1803.01164, 2018.