A Comprehensive Review of Deep Learning and Machine Learning Approaches for Cervical Cancer Detection: Challenges, Advancements, and Future Directions

Vivek Baraiya[1], Nikunj Siddhpura[2], Shivani Bhamani[3], Rutvi Baraiya[4]
Nikul Zinzuvadiya[5]
[1,2]*Assistant Professor, Dept. of Computer Application, Gyanmanjari Innovative University*
[3]*Lecturer, Dept. of Information Technology, Gyanmanjari Innovative University*
[4,5]*Assistant Professor, Dept. of Computer Engineering, Gyanmanjari Innovative University*

***Abstract*—Cervical cancer remains a leading cause of death among women worldwide, making early detection vital for effective treatment and improved survival rates. This review paper explores recent advancements in artificial intelligence (AI), particularly deep learning (DL) and machine learning (ML) models, applied to cervical cancer detection. We examine a variety of state-of-the-art models, including Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and hybrid models, assessing their efficacy in medical image classification tasks. The challenges of data imbalance, model interpretability, and resource limitations are discussed in depth. Furthermore, we propose future research directions aimed at improving diagnostic accuracy, model generalization, and deployment in low-resource settings. This paper highlights the transformative potential of AI in revolutionizing cervical cancer screening and diagnosis, offering a path toward more accessible and accurate healthcare solutions.**

I. INTRODUCTION

Cervical cancer is one of the most prevalent cancers among women, particularly in low- and middle-income countries (LMICs), where access to screening and treatment services is limited. According to the World Health Organization (WHO), nearly 604,000 new cases of cervical cancer were reported globally in 2020, with over 341,000 deaths [5]. The high mortality rate is largely attributed to the late-stage diagnosis of the disease. Early diagnosis and timely intervention are crucial in improving survival rates; however, traditional screening methods, such as the Papanicolaou (Pap) smear test and Human Papillomavirus (HPV) testing, have notable limitations. These conventional methods are often prone to human error, require significant time for interpretation, and depend heavily on the skill and experience of healthcare professionals [4], [3]. In many rural and under-resourced areas, the lack of trained pathologists and proper screening facilities leads to a significant diagnostic gap [1].

Traditional screening techniques, including Pap smear tests, involve microscopic examination of cervical cells, but these methods are subjective and can be inaccurate, especially when performed by inexperienced personnel. Similarly, HPV testing, which detects the virus responsible for the majority of cervical cancers, also faces challenges in terms of sensitivity, false positives, and the lack of immediate, actionable results [5]. Additionally, these methods often require advanced infrastructure and trained medical professionals, which are not always available in LMICs [6].

Recent advancements in Artificial Intelligence (AI), particularly Machine Learning (ML) and Deep Learning (DL), have demonstrated remarkable potential to overcome these limitations. AI technologies can automate the process of medical image analysis, enabling more consistent and accurate detection of cervical cancer. By leveraging large datasets, AI models can identify subtle patterns in cervical images patterns that might be difficult or even impossible for human eyes to detect [3], [7]. AI-based systems have the ability to learn complex features from images through training on labeled datasets, improving diagnostic accuracy, reducing human error, and enabling faster decision-making [4].

Deep learning models such as Convolutional Neural Networks (CNNs) have been particularly successful in medical image analysis, achieving higher accuracy than traditional methods in a variety of applications, including cervical cancer detection. Studies have shown that CNNs can detect abnormal cervical cells with a high degree of precision, automating the process of image segmentation and classification [8]. Additionally, newer models like Vision Transformers (ViTs) have emerged as strong competitors to CNNs, particularly in their ability to capture global dependencies within images and improve model performance across diverse datasets [5], [7].

Despite the advancements in AI, several challenges persist in the adoption of these models for clinical use. One major obstacle is data quality and availability. Many deep learning models require large, high-quality datasets for training, and such datasets are often scarce in medical fields, especially for rare diseases like cervical cancer [3], [6]. Moreover, class imbalance remains a significant issue, as cervical cancer datasets often contain far more normal samples than abnormal ones, leading to biased models that perform poorly on underrepresented classes [2].

Another significant challenge is the interpretability of AI models. While deep learning models can achieve high accuracy, their "black-box" nature makes it difficult for clinicians to trust and understand their decision-making process. Explainable AI (XAI) methods, such as Grad-CAM and LIME, have been integrated into these models to provide visual explanations of predictions, improving transparency and fostering greater trust in automated diagnostic tools [4], [7].

Finally, the real-world deployment of AI models faces barriers such as computational cost, hardware limitations, and the need for regulatory approval. Models that require extensive computational resources may not be feasible for use in resource-limited settings. Moreover, integrating AI-based systems into existing healthcare workflows requires careful consideration of both technical and organizational factors [6], [1].

This review paper explores the state-of-the-art AI methodologies used for cervical cancer detection, with a specific focus on deep learning approaches like CNNs and ViTs. We examine the various challenges these models face in clinical settings and discuss possible solutions to enhance performance and generalization across diverse populations. Additionally, the paper highlights potential future directions in AI research for cervical cancer diagnosis, emphasizing the need for multi-modal data integration, federated learning, and AI deployment in low-resource settings.

The structure of the paper is as follows: Section II reviews the key deep learning models used in cervical cancer detection, including CNNs, ViTs, and hybrid models. Section III outlines potential future research directions, with a focus on improving model generalization and efficiency for global use. Finally, Section VII concludes the paper by summarizing the key findings and offering suggestions for future research.

II. LITERATURE REVIEW

A. Deep Learning Models in Cervical Cancer Detection

AI models, particularly deep learning techniques, have shown significant promise in the domain of cervical cancer detection. These models are capable of learning complex patterns from large medical datasets and can be classified into several categories based on the type of data they analyze and the architecture they employ.

1) Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) have been widely used for medical image classification due to their ability to automatically extract features from images at different levels. CNNs excel in identifying hierarchical features, making them well-suited for image analysis tasks, especially in medical fields where pixel-level patterns are crucial. In the context of cervical cancer, CNNs have been successfully used to classify Pap smear images, HPV test results, and colposcopy images.

A study by Ishak Pacal et al. demonstrated the successful application of CNNs for cervical cancer diagnosis using Pap smear images, achieving high classification accuracy [8]. These models excel in detecting patterns in pixel data and have been successful in automating the interpretation of cytology slides, which are traditionally labor-intensive and prone to human error.

To better understand how CNNs operate, Figure 1 shows the general architecture of a CNN model used for cervical cancer detection. This architecture illustrates the various layers involved in processing and classifying cervical cancer images.

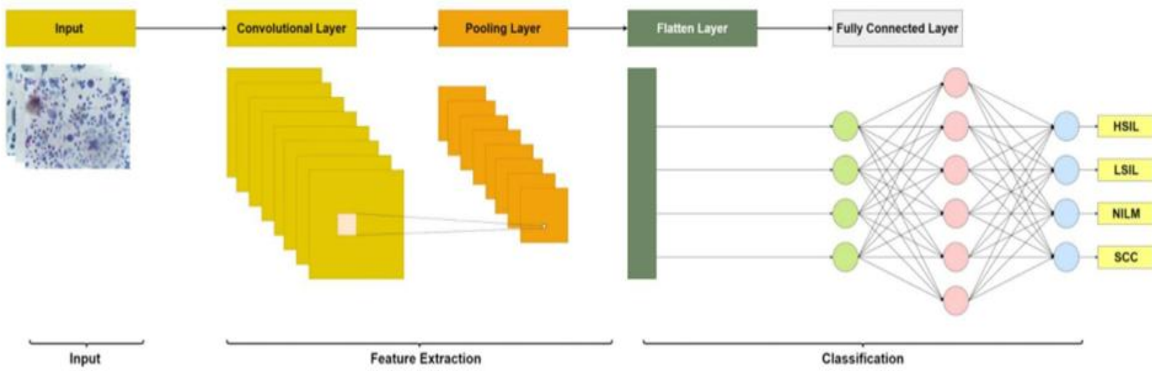

Fig.1 CNN Architecture for Cervical Cancer Detection

2) Vision Transformers (ViTs)

Recent advancements in Vision Transformers (ViTs) have shown that they can outperform CNNs in several image classification tasks. ViTs apply a transformer-based architecture, originally designed for natural language processing (NLP), to image data. The ViT model processes images in patches, allowing it to capture long-range dependencies within the data, which CNNs may miss due to their limited receptive field. This ability to capture both local and global contextual information makes ViTs particularly suitable for complex medical image analysis. In a study conducted by Md. Humaion Kabir Mehedi, ViTs achieved 99.48% classification accuracy in cervical cancer detection, outperforming CNN-based models such as EfficientNetV2 [14]. ViTs have the advantage of being more flexible and capable of handling complex image relationships, which can lead to improved diagnostic accuracy in the detection of cervical cancer.

To visualize the workflow of Vision Transformers and their integration with other deep learning models, Figure 2 presents the architecture and stages of model training, from the dataset preprocessing to the final classification using Vision Transformers.

3) Hybrid Models

Hybrid models, which combine the strengths of both machine learning (ML) and deep learning (DL) techniques, have also shown promise in cervical cancer detection. Sandeep Kumar Mathivanan and colleagues introduced a hybrid deep learning model that fuses CNNs with random forests (RF) to improve classification accuracy [7]. Such hybrid models leverage deep learning for feature extraction, which is effective for capturing intricate patterns in medical images, while using machine learning models like random forests for classification, which are well-known for their robustness and ability to handle small datasets. The fusion of DL and ML techniques leads to enhanced performance and robustness in detecting cervical cancer, particularly in settings where data may be scarce or imbalanced.

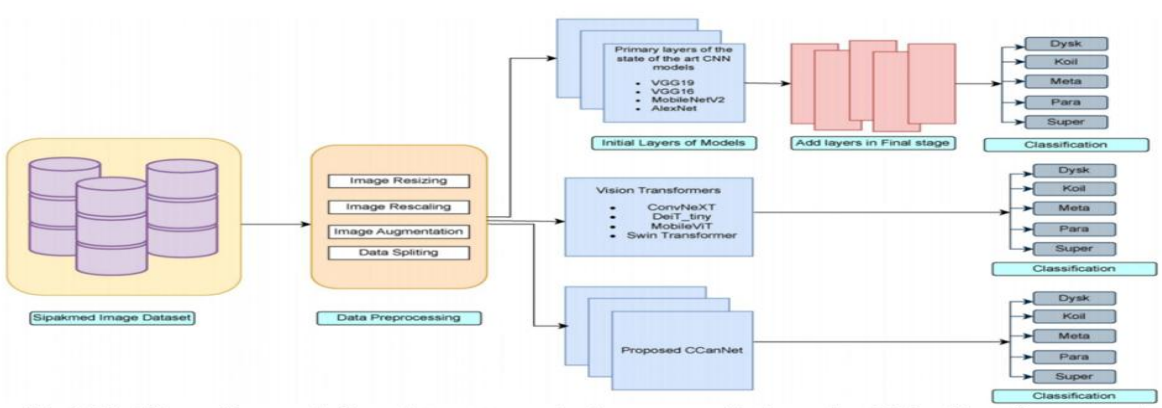

Fig.2 Workflow of research from dataset to cervical cancer prediction using Vision Transformers and CNN models.

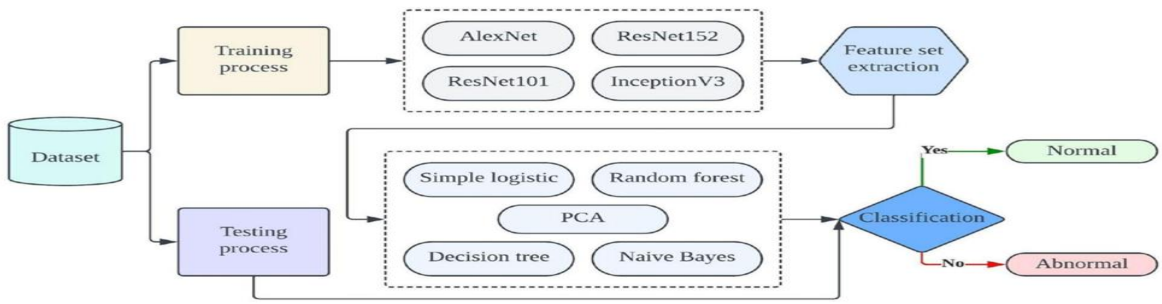

Fig.3 Proposed model for cervical cancer classification using hybrid ML and DL techniques.

To better understand how these hybrid models operate, Figure 2 illustrates the proposed workflow for cervical cancer classification, where deep learning models like CNNs and ViTs are used for feature extraction, and machine learning models such as Random Forests and Naive Bayes are used for classification. The workflow also includes a feature set extraction process, showcasing the integration of both ML and DL methods for efficient classification.

B. Challenges in Cervical Cancer Detection

While the application of AI in cervical cancer detection has shown promise, several challenges remain in the field. One of the primary issues is the lack of large, high-quality datasets for training deep learning models. Many available datasets are small, with limited examples of abnormal cases, leading to class imbalance that can skew model performance. Data augmentation techniques, such as rotation, flipping, and cropping, have been employed to address this issue, allowing models to generalize better across diverse data. Additionally, transfer learning has become a popular method to address the data scarcity problem, where models pre-trained on large datasets (such as ImageNet) are fine-tuned on smaller medical datasets to improve accuracy and performance [6].

Another significant challenge is model interpretability. In medical applications, it is crucial to understand why a model made a particular decision, especially when it is used for critical health diagnoses. Deep learning models, especially CNNs and ViTs, are often considered "black-box" models, making their decision-making process difficult to explain. To address this challenge, Explainable AI (XAI) techniques such as Grad-CAM and LIME have been integrated into deep learning models to provide insights into model predictions and increase trust in automated systems [5].

C. Performance Evaluation of Models

A key aspect of any AI model is its performance, which is typically evaluated using metrics such as accuracy, precision, recall, and F1-score. These metrics help determine how well the model can classify cervical cancer and other related diseases. In the paper by Sohely Jahan et al., the use of Multilayer Perceptron (MLP), Random Forest, and other traditional machine learning algorithms demonstrated the trade-off between computational efficiency and diagnostic accuracy [13]. Traditional models like MLP are computationally less demanding but may not offer the same level of accuracy as more complex models like CNNs and ViTs.

In contrast, CNNs and ViTs often require more computational resources but can achieve significantly higher accuracy, making them suitable for clinical settings where the necessary infrastructure is available. Table 1 below compares the performance of different models in cervical cancer detection based on commonly used metrics.

1) Comparative Performance of Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
CNN (EfficientNet V2)	98.5	98.2	98.3	98.25
ViT (DeiT-tiny)	99.48	99.4	99.5	99.45
Hybrid CNN+RF	98.08	97.9	98.0	97.95
MLP	97.4	97.3	97.6	97.45

Table 1: Performance comparison of different models based on accuracy, precision, recall, and F1-score.

As shown in Table 1, ViTs outperform CNN-based models like EfficientNetV2 and traditional models like MLP and Random Forest, providing the highest accuracy and recall, which makes them a promising choice for cervical cancer detection. Hybrid models, while slightly less accurate than ViTs, offer a good balance between performance and computational efficiency.

III. FUTURE DIRECTIONS

To improve cervical cancer detection, future research should focus on the following areas:

A. Federated Learning

Federated learning offers a promising approach for training machine learning models across multiple decentralized devices or servers without the need to exchange sensitive patient data. This method allows hospitals and medical institutions to collaborate on training robust models while ensuring that patient data remains private and secure. The implementation of federated learning can overcome data privacy concerns, especially in sensitive medical domains like cervical cancer detection, where sharing patient data across institutions could raise ethical and legal challenges. This approach can also lead to better generalization of models by leveraging diverse data sources without compromising patient confidentiality.

B. Multi-modal Data Integration

One of the most exciting avenues for enhancing cervical cancer detection is the integration of multi-modal data. Combining image data with clinical, genomic, and behavioral data can significantly improve the accuracy of predictive models. Medical images, such as Pap smears and HPV tests, can be enriched with additional data such as patient history, genetic factors, and lifestyle information. By fusing these different types of data, machine learning models can develop a more comprehensive understanding of a patient’s condition, improve diagnostic accuracy and enable earlier detection. Multi-modal models have the potential to provide a more holistic view of the patient's health, resulting in better treatment planning and personalized care.

C. Explainable AI

As AI models are increasingly used in healthcare, the need for explainable AI (XAI) becomes critical. In medical applications, it is not enough for a model to provide accurate predictions; it is equally important that healthcare professionals understand why a model arrived at a particular decision. Explainable AI techniques, such as Grad-CAM and LIME, can provide insights into the decision-making process of deep learning models. These techniques allow clinicians to visualize the regions of a medical image that contributed to a diagnosis, making AI decisions more interpretable and trustworthy. The development of more transparent AI systems will help bridge the gap between AI models and clinical practitioners, fostering greater acceptance and integration of AI in medical workflows.

D. Edge Computing

The advent of edge computing has opened new possibilities for real-time diagnostics, particularly in low-resource settings. Developing lightweight machine learning models that can run on mobile devices or low-cost infrastructure would enable healthcare providers in remote and underserved areas to diagnose cervical cancer without the need for expensive equipment or cloud-based services. Edge computing brings computation and data storage closer to the location where it is needed, reducing latency and dependency on high-bandwidth internet connections. By deploying AI models directly on mobile phones or small devices, real-time cervical cancer detection becomes feasible in rural and underserved areas, where access to centralized healthcare services may be limited. This approach can greatly increase the accessibility and affordability of early-stage cervical cancer screening and diagnosis.

IV. CONCLUSION

AI, especially deep learning and machine learning, has the potential to greatly improve cervical cancer detection by automating the analysis of medical images and making diagnoses more accurate. This can lead to earlier detection, which is critical for improving survival rates and outcomes. However, to make the most of these technologies, there are still challenges to overcome, such as the need for more data, making AI models easier to understand, and ensuring these models can be used effectively in real-world clinical settings.

Looking ahead, research should focus on developing more reliable AI models, improving how we share medical data, and making sure AI is used ethically in healthcare. By improving model transparency and combining different types of patient data, we can make these tools more reliable and accessible, helping to detect cervical cancer early and ultimately saving more lives.

REFERENCES

[1] K. P., M. Premkumar, "An In-Depth Exploration of CNN-Based Deep Learning Models in Cervical Carcinoma Analysis," Advances in Biomarker Sciences and Technology, doi: 10.1016/j.abst.2025.11.007.

[2] S. Faisal, "Comparative Analysis and Machine Learning Predictions of Cervical Cancer Incidence: A Multi-National Study," BRAC University, 2023.

[3] M. Rahman, J. Khan, A. S. M. Mostafizur Rahman, R. Haque, M. S. Ali, "A Critical Analysis of Machine Learning and Deep Learning Methods for Cervical Cancer Screening," Journal of Computer and Communications, vol. 11, pp. 64-85, Dec. 2023, doi: 10.4236/jcc.2023.1112005.

[4] M. Dellino, M. Cerbone, A. d'Amati, M. Bochicchio, A. S. Laganà, A. Etrusco, A. Malvasi, A. Vitagliano, V. Pinto, E. Cicinelli, G. Cazzato, E. Cascardi, "Artificial Intelligence in Cervical Cancer Screening: Opportunities and Challenges," AI, vol. 5, pp. 2984–3000, Dec. 2024, doi: 10.3390/ai5040144.

[5] M. Xu, C. Cao, P. Wu, X. Huang, D. Ma, "Advances in Cervical Cancer: Current Insights and Future Directions," Cancer Communications, vol. 45, pp. 77–109, Jan. 2025, doi: 10.1002/cac2.12629.

[6] P. Garg, M. Krishna, P. Kulkarni, D. Horne, R. Salgia, S. S. Singhal, "Machine Learning Models for Predicting Gynecological Cancers: Advances, Challenges, and Future Directions," Cancers, vol. 17, pp. 2799, Aug. 2025, doi: 10.3390/cancers17172799.

[7] S. K. Mathivanan, D. Francis, S. Srinivasan, V. Khatavkar, K. P., and M. A. Shah, "Enhancing Cervical Cancer Detection and Robust Classification through a Fusion of Deep Learning Models," Scientific Reports, vol. 14, Article 10812, 2024, doi: 10.1038/s41598-024-61063-w.

[8] I. Pacal, "Investigating Deep Learning Approaches for Cervical Cancer Diagnosis: A Focus on Modern Image-Based Models," European Journal of Gynaecological Oncology, vol. 46, no. 1, pp. 125-141, 2025, doi: 10.22514/ejgo.2025.012.

[9] A. Gupta, A. Parveen, A. Kumar, P. Yadav, "Advancement in Deep Learning Methods for Diagnosis and Prognosis of Cervical Cancer," Current Genomics, vol. 23, pp. 234-245, 2022, doi: 10.2174/1389202923666220511155939.

[10] S. S. Abrar, S. A. M. Isa, S. M. Hairon, M. P. I. Ismail, M. N. B. N. A. Kadir, "Recent Advances in Applications of Machine Learning in Cervical Cancer Research: A Focus on Prediction Models," Obstetrics & Gynecology Science, vol. 68, no. 4, pp. 247-259, 2025, doi: 10.5468/ogs.25041.

[11] N. Zinzuvadiya, C. Parida, P. K. Samanta, A. Dash, J. J. Jena, and S. Darshana, "X-GAN: Explainable Generative Adversarial Networks for Rare Disease Data Augmentation and Clinical Insights," 4th International Conference on Distributed Computing and Electrical Circuits and Electronics (ICDCECE), 2025, pp. 1-6, doi: 10.1109/ICDCECE65353.2025.11035270.

[12] N. Zinzuvadiya, S. R. Mishra, H. Mohapatra, and P. Mohapatro, "Detecting AI-Generated Images Using Contrastive Learning with Momentum Contrast (MoCo) Framework," International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), Bhubaneswar, India, 2025, pp. 1-6, doi: 10.1109/ASSIC64892.2025.11158049.

[13] S. Jahan, M. D. S. Islam, L. Islam, T. Y. Rashme, A. A. Prova, B. K. Paul, M. D. M. Islam, and M. K. Mosharof, "Automated invasive cervical cancer disease detection at early stage through suitable machine learning model," SN Applied Sciences, vol. 3, Article 806, Sep. 2021, doi: 10.1007/s42452-021-04786-z.

[14] M. H. K. Mehedi, M. Khandaker, S. Ara, M. A. Alam, M. F. Mridha, and Z. Aung, "A lightweight deep learning method to identify different types of cervical cancer," Scientific

Reports, vol. 14, Article 29446, 2024, doi: 10.1038/s41598-024-79840-y.