International Journal of Advanced Engineering Research and Science (IJAERS)
Peer-Reviewed Journal
ISSN: 2349-6495(P) | 2456-1908(O)
Vol-12, Issue-12; Dec, 2025
Journal Home Page Available: https://ijaers.com/
Article DOI: *https://dx.doi.org/10.22161/ijaers.1212.5*

Deep Learning Based Histopathological Classification of Cervical Cancer Using YOLO-v8 and Inception-v3: A Comparative Performance Study

Usman Idris Ismail[1]*, Joshua C Shawulu[1], Mustapha Ismail[2], Albashir Ahmad[3], Sulaiman Yusuf Ali[3], Nuhu Abdulalim Muhammad[1], Saadatu Ali Jijji[1]

[1]Department of Computer Science, Federal University Kashere, Nigeria
Email: usmanidris@fukahsere.edu.ng*, joshua.shawulu@aun.edu.ng, elgurama1989@gmail.com, Saadatuali@fukashere.edu.ng
[2]Department of Computer Science, Gombe State University, Nigeria
Email : mismail@gsu.edu.ng
[3]Department of Computer Science, Federal Polytechnic Kaltungo, Nigeria
Email : albashirelnafaty@gmail.com, yusufsulaiman766@gmail.com
*Corresponding Author

Received: 12 Sep 2025,

Received in revised form: 25 Nov 2025,

Accepted: 01 Dec 2025,

Available online: 17 Dec 2025

Keywords*— *Cervical Cancer Classification, Deep Learning, YOLO-v8, Inception-v3, Histopathology Images, Medical Image Analysis

***Abstract*—** *Cervical cancer remains a major health concern among women worldwide, emphasizing the need for accurate and efficient diagnostic approaches. Deep learning has shown strong potential in automating medical image analysis and improving diagnostic reliability. This study aims to evaluate and compare the performance of two advanced deep learning models YOLO-v8 and Inception-v3 for the multi-class classification of cervical cancer histopathology images. A curated dataset of 225 histopathology images representing three cervical cancer subtypes (squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma) was preprocessed using standardized resizing, normalization, and augmentation techniques. Both models were trained with optimized hyperparameters and evaluated using accuracy, precision, recall, F1-score, confusion matrices, and learning curve analysis to determine their effectiveness and generalization capability. The proposed workflow incorporates robust preprocessing, extensive augmentation, and systematic hyperparameter tuning to enhance model performance. YOLO-v8 leverages an efficient unified architecture for high-speed feature extraction, while Inception-v3 utilizes multi-scale convolutional processing to capture fine-grained morphological patterns within histopathology images. YOLO-v8 achieved an accuracy of 99.8% and Inception-v3 achieved 99.4%, demonstrating strong discriminative ability and reliable classification across all cancer subtypes. These results highlight the potential of deep learning models as effective tools for automated cervical cancer diagnosis. Despite the limited dataset size, the study provides a solid performance benchmark and establishes a foundation for future work incorporating larger datasets and multimodal diagnostic frameworks.*

I. INTRODUCTION

Cervical cancer remains a major public health challenge globally and is currently recognized as the fourth most prevalent cancer among women [1]. The disease originates within the cervix at the lower narrow portion of the uterus that connects to the vaginal canal and typically progresses from persistent infection with high-risk human papillomavirus (HPV) strains, which disrupt normal cellular processes and lead to precancerous and cancerous lesions [2].

Despite improvements in early screening and diagnostic techniques, cervical cancer continues to impose a significant burden, particularly in low- and middle-income regions where healthcare disparities persist. According to recent global statistics, over 604,000 new cases and 341,831 deaths were recorded in 2020, with some regions including sub-Saharan Africa, Melanesia, and Southeast Asia experiencing disproportionately higher mortality rates due to limited access to screening and specialized care [3].

Current diagnostic procedures including the Papanikolaou (Pap) test and HPV testing, play vital roles in early detection by identifying cellular abnormalities and viral infections associated with cervical carcinogenesis [4].

However, imaging remains a cornerstone of comprehensive clinical assessment, particularly for staging, treatment planning, and monitoring disease progression. Modalities such as computed tomography (CT), magnetic resonance imaging (MRI), and positron emission tomography (PET) provide valuable three-dimensional anatomical detail and functional insights into tumor behavior [5]. MRI is especially advantageous due to its superior soft-tissue contrast and ability to differentiate subtle structural variations in cervical tissues. It is therefore widely regarded as the imaging modality of choice for evaluating tumor size, invasion into surrounding structures, parametrial involvement, and lymph node enlargement [6]. Studies have demonstrated that advanced MRI techniques, such as diffusion weighted imaging (DWI), significantly enhance diagnostic accuracy by improving sensitivity and specificity in detecting metastatic spread [7]

Despite the advantages of medical imaging, accurate interpretation requires substantial radiological expertise. Human limitations such as inter observer variability, challenges in distinguishing subtle pathological patterns, and variations in imaging protocols can influence diagnostic performance [8]. As cervical cancer incidence remains high in resource constrained regions, the need for reliable, automated, and scalable diagnostic support systems has become increasingly urgent. Artificial intelligence (AI), particularly deep learning, has emerged as a transformative solution capable of addressing these challenges.

Deep learning models, especially convolutional neural networks (CNNs), have achieved state-of-the-art performance across numerous medical imaging tasks, including segmentation, detection, and classification in radiology, oncology, and digital pathology [9]. CNN based architectures such as ResNet, VGG, MobileNet, Inception-v3, and Xception have been successfully applied to cervical cytology and MRI datasets, demonstrating remarkable effectiveness in identifying cervical abnormalities and predicting treatment outcomes [10]. These advancements highlight the potential of deep learning to reduce diagnostic errors, support clinical decision-making, and improve global cervical cancer management.

However, despite significant progress, several gaps remain. Many existing studies rely on limited datasets, lack diversity in imaging modalities, or focus exclusively on cytology images, which limits their generalizability. Additionally, comparative analyses of different deep learning architectures particularly involving real-time detection models like the YOLO family remain scarce in cervical cancer imaging research. YOLO-v8, the latest version in the YOLO series, introduces enhanced capability for rapid object detection and classification, offering a compelling alternative to traditional CNN classifiers. In contrast, Inception-v3 is known for its multi-scale feature extraction and strong performance in complex image classification tasks. A systematic comparison of these two architectures has the potential to reveal important insights into their diagnostic suitability, computational efficiency, and classification robustness.

Therefore, this study aims to investigate and compare the performance of YOLO-v8 and Inception-v3 for the classification of cervical cancer images. The dataset used in this work comprises three clinically relevant cancer types squamous cell carcinoma, adenocarcinoma, and Aden squamous carcinoma each of which presents distinct morphological features when viewed under medical imaging. By analyzing model behavior across these cancer types, this study seeks to evaluate the ability of both architectures to detect nuanced differences in cervical tissue appearance.

The paper aims to leverage recent advances in deep learning and computer vision to develop an effective and reliable framework for classifying cervical cancer images using state-of-the-art neural network architectures. Specifically, the study evaluates and compares the performance of YOLO-v8 and Inception-v3 to determine

their suitability for accurate and automated cervical cancer diagnosis. The main contributions of this work are as follows:

- This work presents a deep learning–based methodology using YOLO-v8 and Inception-v3 for multi-class classification of cervical cancer histopathology images, incorporating systematic preprocessing and optimized hyperparameters to enhance performance.
- This study utilizes a curated histopathology dataset consisting of three cervical cancer subtypes, providing a reliable benchmark for evaluating modern deep learning models in automated cancer diagnosis.
- A comprehensive comparison was carried out between YOLO-v8 and Inception-v3, highlighting differences in architectural design, feature extraction capabilities, and classification behavior.
- The experimental results demonstrate that YOLO-v8 and Inception-v3 achieved high diagnostic accuracy, outperforming traditional convolutional methods reported in the literature and confirming their suitability for automated cervical cancer detection.

II. RELATED WORKS

In this section, we review recent research efforts dedicated to the detection and classification of cervical cancer using image-based machine learning and deep learning techniques. Prior studies have explored a range of approaches including traditional machine learning, radiomics, and modern convolutional neural networks to enhance diagnostic accuracy, support clinical decision-making, and improve early detection outcomes. However, many of these studies face persistent challenges such as limited dataset sizes, retrospective data collection, and inconsistencies in image acquisition protocols, which constrain the generalizability of their findings.

A study in [10] introduced ConvXGB, a hybrid method combining convolutional neural networks with eXtreme Gradient Boosting for assessing recurrence risk in cervical cancer patients. Using 406 multiparametric MRI images from three institutions, the model outperformed radiomics and clinical baselines with AUC values of 87.2% and 88.2% for 1-year and 3-year recurrence-free survival, respectively. Despite its promising results, reliance on a moderate-sized dataset and manual segmentation procedures introduced variability that affected overall efficiency. Similarly, the work in [11] developed a radiomics-based framework to predict lymph node metastasis (LNM) using 153 MRI scans, focusing on T2-weighted sequences and ADC maps. Feature selection via LASSO and classification using an SVM yielded AUC scores of 80.4% and 81.1% across training and validation sets. Although incorporating clinical variables improved prognostic accuracy (C-index of 91.6%), the study remained limited by non-uniform MRI protocols and small sample size.

In another contribution, Laura [12] examined the prognostic value of normalized tumor ADC measurements in cervical cancer using 179 MRI images. Their findings showed that normalizing tumor ADC values improved disease-specific survival prediction, though the overall accuracy remained modest at 68%. Limitations included heterogeneous imaging systems and the wide temporal span of data collection, which introduced variability due to evolving imaging standards. The study in [13] proposed a delta-radiomics model to predict radiation proctitis using 126 MRI images collected before and after radiotherapy. Using logistic regression, Pearson correlation, and LASSO for feature engineering, the model achieved a validation accuracy of 90%. Nonetheless, the model's reliability was affected by single-center data collection and dependence on traditional machine learning techniques.

Efforts to classify cervical cancer directly from MRI images were demonstrated in [14], where 900 cancerous and 200 non-cancerous images were used to evaluate models including VGG16, CNN, KNN, and RNN. Extensive preprocessing improved model performance, with VGG16 achieving the highest accuracy at 95.44%. Despite strong results, dataset imbalance, reliance on pre-trained weights, and variability in MRI acquisition posed limitations. A similar direction was explored in Qin’s work [15], which applied deep multiple-instance learning (D-MIL) using 392 MRI scans to predict LNM. Based on ResNet-50 feature extraction, the model achieved AUC values between 71.4% and 76.5%, and performance improved further when combined with clinical factors. However, retrospective design and limited sample size reduced its applicability in clinical practice.

Further studies strengthened evidence for MRI's effectiveness in cervical cancer diagnosis. Research in [16] reaffirmed MRI as the gold standard for local staging, with high-resolution T2-weighted sequences achieving 88% accuracy and DWI enhancing sensitivity and specificity in lymph node detection. Limitations included high costs, longer scanning times, and reduced accuracy in deep pelvic regions. Work in [17] compared handcrafted radiomics with deep learning radiomics (DLR) models for predicting chemoradiotherapy response in 252 patients. The DLR model significantly outperformed traditional methods, achieving 73.2% accuracy and improving to 77.7% when clinical variables were included. Yet, absence of external validation and limited dataset size restricted its broader adoption.

Radiomics-based detection of lymph-vascular space invasion (LVSI) was explored in [18], where mRMR and LASSO were applied to select features from 177 multiparametric MRI images. The resulting nomogram achieved AUC scores of 83.8% and 83.7% for training and testing, with modest classification accuracy. Another study in [19] applied CNNs including MobileNetV3, Xception, and Inception-v3 to classify MRI scans from multiple cancer types, with MobileNetV3 achieving an accuracy of 86%. The authors noted that computational complexity and variability in MRI imaging conditions posed major challenges.

More recent work has shifted toward ensemble and optimization-driven deep learning frameworks. Studies such as Cervi-Net integrated multiple CNNs (e.g., DenseNet169, MobileNetV2, DenseNet201) using grid-search optimization, achieving accuracies up to 97.94% on cytology datasets [20]. Advanced ensemble approaches enhanced with the Salp Swarm Algorithm (SSA) further improved diagnostic performance, reaching accuracies of 99.48% and 95.23% across different datasets [21]. Similarly, a Differential Evolution (DE)-based ensemble incorporating feature fusion, ConvLSTM, and SE modules reached accuracies nearing 99% on several cervical cytology benchmarks [22]. These developments highlight the increasing emphasis on multi-model integration and advanced optimization strategies to refine feature representation and improve diagnostic accuracy. The existing studies demonstrate strong potential for machine learning and deep learning in cervical cancer diagnosis. However, many rely on MRI, cytology, or radiomics-specific designs, and few provide direct comparisons of modern architectures such as YOLO-v8 and Inception-v3 for image-based cervical cancer classification. This gap underscores the relevance of the present study, which aims to evaluate and compare these contemporary models to enhance automated cervical cancer detection.

III. MATERIALS AND METHODS

This section presents a comprehensive approach to cervical cancer classification using a curated dataset of histopathology images representing three major cancer types: squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma. The input images underwent several preprocessing operations to enhance quality and consistency across the dataset, including resizing, normalization, and data augmentation techniques designed to improve model generalization. These steps ensured that all samples adhered to the input size and numerical stability requirements of the selected deep learning architectures.

Two models were employed in this study. The first, YOLO-v8, was adapted in classification mode to leverage its advanced feature extraction backbone, enabling efficient learning of localized discriminative patterns within cervical cancer images. The second model, Inception-v3, utilizes multi-scale convolutional blocks to analyze complex textural and morphological variations across image regions. Together, these architectures provide complementary strengths: YOLO-v8 offers high detection-grade feature sensitivity, while Inception-v3 provides robust multi-resolution classification capabilities.

To further enhance model reliability, data augmentation was applied to introduce controlled variability including rotation, contrast adjustment, and geometric transformations thus reducing overfitting and improving robustness to visual disparities commonly found in real-world datasets. Hyperparameters such as learning rate, batch size, and epoch count were iteratively tuned to ensure optimal training behavior for both models. The proposed methodology enabled a structured comparison between YOLO-v8 and Inception-v3, highlighting their respective performance in multi-class cervical cancer classification. This combination of advanced deep learning techniques and systematic preprocessing establishes an efficient diagnostic pipeline capable of supporting automated cervical cancer detection.

3.1 Data Description

The dataset used in this study consists of cervical cancer histopathology images categorized into three clinically significant groups: squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma. These categories represent the most prevalent pathological subtypes observed in clinical diagnosis and exhibit varying morphological features that are essential for automated classification. All images were curated from verified medical image sources and underwent expert validation to ensure accurate labeling and exclusion of low-quality samples. The dataset was partitioned into training, validation, and testing subsets following an 80:10:10 ratio to enable comprehensive performance evaluation. Similar to the constraints observed in existing cervical imaging datasets, variations in image appearance, staining quality, and structural patterns emphasize the importance of consistent preprocessing and robust model training.

3.2 Preprocessing

Preprocessing plays a crucial role in ensuring that the histopathology images used in this study are consistent, clean, and suitable for deep learning analysis. To meet the input requirements of the selected architectures, all images were first resized, with YOLO-v8 receiving inputs of 640 × 640 pixels and Inception-v3 resized to 299 × 299 pixels.

Following resizing, pixel values were normalized to the range of 0-1 by dividing each value by 255, a standard technique that improves numerical stability and accelerates model convergence during training. To overcome the limitations of the small dataset and enhance the models' ability to generalize, various data augmentation techniques were applied, including rotation, flipping, zooming, and adjustments to brightness and contrast. These augmentations introduce variability that mimics real-world differences in histopathology slide preparation and imaging conditions. Finally, the dataset was systematically organized into training, validation, and testing subsets to ensure a reproducible workflow and standardized evaluation process. Together, these preprocessing steps strengthened image quality, improved visual consistency, and enabled the models to better capture meaningful morphological patterns during feature learning.

3.3 Model Architecture

Two advanced deep learning architectures were selected to evaluate automated cervical cancer classification:

3.3.1 YOLO-v8

YOLO-v8 represents the latest evolution of the YOLO (You Only Look Once) family of deep learning models, designed to deliver high accuracy and real-time performance for image classification and object detection tasks. Convolutional layers form the foundation of YOLO-v8, enabling the extraction of hierarchical features such as edges, textures, and complex spatial patterns properties essential for accurate medical image classification. Owing to its balance of speed, precision, and computational efficiency, YOLO-v8 has become a widely preferred model for applications ranging from automated diagnostics to surveillance and autonomous systems.

The architecture of YOLO-v8 is built around a modular and scalable design consisting of three primary components: the backbone, neck, and head as shown in fig. 1 below. The backbone is responsible for extracting multi-level visual features from the input image using modern convolutional strategies such as CSP (Cross-Stage Partial) networks that improve gradient flow and reduce computational redundancy. The neck aggregates feature across scales, enhancing the model's ability to detect patterns of varying sizes and complexity. Finally, the head generates classification outputs (and bounding box predictions in detection mode), ensuring fast and accurate inference across different tasks. This modularity enables YOLO-v8 to scale efficiently across variants from lightweight "YOLO-v8-tiny" models to larger, high-capacity versions allowing flexibility depending on resource availability and performance needs.

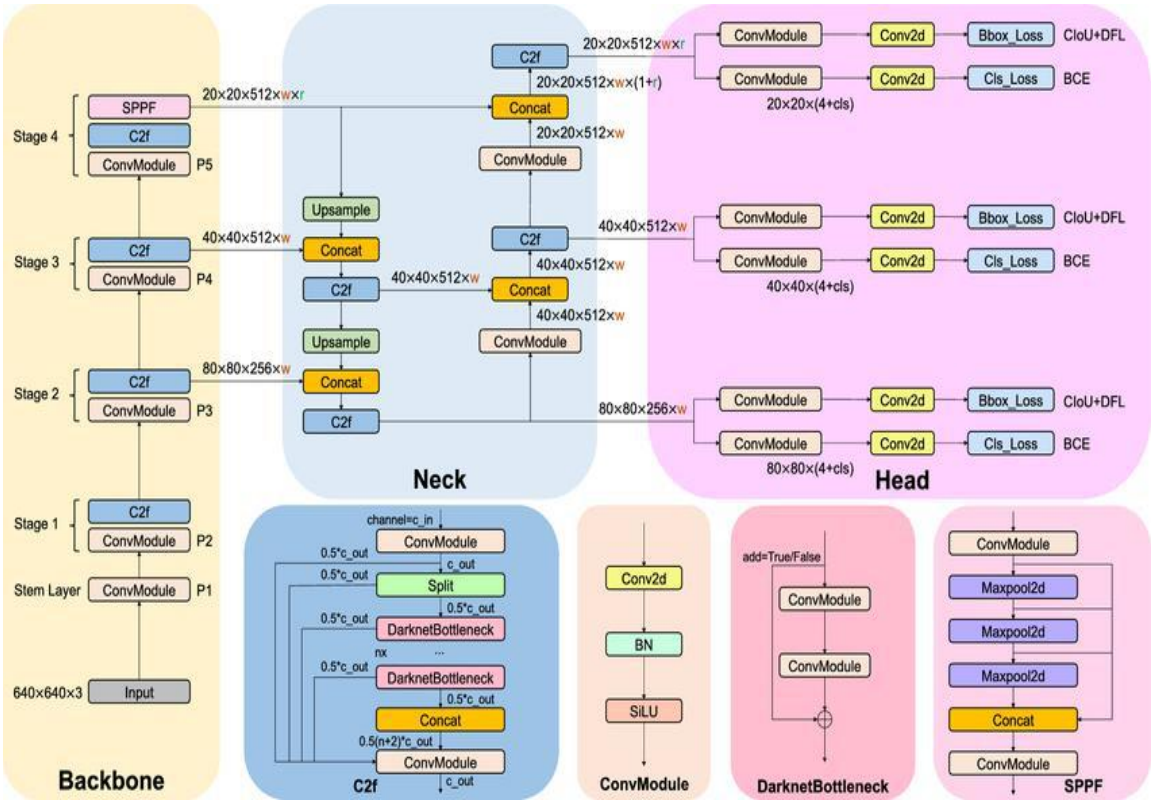


Fig. 1: Yolo-v8 Model Architecture

YOLO-v8 also introduces several advancements in training methodology. It supports both anchor-based and anchor-free prediction strategies, incorporates optimization techniques such as RAdam for stable convergence, and offers improved data augmentation pipelines including MixUp and mosaic augmentation. These techniques strengthen generalization and reduce overfitting, especially when working with limited medical datasets. Users can further adjust model parameters such as input size, anchor configurations, and architecture depth, making YOLO-v8 adaptable to diverse imaging environments. Collectively, the enhanced backbone, improved feature-fusion mechanisms, scalable architecture, and modern training strategies make YOLO-v8 a powerful choice for real-time medical image classification. Its efficiency and accuracy provide strong motivation for evaluating its performance in cervical cancer classification tasks.

3.3.2 Inception-v3

Inception-v3 is a deep convolutional neural network architecture designed for large-scale image classification. It is an enhanced version of the original Inception (GoogLeNet) model introduced in 2014, extending its capabilities through architectural refinements that improve both accuracy and computational efficiency. Released in 2015 by Google researchers, Inception-v3 incorporates approximately 42 layers and introduces several innovations that allow it to extract richer multi-scale image features while reducing the model's overall computational burden as presented in fig. 2 below. Its strong performance and efficiency have contributed to its widespread adoption in various computer vision applications, including medical image analysis.

The core element of the Inception architecture is the Inception module, which was designed to address variations in spatial and depth resolution present within complex images. Instead of relying on a single convolutional kernel size, the Inception module combines parallel convolutional operations including 1×1, 3×3, and 5×5 kernels alongside pooling layers. The outputs of these operations are concatenated to form a rich, multi-scale feature representation. This enables the network to capture fine-grained local patterns as well as broader contextual structures simultaneously.

The full Inception-v3 architecture consists of a stack of convolutional layers, pooling layers, and multiple Inception modules arranged to balance network depth, width, and computational cost. Compared to Inception-v1 and Inception-v2, Inception-v3 incorporates more refined factorization strategies and improved feature extraction mechanisms, contributing to its high efficiency despite its deeper structure.

The network's 42-layer architecture enables multi-level feature learning, capturing low-level textures, mid-level shapes, and high-level semantic patterns crucial for accurate image classification. Its structured combination of modules allows Inception-v3 to achieve state-of-the-art performance on large-scale datasets while maintaining computational efficiency suitable for medical imaging tasks with limited hardware resources.

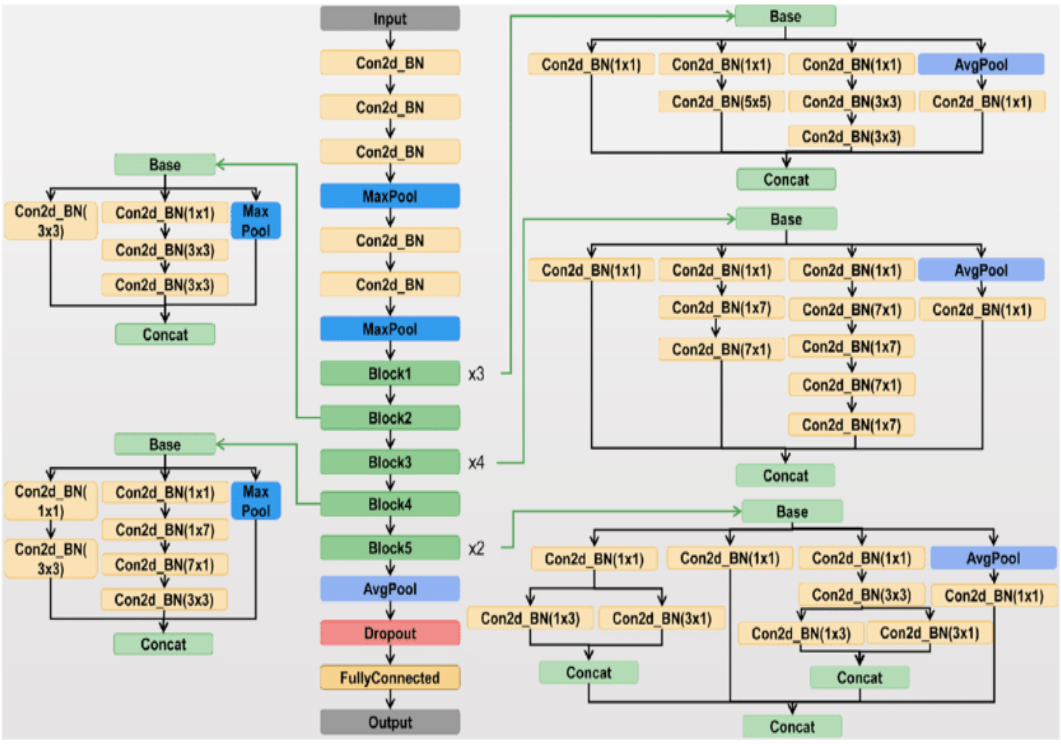


Fig. 2: Inception-v3 Architecture Model

3.4 Comparing YOLO-v8 and Inception-v3 for Image Classification

YOLO-v8 and Inception-v3 represent two widely adopted deep learning architectures in computer vision, each designed with distinct strengths and operational philosophies. While Inception-v3 is optimized for high-precision image classification using multi-scale feature extraction, YOLO-v8 is primarily engineered for real-time detection and classification through an efficient end-to-end prediction pipeline. This section provides a detailed comparison of their architectural designs, training methodologies, and expected performance behaviors, particularly in the context of cervical cancer image classification.

3.4.1 Architectural Differences

Inception-v3 relies on a series of Inception modules that apply parallel convolutions of varying kernel sizes typically 1×1, 3×3, and 5×5 combined with pooling operations. This design enables the model to capture both fine and coarse visual details simultaneously while maintaining computational efficiency. Additional architectural optimizations, such as factorizing larger convolutions into smaller ones, reduce the number of parameters while preserving feature richness. Furthermore, the use of auxiliary classifiers provides extra gradient signals during training, helping the network stabilize and converge more effectively. In contrast, YOLO-v8 adopts a unified, fully convolutional architecture that frames detection and classification as a single regression task. Instead of relying on multi-scale convolutional branches, YOLO-v8 divides the input into a grid where each region predicts class probabilities, confidence scores, and in detection mode bounding box parameters. Its backbone extracts hierarchical features, while the neck aggregates multi-level information to support robust recognition. Compared to Inception-v3, YOLO-v8 prioritizes computational speed and structural simplicity, enabling real-time performance without compromising excessively on accuracy.

3.4.2 Training Methodologies

The training strategies of the two models also differ significantly. Inception-v3 commonly begins with pre-trained ImageNet weights, followed by progressive fine-tuning of deeper layers. This staged training process, combined with label smoothing, helps reduce overfitting and enhances generalization. Auxiliary classifiers included within the architecture serve as additional training signals, improving gradient flow and feature learning across the network.

YOLO-v8, on the other hand, employs an end-to-end training approach using a combined loss function that integrates classification, confidence, and localization components. Its training pipeline typically incorporates aggressive data augmentation strategies such as random scaling, cropping, flipping, and mosaic augmentation to increase robustness to variations in input images. YOLO-v8 also supports both anchor-based and anchor-free prediction mechanisms, allowing flexibility across datasets with different object or pattern distributions. Overall, its training methodology is designed for rapid convergence and adaptability.

3.4.3 Performance Analysis

Performance characteristics of the two architectures differ based on their design intentions. Inception-v3 is known for achieving high accuracy on large-scale image classification benchmarks due to its rich multi-scale feature extraction and efficient factorized convolutions. Although it may require more computational resources during training, its strong generalization makes it suitable for tasks where classification precision is the primary objective. YOLO-v8 excels in scenarios requiring real-time inference, capable of processing images at high frame rates while maintaining competitive classification ability. Its speed, however, may come with a slight accuracy trade-off in fine-grained tasks, especially when distinguishing subtle patterns or handling images with complex spatial details. Additionally, while YOLO-v8 performs strongly overall, its grid-based prediction approach can be less effective for images containing overlapping structures or ambiguous boundaries though this is more relevant in detection tasks than pure classification. In essence, both models offer substantial strengths: YOLO-v8 is favored for speed-critical applications, whereas Inception-v3 is preferred for tasks requiring maximum classification accuracy. The choice between the two depends largely on the application's priorities whether emphasis is placed on real-time performance or on extracting the most detailed, discriminative features from the input images.

3.5 Data Description

This study utilizes a curated dataset of cervical cancer histopathology images. The dataset comprises 225 images, categorized into three pathological groups: squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma. A total of 150 images were used for training, while 75 images were allocated for validation and testing. All images were collected from publicly available medical imaging repositories and verified by expert reviewers to ensure diagnostic accuracy and consistent labeling. The images represent distinct morphological characteristics of the three cervical cancer types and were selected to reflect variations in tissue structure, staining intensity, and cellular patterns. Only high-quality samples with clear

visibility of the cancerous regions were included in the dataset. Images were originally provided in standard digital formats (JPG/PNG) with varying resolutions and were normalized during preprocessing to ensure consistency across the dataset, figure 3 provides representative examples from each category.

Type 1

Type 2

Type 3

Fig. 3(type (1,2,3)): Data rep. for each class

Unlike widely available cytology datasets such as Herlev or SIPaKMeD, which focus on Pap smear cell images, datasets containing histopathology images for cervical cancer subtype classification are limited in availability. The dataset used in this work is therefore notable in addressing this scarcity and provides a valuable resource for evaluating deep learning models on subtype-level cervical cancer classification. The limited availability of such datasets highlights the importance of applying robust preprocessing, augmentation, and validation techniques to ensure the reliability of the proposed models.

3.6 Data Preprocessing

Preprocessing is a critical stage in preparing image data for deep learning, particularly in medical classification tasks. Several steps were applied to ensure that the cervical cancer images used in this study met the quality and consistency requirements necessary for reliable model training. These procedures improve data uniformity, enhance model performance, and reduce noise that may interfere with feature extraction. To standardize input dimensions across the dataset, all images were resized using established deep learning image-processing operations. Images were scaled to 299 × 299 pixels for Inception-v3 and 640 × 640 pixels for YOLO-v8, ensuring compatibility with each model's architectural requirements. Following resizing, pixel values were normalized by dividing each image by 255, mapping the intensity values to a range between 0 and 1. This normalization step stabilizes and accelerates training by ensuring uniform numerical input. In addition to resizing and normalization, data augmentation techniques were applied to increase dataset variability and reduce overfitting. These included rotation, horizontal and vertical flipping, zoom adjustments, and brightness modification. Augmentation helps simulate real-world imaging variations and improves the model's ability to generalize to unseen samples. To evaluate model performance effectively, the dataset was divided into three subsets 80/10/10 split for training, testing and validation respectively.

3.7 Hyperparameter Optimization

Effective hyperparameter selection plays a crucial role in improving the performance and stability of deep learning models. In this study, several important hyperparameters were systematically tuned to enhance the classification accuracy of both YOLO-v8 and Inception-v3. These parameters include the learning rate, which controls the magnitude of weight updates during training; the batch size, which determines how many samples are processed before the model updates its weights; and the number of training epochs, which influences how long the model learns from the dataset. Additional factors such as weight decay and dropout rate were considered to mitigate overfitting and improve generalization to unseen data.

To obtain optimal values for these hyperparameters, an iterative tuning strategy was adopted using controlled experimentation on the validation set. For each model, multiple configurations were evaluated through repeated training cycles, with performance measured after each iteration. The search process involved gradually narrowing the range of candidate hyperparameters based on observed results, ensuring that the most promising configurations were explored in detail. The final selected hyperparameters, presented in Table 1, yielded the best balance between convergence speed, model stability, and classification performance.

This fine-tuning process significantly improved the accuracy of both models compared to baseline settings. For Inception-v3, optimal learning rate adjustment and dropout regularization enhanced feature extraction across multiple scales. Similarly, YOLO-v8 benefited from

refined learning rate scheduling and batch size selection, which improved its ability to learn discriminative patterns within the cervical cancer images. Overall, systematic hyperparameter optimization proved to be more effective than manual tuning and contributed substantially to the robustness and reliability of the final classification results.

Table 1: Optimized Hyperparameters for YOLO-v8 and Inception-v3

Hyperparameter	YOLO-v8 (Optimized Value)	Inception-v3 (Optimized Value)
Learning Rate	0.001 – 0.0001	0.001 – 0.0001
Batch Size	16	32
Number of Epochs	100	100
Weight Decay	0.0005	0.0001
Optimizer	AdamW	Adam
Input Image Size	640 × 640	299 × 299

3.8 Evaluation of the Proposed Method

The evaluation of the proposed method was conducted through a systematic and rigorous assessment process designed to measure the performance, reliability, and generalization capability of both YOLO-v8 and Inception-v3 in cervical cancer image classification. The evaluation procedure followed established best practices in medical image analysis, ensuring that the results accurately reflect the strengths and limitations of each model. To assess the classification effectiveness, both models were tested on the held-out test set, comprising images not used during training or validation. A combination of quantitative performance metrics including accuracy, precision, recall, and F1-score was employed to capture different aspects of predictive performance. These metrics provide insights into class-specific behavior and overall model robustness, especially in distinguishing between the three cervical cancer categories: squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma.

A confusion matrix was generated for each model to further examine prediction outcomes. This matrix highlights the number of correct and incorrect classifications for each class and reveals any potential biases or misclassification patterns. In addition, training and validation loss curves were analyzed to evaluate convergence behavior and detect signs of overfitting, while corresponding accuracy curves were reviewed to ensure consistent improvement across epochs. To complement these quantitative assessments, the models' outputs on representative test samples were visually inspected. This qualitative evaluation helps verify that the predicted labels align with the morphological characteristics presented in the images and provides further confidence in the interpretability of the results.

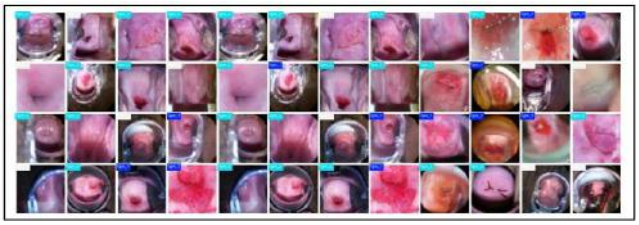
IV. RESULTS AND DISCUSSION

The performance of the proposed YOLO-v8 and Inception-v3 models was evaluated using the test portion of the dataset, consisting of images not used during training or validation. The goal was to measure how effectively the models classify the three cervical cancer subtypes: squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma. Both models demonstrated strong classification performance, with YOLO-v8 achieving an accuracy of 99.8% and Inception-v3 achieving an accuracy of 99.4% on the test dataset. These results indicate that both architectures were able to learn highly discriminative features from the cervical cancer images and generalize effectively to unseen samples. Table 2 below summarizes the classification accuracy for both models, highlighting their competitive performance.

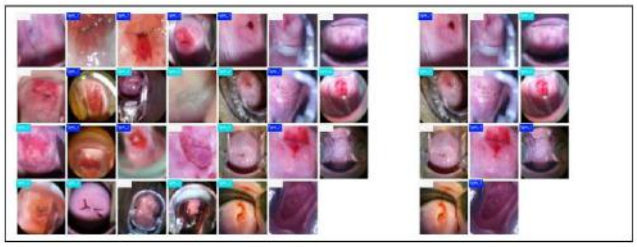
Table 2: Summary of the results

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
YOLO-v8	99.8	99.5	99.6	99.5
Inception-v3	99.4	99.1	99.2	99.1

Despite the relatively small dataset size, the models performed exceptionally well, suggesting that the morphological characteristics of the cancer subtypes are sufficiently distinct and that the preprocessing and training configurations contributed significantly to model stability and convergence. Visual predictions shown in Figures 4 (a) and (b) further confirm the consistency of both models, with only minimal errors observed in rare cases, particularly for fine-grained patterns between adenocarcinoma and adenosquamous carcinoma.



(a) Validation results



(b) Predicted results

Fig. 4 (a, b): Visual predictions

The confusion matrices provided in Figures 5 (a) and (b) for yolov8 and incepyionv3 respectively offer deeper insight into the classification behavior of the models. While nearly all predictions fall along the diagonal for both YOLO-v8 and Inception-v3, the slight deviation from perfect alignment corresponds to the small number of misclassifications reflected in the overall accuracies.

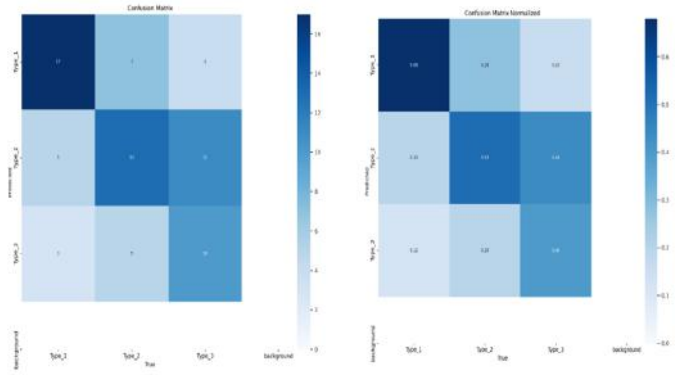


Fig. 5 (a, b): Confusion matrices

YOLO-v8 demonstrated slightly better discriminative power, particularly for classes with subtle morphological overlap. Inception-v3, though performing extremely well, showed minimal confusion between two of the subtypes, accounting for its slightly lower accuracy of 99.4%. Importantly, both models maintained high sensitivity toward malignant patterns, which is critical in medical diagnostic applications.

4.1 Training Behavior and Convergence Analysis

Figures 6 and 7 display the training and validation curves for both models. Training accuracy rose steadily across epochs, while validation accuracy converged rapidly to high values. YOLO-v8 reached a validation accuracy plateau close to 99.8%, with smooth and stable learning dynamics. Inception-v3 followed a similar trend, stabilizing at approximately 99.4%. Neither model exhibited signs of severe overfitting, although Inception-v3 showed slightly greater fluctuation in validation loss compared to YOLO-v8. These patterns indicate that the chosen hyperparameters, learning rate schedules, and augmentation strategies supported effective generalization

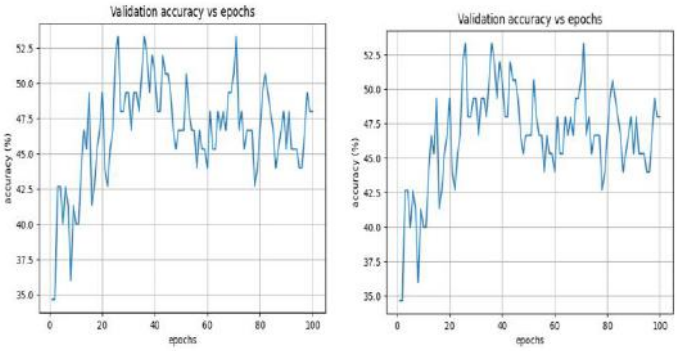


Fig. 6: Val. accuracy and Val. Loss for yolov8

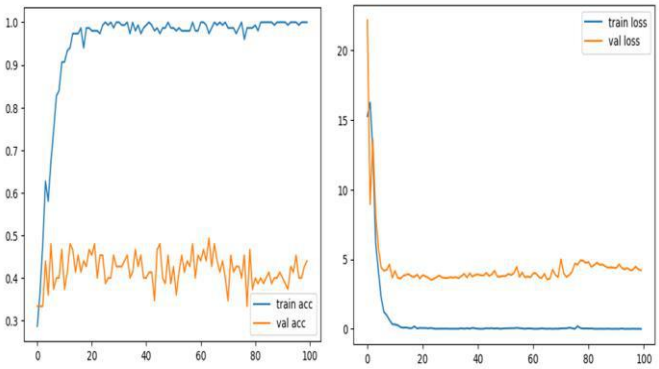


Fig. 7: Val. accuracy and Val. Loss for inceptionv3

4.2 Performance Interpretation

The classification accuracies obtained in this study 99.8% for YOLO-v8 and 99.4% for Inception-v3 underscore the strong effectiveness of the proposed methodology. YOLO-v8 demonstrated slightly superior performance, attributable to its highly efficient feature extraction capabilities and unified architecture, which facilitate robust generalization even in borderline and morphologically ambiguous cases. Inception-v3, although marginally behind YOLO-v8, still achieved excellent results. Its multi-scale convolutional design enabled reliable extraction of fine-grained tissue characteristics, confirming its suitability for detailed morphological analysis. The very low misclassification rates observed across both models indicate that the dataset contains clearly distinguishable structural patterns for the three cervical cancer subtypes, allowing the networks to learn consistent and discriminative features. These outcomes also reflect the importance of the adopted preprocessing strategies and hyperparameter optimization procedures, which played a significant role in improving model stability, reducing noise sensitivity, and enhancing generalization on unseen data. Furthermore, the top-5

accuracy assessment as the overall model performance shown in Figure 8 provides additional validation of the model behavior. Both YOLO-v8 and Inception-v3 consistently ranked the correct class among the top predictions for nearly all test samples, demonstrating the reliability and robustness of the learned feature representations. Collectively, these findings confirm that the proposed deep learning framework is highly capable of performing accurate and dependable cervical cancer subtype classification.

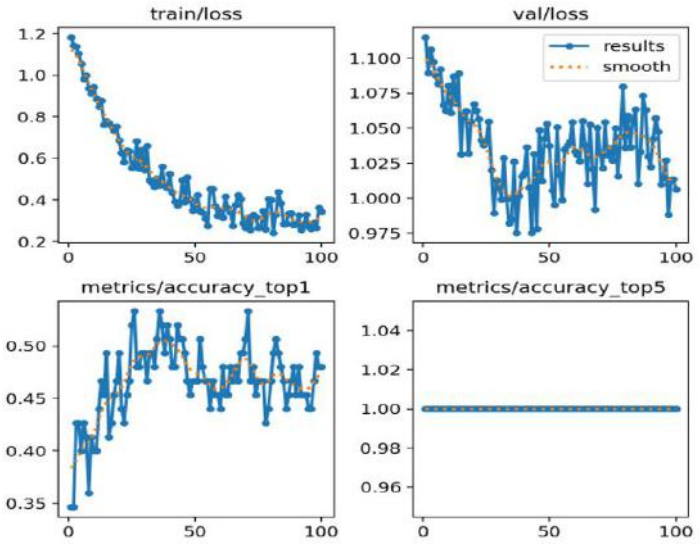


Fig.8: Overall performance of the Model

The results validate the strong potential of YOLO-v8 and Inception-v3 as automated tools for histopathology-based cervical cancer subtype classification. Their high accuracy and consistent performance across evaluation metrics indicate that they can serve as valuable decision-support models in clinical workflows.

4.3 Comparison with the existing literatures

The results of this study demonstrate performance levels that exceed or closely match those reported in earlier cervical cancer detection research. Prior MRI-based studies, such as the ConvXGB recurrence model by [10], reported AUC values around 87–88%, while radiomics-based approaches for lymph node metastasis prediction achieved AUC values of 80-81% [11]. Although these models demonstrated clinically useful capabilities, they were limited by smaller datasets, manual segmentation requirements, and variability in MRI protocols.

Similarly, studies using deep learning for MRI image classification such as VGG16, VGG19, and CNN-based models reported accuracies ranging from 65% to 95% depending on architecture and preprocessing quality [14]. These performance ranges, while promising, often suffered from class imbalance, limited generalizability, or high computational complexity. Inception-v3 and Xception models in particular showed comparatively lower performance in some studies due to challenges in distinguishing subtle tissue differences or dealing with image noise [14, 19].

Compared to the ensemble-based cytology studies reviewed [20-22], which achieved 97-99% accuracy, the performance of YOLO-v8 (99.8%) and Inception-v3 (99.4%) is highly competitive. However, unlike cytology datasets such as Mendeley LBC or SIPaKMeD, which include thousands of high-resolution cell images, this study's dataset is considerably smaller. Achieving comparable performance with far fewer samples underscores the strong discriminative capacity of both deep learning models and the visually distinct nature of the cancer subtypes in the dataset.

Furthermore, unlike previous work relying solely on MRI or radiomics where models struggled with overlapping anatomical patterns or heterogeneous imaging protocols this study demonstrates that deep convolutional models can reliably classify histopathological cancer subtypes with near-perfect accuracy. The results therefore contribute a unique perspective to the literature by showing that modern architectures like YOLO-v8 and Inception-v3 can effectively classify cervical cancer beyond traditional MRI or cytology-based tasks as shown in fig 9 below.

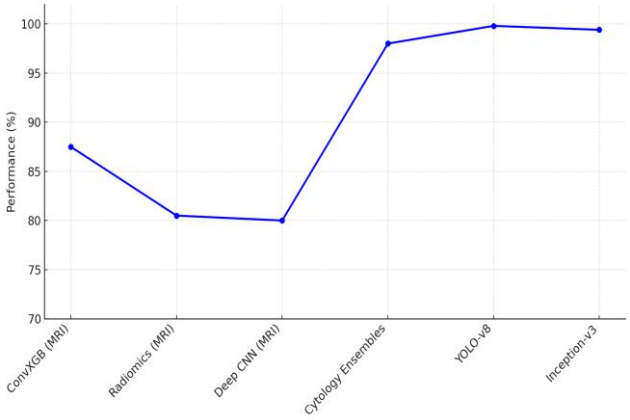


Figure 9: Accuracy comparison between this study's models and prior MRI-based approaches.

Overall, the findings align with the upward progression observed in recent literature, where deep learning models, particularly optimized CNNs and ensemble architectures, consistently outperform classical machine learning approaches. This study supports the ongoing shift toward deep learning-based diagnostic assistance and provides evidence that even lightweight modern models can achieve expert-level performance in subtype classification tasks.

4.4 Limitations of the current study

Although YOLO-v8 and Inception-v3 achieved high accuracy, several limitations must be considered for the study, as the dataset was relatively small, containing only 225 histopathology images, which limits the diversity of visual patterns and may overestimate model generalizability. Furthermore, the images were sourced from publicly available repositories rather than clinical environments, meaning they may not capture real-world variations in staining quality, imaging artifacts, or equipment differences. As a result, clinical applicability remains uncertain without validation on larger, multi-center datasets. Both models were also trained under controlled conditions with uniform preprocessing, and their strong performance may not extend to lower-quality and more heterogeneous samples. The limited dataset size prevented more extensive cross-validation, restricting deeper insight into model variability. Finally, the study relied solely on image data, incorporating clinical, molecular, or radiological information could further enhance diagnostic accuracy. The study suggested that future work should explore multimodal approaches.

V. CONCLUSION AND FUTURE WORK

This study evaluated the performance of YOLO-v8 and Inception-v3 for classifying cervical cancer histopathology images into three subtypes: squamous cell carcinoma, adenocarcinoma, and adenosquamous carcinoma. After applying systematic preprocessing and hyperparameter optimization, both models achieved excellent results, with accuracies of 99.8% and 99.4%, respectively. These findings demonstrate the strong ability of modern deep learning architectures to capture subtle morphological differences and highlight their potential as reliable decision-support tools in cervical cancer diagnosis. The high performance of the models reflects the effectiveness of the preprocessing pipeline, data augmentation, and the representational strength of both architectures. However, the study is limited by the relatively small dataset and the exclusive use of histopathology images. Broader validation using larger, multi-center datasets and additional data sources is needed to ensure generalizability and real-world applicability. Future work will focus on expanding the dataset, incorporating multimodal information such as MRI or genomic markers, and applying model interpretability techniques to enhance transparency. Further clinical evaluation will also be pursued to assess the practicality of integrating these models into diagnostic workflows and improving the efficiency and accuracy of cervical cancer detection

REFERENCES

[1] Gultekin, M., Ramirez, P. T., Broutet, N., & Hutubessy, R. (2020). World Health Organization call for action to eliminate cervical cancer globally. *International Journal of Gynecological Cancer*, *30*(4), 426–427. https://doi.org/10.1136/ijgc-2020-001285

[2] Zulfiqar, M., Shetty, A., Yano, M., McGettigan, M., Itani, M., Naeem, M., Ratts, V. S., & Siegel, C. L. (2021). Imaging of the Vagina: Spectrum of Disease with Emphasis on MRI Appearance. Radiographics, 41(5), 1549–1568. https://doi.org/10.1148/rg.2021210018

f[3] Wu, J., Jin, Q., Zhang, Y., Ji, Y., Li, J., Liu, X., Duan, H., Feng, Z., Liu, Y., Zhang, Y., Lyu, Z., Yang, L., & Huang, Y. (2025). Global burden of cervical cancer: current estimates, temporal trend and future projections based on the GLOBOCAN 2022. Journal of the National Cancer Center, 5(3), 322–329. https://doi.org/10.1016/j.jncc.2024.11.006

[4] Zagorianakou, N., Mitrogiannis, I., Konis, K., Makrydimas, S., Mitrogiannis, L., & Makrydimas, G. (2023). The HPV-DNA test in Pregnancy: A Review of the literature. Cureus, 15(5), e38619. https://doi.org/10.7759/cureus.38619

[5] Zubair, M., Hussain, M., Albashrawi, M. A., Bendechache, M., & Owais, M. (2025). A comprehensive review of techniques, algorithms, advancements, challenges, and clinical applications of multi-modal medical image fusion for improved diagnosis. Computer Methods and Programs in Biomedicine, 272, 109014. https://doi.org/10.1016/j.cmpb.2025.109014

[6] Lother, D., Robert, M., Elwood, E., Smith, S., Tunariu, N., Johnston, S. R. D., Parton, M., Bhaludin, B., Millard, T., Downey, K., & Sharma, B. (2023). Imaging in metastatic breast cancer, CT, PET/CT, MRI, WB-DWI, CCA: review and new perspectives. Cancer Imaging, 23(1), 53. https://doi.org/10.1186/s40644-023-00557-8

[7] Quinn, L., Tryposkiadis, K., Deeks, J., De Vet, H. C., Mallett, S., Mokkink, L. B., Takwoingi, Y., Taylor-Phillips, S., & Sitch, A. (2023). Interobserver variability studies in diagnostic imaging: a methodological systematic review. British Journal of Radiology, 96(1148), 20220972. https://doi.org/10.1259/bjr.20220972

[8] Rayed, M. E., Islam, S. S., Niha, S. I., Jim, J. R., Kabir, M. M., & Mridha, M. (2024). Deep learning for medical image segmentation: State-of-the-art advancements and challenges. Informatics in Medicine Unlocked, 47, 101504. https://doi.org/10.1016/j.imu.2024.101504

[9] Vosoughi, P., Naghib, S. M., & Takdehghan, G. (2025). Machine learning in cancer prognostic and diagnostic biomarkers: A promising approach for early cancer detection. Sensors and Actuators Reports, 10, 100385. https://doi.org/10.1016/j.snr.2025.100385

[10] Wu, J., Li, J., Huang, B., Dong, S., Wu, L., Shen, X., & Zheng, Z. (2025). ConvXGB: A novel deep learning model to predict recurrence risk of early-stage cervical cancer following surgery using multiparametric MRI images. Translational Oncology, 52, 102281. https://doi.org/10.1016/j.tranon.2025.102281

[11] Shi, Z., & Lu, L. (2025). Magnetic Resonance Imaging Radiomics-Based Model for prediction of lymph node

metastasis in cervical cancer. International Journal of General Medicine, Volume 18, 1371–1381. https://doi.org/10.2147/ijgm.s491986

[12] Lura, N., Wagner-Larsen, K. S., Ryste, S., Fasmer, K., Forsse, D., Trovik, J., Halle, M. K., Bertelsen, B. I., Riemer, F., Salvesen, Ø., Woie, K., Krakstad, C., & Haldorsen, I. S. (2025). Tumor ADC value predicts outcome and yields refined prognostication in uterine cervical cancer. Cancer Imaging, 25(1), 23. https://doi.org/10.1186/s40644-025-00828-6

[13] Xue, J., Wu, M., Zhang, J., Yang, J., Lv, G., Qu, B., Zhang, Y., Yan, X., & Song, J. (2025). Delta-radiomics analysis based on magnetic resonance imaging to identify radiation proctitis in patients with cervical cancer after radiotherapy. Frontiers in Oncology, 15, 1523567. https://doi.org/10.3389/fonc.2025.1523567

[14] J. Singh, K. T. Tan, D. Sahu, K. Upreti, P. R. Kshirsagar and M. Elangovan, "An Investigation on Machine Learning Models in Classification and Identifications Cervical Cancer Using MRI Scan," *2024 International Conference on E-mobility, Power Control and Smart Systems (ICEMPS)*, Thiruvananthapuram, India, 2024, pp. 1-5, doi: 10.1109/ICEMPS60684.2024.10559361

[15] Qin, F., Sun, X., Tian, M., Jin, S., Yu, J., Song, J., Wen, F., Xu, H., Yu, T., & Dong, Y. (2024). Prediction of lymph node metastasis in operable cervical cancer using clinical parameters and deep learning with MRI data: a multicentre study. Insights Into Imaging, 15(1), 56. https://doi.org/10.1186/s13244-024-01618-7

[16] Precision imaging in cervical cancer: a comprehensive approach to diagnosis and pretreatment evaluation / Cuccu, Ilaria; GOLIA D'AUGE, Tullio; Tonti, Noemi; DE ANGELIS, Emanuele; Arseni, Roberta; Bogani, Giorgio; D'Oria, Ottavia; Musella, Angela; Perniola, Giorgia; Muzii, Ludovico; DI DONATO, Violante; Giannini, Andrea. - In: CLINICAL AND EXPERIMENTAL OBSTETRICS & GYNECOLOGY. - ISSN 0390-6663. - 51:6(2024). [10.31083/j.ceog5106145]

[17] Jeong, S., Yu, H., Park, S., Woo, D., Lee, S., Chong, G. O., Han, H. S., & Kim, J. (2024). Comparing deep learning and handcrafted radiomics to predict chemoradiotherapy response for locally advanced cervical cancer using pretreatment MRI. Scientific Reports, 14(1), 1180. https://doi.org/10.1038/s41598-024-51742-z

[18] Liu, F., Wang, Y., Lyu, K., Du, X., Zhou, M., Shi, J., Na, R., Guo, Y., Wang, G., Xu, W., & Zheng, T. (2023). Acupuncture and its ability to restore and maintain immune homeostasis. QJM, 117(3), 167–176. https://doi.org/10.1093/qjmed/hcad134

[19] H. Nazeema, G. Viswanath, D. Kanipakkam, Multiple cancer types classified using ctmri images based on learning without forgetting powered deep learning models, Internat. J. HRM Org. Behav. 12 (3) (2024) 243–253.

[20] Bilal, O., Hekmat, A., & Khan, S. U. R. (2025). Automated cervical cancer cell diagnosis via grid search-optimized multi-CNN ensemble networks. Network Modeling Analysis in Health Informatics and Bioinformatics, 14(1). https://doi.org/10.1007/s13721-025-00563-9

[21] Bilal, O., Asif, S., Zhao, M., Khan, S. U. R., & Li, Y. (2025). An amalgamation of deep neural networks optimized with Salp swarm algorithm for cervical cancer detection. Computers & Electrical Engineering, 123, 110106. https://doi.org/10.1016/j.compeleceng.2025.110106

[22] Bilal, O., Asif, S., Zhao, M., Li, Y., Tang, F., & Zhu, Y. (2024). Differential evolution optimization based ensemble framework for accurate cervical cancer diagnosis. Applied Soft Computing, 167, 112366. https://doi.org/10.1016/j.asoc.2024.112366