Highlights

TransPapCanCervix: An Enhanced Transfer Learning-based Ensemble Model for Cervical Cancer Classification

Barkha Bhavsar,Bela Shrimali

- Create a deep Neural Network model architecture for precise identification and categorization of squamous cell carcinoma cells through single-cell classification.

- Assess the model's effectiveness using diverse evaluation metrics, encompassing accuracy, precision, recall, and F1 score.

- Investigate the impact of various hyperparameters, such as the number of layers, neurons, activation functions, and optimization algorithms on the performance of the developed model.

- Explore the potential for future advancements in the proposed approach, such as the integration of additional deep learning architectures and ensemble techniques, to further enhance Squamous Cervical Cancer detection accuracy and streamline clinical decision-making.

TransPapCanCervix: An Enhanced Transfer Learning-based Ensemble Model for Cervical Cancer Classification

Barkha Bhavsar[a,*,1], Bela Shrimali[b,2]

[a]*LDRP-ITR, Gandhinagar, KSV, SVKM, KH-5 Circle, Near ITI, Gandhinagar, 382015, India*
[b]*Department of Computer Science and Engineering, Institute Of Technology, Nirma University, Ahmedabad, 382481, Gujarat, India*

ARTICLE INFO

ABSTRACT

Cervical cancer, like many other cancers, is most treatable when detected at an early stage. Using classification methods helps find early signs of cancer and small tumors. This allows doctors to act quickly and offer treatments that might cure the cancer. This study presents a comprehensive approach to the classification of squamous cell carcinoma (SCC) leveraging a dataset comprising 1140 single-cell images sourced from Herlev. A variety of deep learning models, including DenseNet121, DenseNet169, InceptionResNet, XceptionNet, ResNet50, and ResNet101, are employed both individually and in ensembles, demonstrating their efficacy in classifying diverse cellular features. To validate the robustness of results, k-fold cross-validation is conducted, further affirming the effectiveness of the proposed methodology. Thorough exploration produces a precise and effective model for SCC classification, providing detailed insights into both normal and abnormal cell types. These findings show that transfer learning-based deep neural networks and ensemble methods can improve the diagnostic capabilities by 98% accuracy, of SCC classification systems for different cell types.

1. Introduction

Cervical cancer is a significant global health concern, particularly affecting women in both developed and developing countries Huang and Feng (2022); Tan, Selvachandran, Ding, Paramesran and Kotecha (2023); Youneszade, Marjani and Pei (2023); Alyafeai and Ghouti (2020); Stelzle, Tanaka, Lee, Khalil, Baussano, Shah, McAllister, Gottlieb, Klug, Winkler et al. (2021). It arises from the uncontrolled growth of cells in the cervix, often triggered by persistent infection with certain types of Human Papillomavirus (HPV) Lu, Song, Ghoneim and Alrashoud (2020); Vargas-Cardona, Rodriguez-Lopez, Arrivillaga, Vergara-Sanchez, García-Cifuentes, Bermúdez and Jaramillo-Botero (2023); Habtemariam, Zewde and Simegn (2022); Park, Kim, Ju, Nam, Kim and Kim (2021). Despite advancements in screening programs and HPV vaccination, cervical cancer continues to be a major cause of death and illness from cancer in women around the world Alyafeai and Ghouti (2020); Fekri-Ershad and Ramakrishnan (2022); Hull, Mbele, Makhafola, Hicks, Wang, Reis, Mehrotra, Mkhize-Kwitshana, Kibiki, Bates et al. (2020); Huang and Feng (2022); Tan et al. (2023); Xue, Wang, Qin, Yan, Qu, Seery, Jiang and Qiao (2022). According to the recent reports and statistics by the Catalan Institute of Oncology (ICO) and the International Agency for Research on Cancer (IARC) for information center on HPV and Cancer, India has 511.4 million women aged 15 and older at risk of cervical cancer, with approximately 123,907 new cases and 77,348 deaths annually, ranking it as the second most frequent cancer among women in the country Youneszade et al. (2023); Priya, Karthikeyan and Palanikkumar (2023). About 5.0% of women harbor cervical HPV-16/18 infection at any given time, with 83.2% of invasive cervical cancers attributed to these HPV types Bruni, Albero, Serrano, Mena, Collado, Gómez, Muñoz, Bosch and Sanjosé (2023); Hussain, Mahanta, Das and Talukdar (2020); Tripathi, Arora and Bhan (2021); Vargas-Cardona et al. (2023). The WHO's Global Strategy to Accelerate the Elimination of Cervical Cancer as a Public Health Problem aims to achieve vaccination coverage for 90% of girls by 15 years of age and ensure that 70% of women receive cervical cancer screening by 35 and 45 years of age Gabutti and Ogliastro (2024); Vargas-Cardona et al. (2023).

Limited access to preventive measures and treatment makes cervical cancer a critical health challenge in developing countries Fekri-Ershad and Ramakrishnan (2022); Hull et al. (2020), exacerbated by the scarcity of expert resources for timely diagnosis Park et al. (2021); Xue et al. (2022). Tragically, in 2020, cervical cancer claimed the lives of

*Corresponding author

✉ barkha.bhavsar@gmail.com (B. Bhavsar); bela.shrimali@gmail.com (B. Shrimali)
ORCID(s): 0000-0001-6592-5440 (B. Bhavsar); 0000-0002-7543-5389 (B. Shrimali)

over 340,000 people worldwide, with an estimated 600,000 new cases diagnosed Attallah (2023a); Ji, Zhang, Dong, Lin, Geng and Hong (2023); Sung, Ferlay, Siegel, Laversanne, Soerjomataram, Jemal and Bray (2021). The National Health Portal of India reports that cervical cancer ranks as the second most common cancer among Indian women, with an estimated 96,922 new cases and 60,078 deaths in 2020 Singh, Jha, Shri, Bhattacharyya, Patel and Dhamnetiya (2022); Tripathi et al. (2021). Recent studies highlight the alarming gap in access to diagnostic expertise, contributing to delayed detection and treatment initiation Arora, Dhawan and Singh (2021); Tripathi et al. (2021); Chandran, Sumithra, Karthick, George, Deivakani, Elakkiya, Subramaniam, Manoharan et al. (2021); Wang, Tian, Xu, Zhang, Li, Zhao and Bai (2022). With limited screening programs and a shortage of trained healthcare professionals, many cases go undetected until advanced stages, drastically reducing survival rates Arifianto and Agoes (2021); Alyafeai and Ghouti (2020); Rahaman, Li, Yao, Kulwa, Wu, Li and Wang (2021). Moreover, the absence of immediate reporting mechanisms further compounds the problem, prolonging the diagnostic process and increasing patient anxiety Huang and Feng (2022); Lu et al. (2020); Li, Xu, Shen, Zhou, Xiao and Li (2021); Lilhore, Poongodi, Kaur, Simaiya, Algarni, Elmannai, Vijayakumar, Tunze and Hamdi (2022). The combination of subjective interpretations and the possibility of errors in medical opinions further hinders effective cervical cancer management due to the absence of expert supervision Tan et al. (2023); Vargas-Cardona et al. (2023); Wang et al. (2022). It is crucial to allocate resources promptly and give urgent attention to this pressing issue to ensure all women in developing countries have fair access to precise diagnosis, ultimately leading to improved outcomes Hussain et al. (2020); Tan et al. (2023). Cervical Cancer cell stratification is essential for early detection, treatment planning, prognostic assessment, research advancement, and public health initiatives Pramanik, Biswas, Sen, de Souza Júnior, Papa and Sarkar (2022); Park et al. (2021); Plissiti, Dimitrakopoulos, Sfikas, Nikou, Krikoni and Charchanti (2018). By accurately characterizing tumors based on their histopathological, molecular, and clinical features, classification systems play a pivotal role in improving patient outcomes and reducing the global burden of cervical cancer Chandran et al. (2021); Tan et al. (2023). Classification systems provide valuable prognostic information by predicting the likelihood of disease progression, recurrence, and patient survival Arora, Dhawan and Singh (2023); Bhavsar and Shrimali (2024); Tripathi et al. (2021).

Machine learning (ML) and deep learning (DL) are trending topics in Artificial Intelligence (AI), as they allow machines to learn from vast amounts of data and make increasingly accurate predictions and decisions Bhavsar and Shrimali (2024); Devi, Gaikwad and Harikrishnan (2023); Lu et al. (2020); Park et al. (2021); Paul, Bhowmik and Bhattacharjee (2015); Tanimu, Hamada, Hassan, Kakudi and Abiodun (2022). While machine learning relies on handcrafted features and simpler models like decision trees or support vector machines, deep learning utilizes neural networks with multiple layers to automatically learn hierarchical representations of data, eliminating the need for manual feature engineering Ghoneim, Muhammad and Hossain (2020); Rahimi, Akbari, Asadi and Emami (2023); Vargas-Cardona et al. (2023); Youneszade et al. (2023). Deep learning models, with their complex architectures and millions of parameters, excel in tasks like image and speech recognition and classification, natural language processing, and autonomous driving, but they require large datasets and significant computational resources for training Attallah (2023a); Hussain et al. (2020); Kalbhor and Shinde (2023). In contrast, traditional machine learning techniques are more suitable for scenarios with limited data and computational resources, finding applications across various domains such as finance, healthcare, and marketing Devi et al. (2023); Karani, Gangurde, Dhumal, Gautam, Hiran and Marathe (2022); Lu et al. (2020); Park et al. (2021); Sundarambal, Karthikeyini, Bommi, Subramanian and Jacintha (2022); Ratul, Al-Monsur, Tabassum, Ar-Rafi, Nishat and Faisal (2022).

DL has emerged as a highly suitable approach for healthcare applications due to its capacity to analyze extensive and intricate datasets such as medical images, genomic information, and electronic health records Hemalatha and Vetriselvi (2022); Huang and Feng (2022); Wang et al. (2022); Nambu, Mariya, Shinkai, Umemoto, Asanuma, Sato, Hirohashi, Torigoe, Fujino and Saito (2022). This enables more precise diagnosis, personalized treatment plans, and predictive analytics. In our proposed framework, DL models, particularly Convolutional Neural Networks (CNNs), are integrated for the classification of cervical cells Arifianto and Agoes (2021); Bhatt, Ganatra and Kotecha (2021); Bhavsar and Shrimali (2023); Tanimu et al. (2022). CNNs excel in tasks such as tumor detection and disease classification through image analysis, extracting insights from unstructured clinical data, and predicting patient outcomes based on historical records Youneszade et al. (2023); Bhavsar and Shrimali (2023); Kruczkowski, Drabik-Kruczkowska, Marciniak, Tarczewska, Kosowska and Szczerska (2022). DL plays a pivotal role in cervical cancer cell classification by facilitating accurate analysis of cytology images, aiding in early detection and diagnosis Khobragade, Jain and Sisodia (2020); Tan et al. (2023); Wang et al. (2022). Transfer learning (TL) proves invaluable in addressing challenges posed by limited datasets, enhancing classification accuracy despite data scarcity, and reducing training time and computational resources Dhawan, Singh and Arora (2021); Ji et al. (2023); Hemalatha and Vetriselvi (2022); Tripathi

et al. (2021); Zhuang, Qi, Duan, Xi, Zhu, Zhu, Xiong and He (2020). Combining TL with an Ensemble approach holds great promise for cervical cancer cell classification Ghoneim et al. (2020); Zhuang et al. (2020). It allows leveraging pre-trained models to extract relevant features from images, while ensemble methods aggregate the predictions of multiple models to enhance overall accuracy and robustness Attallah (2023a,b). This approach not only improves classification accuracy, sensitivity, specificity, and F1 scores but also enhances the system's ability to generalize to new and unseen data Khobragade et al. (2020); Tan et al. (2023). Moreover, the ensemble strategy helps mitigate the risk of overfitting and reduces the likelihood of erroneous predictions, making the classification system more reliable for clinical use Khamparia, Gupta, de Albuquerque, Sangaiah and Jhaveri (2020). As the integration of TL and ensemble methods offers a powerful and effective approach to cervical cancer cell classification Chen, Li, Gao and Shen (2020); Hemalatha and Vetriselvi (2022), and the ability to improve patient outcomes through more accurate diagnoses and treatment decisions, the Ensemble Transfer Learning approach is incorporated in our proposed architecture.

We proposed a working model that focuses on enhancing the classification accuracy of Squamous Cell Carcinoma (SCC) for the cervix by employing advanced deep-learning models and ensemble techniques. Utilizing the Herlev dataset, which includes 917 images of both multi-cell and single-cell samples, our contributions are as follows:

(1) Our research pipeline involves a comprehensive pre-processing stage encompassing image acquisition, normalization, and enhancement techniques like Clahe, median filtering, and NLM filtering. Subsequently, we employ data augmentation strategies to enrich the processed dataset.

(1) To improve the predictive capability of our models, we employ transfer learning, leveraging pre-trained architectures including ResNet50V2, ResNet101, DenseNet121, DenseNet169, XceptionNet, and InceptionResNetv2. Subsequently, we apply Ensemble techniques to amalgamate the strengths of individual models, resulting in an ensemble model with an improved accuracy, specificity, sensitivity, and F1 score as compared to the cutting-edge state-of-the-art.

(1) Our final model effectively classifies cellular images into normal and abnormal categories, providing a reliable tool for early detection and intervention in cases of Squamous Cell Carcinoma cervical cells. The success of our approach underscores its potential for significant contributions to the field of medical image analysis and cancer diagnosis.

The structure of the paper follows a systematic approach, beginning with an introduction to the research problem of improving the classification accuracy of Squamous Cell Carcinoma (SCC) of cervical cells through deep learning and ensemble techniques. This is followed by a thorough review of related literature, providing context and background for the study. The methodology section outlines the various steps undertaken, starting with data acquisition from the Herlev dataset and pre-processing steps including image normalization, enhancement, and augmentation. The application of transfer learning and the integration of multiple pre-trained architectures are detailed, along with the ensemble technique employed to combine individual models. The following section dives into the evaluation of our proposed approach. We will analyze its performance using accuracy metrics and compare it to existing methods. Lastly, the discussion interprets the findings, highlighting the significance of the achieved accuracy in SCC classification and potential avenues for future research. The paper concludes with a summary of key findings and contributions to the field of medical image analysis and cancer diagnosis.

2. Related Work

Despite being preventable, cervical cancer remains a major health crisis for women in low- and middle-income nations. Increased access to healthcare is crucial Bhatt et al. (2021). The human papillomavirus (HPV), particularly persistent infection with high-risk strains, is the main cause of cervical cancer Allahqoli, Laganà, Mazidimoradi, Salehiniya, Günther, Chiantera, Karimi Goghari, Ghiasvand, Rahmani, Momenimovahed et al. (2022); Habtemariam et al. (2022); Khan, Han, Ilyas, Lee and Lee (2023). Numerous studies highlight the prevalence of HPV infection as a crucial risk element for cervical cancer. In India, cervical cancer ranks as a prominent cause of cancer-related illness and death among women Bhavsar and Shrimali (2024). Despite its preventable nature, cervical cancer incidence remains high in India, emphasizing the pressing need for comprehensive prevention and early detection strategies. Various factors, including limited access to healthcare services and low awareness levels, contribute to the burden of cervical cancer Tan et al. (2023). The absence of regular screening, notably Pap smears, poses a significant obstacle, as it hampers the timely detection of precancerous lesions Paul et al. (2015); Plissiti et al. (2018). The literature underscores
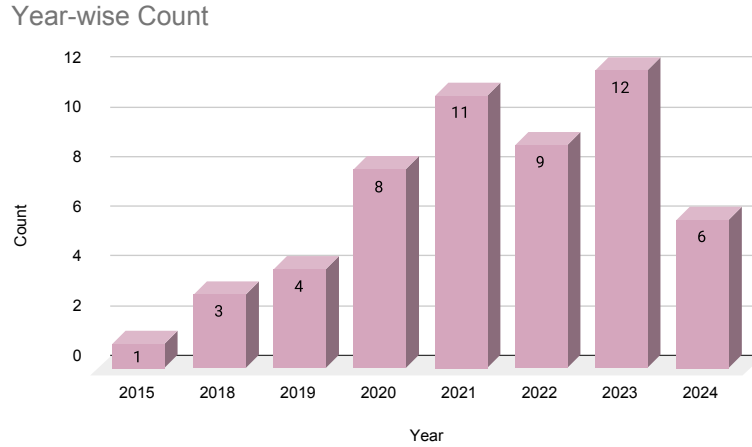
Figure 1: Year-wise Breakdown of current Literature

the necessity for a computer-assisted system to address cervical cancer comprehensively, considering the complex array of risk factors and the diverse healthcare environments worldwide. By consolidating existing knowledge and leveraging evidence-based approaches, advancements in cervical cancer prevention and early detection can substantially enhance women's health and overall well-being Allahqoli et al. (2022); Bhavsar and Shrimali (2023, 2024); Khan et al. (2023).

The state-of-the-art is gathered through various well-known publishers by applying keywords including cervical cancer classification, Pap-smear, Deep learning, transfer learning, ensemble deep neural network, cervical cancer cell detection, machine learning etcetera. We analyze the state-of-the-art research in cervical cancer detection and classification. Table 1 summarizes the number of publications per year, while Figure 1 visually depicts this trend. A wide range of esteemed publishers, such as Elsevier, IEEE, and Springer, provide a vast collection of research articles on cervical cancer detection and stratification. This article also includes publishers like MDPI, Nature Publishing Group, Hindawi, and many more. The graph shown in figure 2 displays the summary of the current state-of-the-art count for better understanding.

Table 1
Breakdown of State-of-an-art Year-wise

Year	2015	2018	2019	2020	2021	2022	2023	2024
Count	1	3	5	9	12	16	15	6

The proposed architecture presented by Bhaswati Singha Deo et al. Deo, Pal, Panigrahi and Pradhan (2024) explores CerviFormer, a novel approach for classifying cervical cancer in Pap smear images. CerviFormer utilizes Transformers, offering flexibility in handling various image sizes. It incorporates a cross-attention technique to efficiently condense data into a compact module, making it suitable for large-scale datasets. Evaluated on public datasets, CerviFormer achieved impressive results. On the Sipakmed dataset (3-class classification), it reached an accuracy of 93.70%. Similarly, on the Herlev dataset (2-class classification), it achieved 94.57% accuracy. These findings suggest CerviFormer's competitiveness compared to existing methods, positioning it as a comprehensive model for cervical cancer detection in Pap smears. Integrating such models into clinical practice could empower medical professionals to deliver more effective cervical cancer treatment through improved screening processes.

The paper presented by Nina Youneszade et al. Youneszade et al. (2023) provides an in-depth review of current research on deep learning (DL) solutions utilizing image segmentation and classification techniques for analyzing and categorizing cervical screening images. It emphasizes the significance of DL techniques in processing and classifying cervical cytology and colposcopy images, with convolutional neural networks (CNNs) showcasing exceptional performance. This review explores the potential of compound algorithms to enhance classification performance in cervical cancer. It specifically examines the role of image segmentation in analyzing both cervical cytology (Pap

Publisher-wise Count
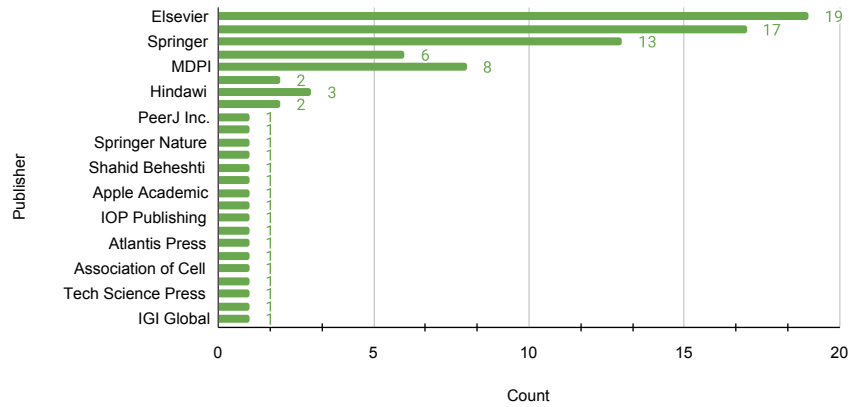
Figure 2: Publisher-wise Breakdown of current Literature

smear) and colposcopy images. Currently, Convolutional Neural Networks (CNNs) like ANFIS, CapsNet, ResNet, VGGNet, and AlexNet are widely used. The future holds promise for incorporating mixed feature selection techniques that leverage Deep Learning algorithms like RCNN, Faster RCNN, and VGG19. This combined approach could lead to even more accurate cervical cancer detection. ML and DL-powered computer-aided diagnostics (CAD) are gaining traction as critical tools for boosting cervical cancer detection accuracy, especially in resource-limited regions with high cervical cancer mortality rates. The review offers insights, research opportunities, and future directions for DL-based approaches in cervical cancer screening, aiming to advance diagnostic capabilities in this field.

The study presented by Sher Lyn Tan et al. Tan et al. (2023) tackles the limitations of traditional cervical cancer detection by proposing deep learning models that bypass cell segmentation and manual feature engineering. Employing transfer learning with pre-trained convolutional neural networks (CNNs), the research extensively evaluates 13 models on a seven-class classification task using the Herlev dataset. DenseNet-201 stands out as the most accurate and efficient model. By leveraging transfer learning to address data scarcity and class imbalance, the study demonstrates the efficacy of CNNs for multi-class classification without pre-processing. Notably, DenseNets, especially DenseNet-201, outperforms other pre-trained models, highlighting their potential for accurate and efficient cervical cancer detection without separate segmentation or feature extraction steps.

The research presented by Madhura M. Kalbhor et al. Kalbhor and Shinde (2023) suggests that automated recognition and classification of Pap smear cells are crucial for early disease diagnosis. Unlike traditional methods that rely on manually designed features, deep learning offers a powerful approach to tumor identification. This study explores two deep learning strategies: utilizing pre-trained models for feature extraction and leveraging transfer learning for classification. In the first approach, ResNet-50 achieved a 92.03% classification accuracy by extracting informative features from the data. Employing transfer learning with fine-tuning, GoogleNet surpassed this performance, reaching a remarkable 96.01% accuracy.

The study presented by Rishav Pramanik et al. Pramanik et al. (2022) investigates using three transfer learning models: InceptionV3, MobileNetV2, and InceptionResNetV2. To improve their performance for cervical cancer detection in Pap smear images, the models are enhanced with additional layers that can capture features specific to this task. A new ensemble method is introduced that combines the predictions from these models. This method focuses on minimizing the difference between the models' predictions and the actual labels (ground truth). It achieves results by incorporating various distance measures like Euclidean, Manhattan, and Cosine distance to effectively combine the strengths of each model. Experimental results demonstrate significant performance improvements, with the ensemble achieving an accuracy of 96.96%, surpassing the performance of individual models. Evaluation of public datasets demonstrates that the proposed end-to-end classification technique is competitive with existing cutting-edge methods. This achievement suggests its potential to significantly improve cervical cancer detection. By offering more accurate

results, the approach could empower medical professionals to make better treatment decisions and ultimately enhance the overall effectiveness of Pap smear testing.

Habtemariam et al. Habtemariam et al. (2022) addressed challenges in cervical cancer diagnosis by developing an integrated deep-learning system for automatic cervix-type classification. The system analyzed colposcopy and histopathology images. Pre-trained models were first evaluated for cervix-type classification. Then, a MobileNetv2-YOLOv3 model identified the relevant region (cervix) within the image, followed by classification using an Efficient-NetB0 model. For cancer detection, a separate EfficientNetB0 model, trained on specially prepared histopathology images, was employed. The system achieved impressive results: 99.88% mean average precision for region extraction and 96.84% and 94.5% test accuracy for cervix type and cancer classification, respectively. These findings highlight the system's potential as a valuable decision-making aid in diagnosing cervical cancer, especially in areas with limited resources and expertise.

The research conducted by Hiam Alquran et al. Alquran, Mustafa, Qasmieh, Yacob, Alsalatie, Al-Issa and Alqudah (2022) presents a computer-aided system for analyzing cervical cancer in women through automatic Pap smear image analysis. The proposed hybrid system utilizes a polynomial SVM classifier to classify test Pap smear cell images into seven classes, including Superficial, Intermediate, and Columnar Normal classes, as well as Mild, Moderate, and Severe Carcinoma in situ for abnormal images. Feature extraction is performed automatically using a predefined CNN, ResNet101, with further feature reduction accomplished via PCA to enhance accuracy. The system utilizes a cascade of five Support Vector Machine (SVM) models to achieve high confidence in its predictions, ensuring its suitability for real-world clinical applications in hospitals. The methodology is fast and accurate, diagnosing an image in less than one second, and is the first of its kind to detect and quantify abnormality levels. The system overcomes challenges faced by previous research, such as segmentation issues and image enhancement, by automatically extracting features without image preprocessing. This approach maintains computational efficiency and preserves image information, resulting in high-confidence results that meet clinical requirements. Despite its limitations, such as the need for additional datasets and high-performance CPUs, the system demonstrates the potential for improving accuracy and specificity in cervical cancer diagnosis.

The research carried out by Anurag Tripathi et al. Tripathi et al. (2021) contributes by applying deep transfer learning methods to the SIPAKMED dataset and implementing various architecture DL models including Resnet-50, Resnet-152, VGG-16, and VGG-19. The proposed architecture is achieving a high classification accuracy of 94.89% using the ResNet-152 architecture. Classification targets included Dyskeratotic, Koilocytotic, Metaplastic, Parabasal, and Superficial-intermediate cell types. The researchers also suggest that future enhancements could involve integrating multiple datasets to improve robustness and optimizing training runtime to expedite diagnosis and treatment decisions.

Park et al. Park et al. (2021) compared machine learning (ML) and deep learning (DL) models for detecting cervical cancer in cervicography images. They utilized ResNet-50 for DL and XGBoost, SVM, and Random Forest for ML. The study classified 4119 images (positive or negative for cancer) and validated feature extraction through cross-validation. The DL model, ResNet-50, significantly outperformed ML models (AUC: 0.97 vs. 0.79-0.84). This suggests ResNet-50's potential for more accurate diagnoses. The research highlights the promise of DL for robust computer-aided diagnosis (CAD) tools in cervical cancer. Moreover, it emphasizes the importance of considering diagnostic factors identified by various model architectures to improve clinical decision-making.

Thorough research conducted by Venkatesan Chandran et al. Chandran et al. (2021) introduces two DL CNN architectures, VGG19 (TL) and CYENET, for the automated detection of cervix carcinoma cells from colposcopy images. VGG19 (TL) utilizes transfer learning with satisfactory results, achieving a classification accuracy of 73.3% and a moderate Cohen's Kappa score. CYENET, a new dedicated model, outperforms VGG19 (TL) significantly, achieving a classification accuracy of 92.3%, with noticeably improved sensitivity, specificity, and kappa scores. The proposed CYENET model exhibits promise as a diagnostic assistance tool for clinicians, offering improved classification efficiency and accuracy in cervical cancer detection through colposcopy screening. Future work may involve validating the model with different datasets and enhancing its performance through advanced pre-processing techniques and CNN architectures.

The paper authored by Vinay Khobragade et al. Khobragade et al. (2020) presents a novel method for cervical cancer screening using transfer learning with a deep EfficientNet model. This method aims to mitigate manual errors and time limitations by eliminating the need for segmentation. The model undergoes training first on the ImageNet dataset and then on a dataset containing microscopic images of cervical cells. Unlike previous techniques that focus on individual cells, this approach directly processes images containing multiple cells, thereby enhancing the efficiency of cell analysis per unit time. Through ten-fold cross-validation on the Herlev Pap smear dataset, the study demonstrates the model's

efficacy in classification. Notably, EfficientNet outperforms other methods in terms of accuracy and performance metrics while also reducing processing time, indicating its potential as a valuable tool for cervical cancer screening.

Khamparia et al. Khamparia et al. (2020) introduces an innovative approach for detecting and categorizing cervical cancer in Pap smear images, utilizing a framework driven by the Internet of Health Things (IoHT) and deep learning. Transfer learning is employed, combining various pre-trained CNN models with traditional machine learning techniques. Evaluation of the Herlev dataset demonstrates that the ResNet50 CNN pre-trained model achieves a classification rate of 97.89% when combined with the random forest classifier, with minimal training and testing times of 0.032 and 0.006 seconds, respectively.

The study conducted by Elima Hussaina et al. Hussain et al. (2020) assesses the effectiveness of 6 deep CNN architectures—AlexNet, VGGNet (VGG-16 and VGG-19), ResNet (ResNet-50 and ResNet-101), and GoogLeNet—for diagnosis of cervical precancerous and cancer lesions across 4 classes. An ensemble classifier, formed by combining 3 top-performing models, is introduced to improve multi-class classification accuracy. Training and validation are conducted on a diverse pap smear dataset from conventional and liquid-based cytology methods, alongside the Herlev dataset for benchmarking. The evaluation process includes image resizing, dataset division, classifier training, test dataset prediction, and performance metric analysis such as accuracy, precision, recall, specificity, false-positive and false-negative rates, and area under the curve (AUC). The ensemble classifier excels in reducing false-positive and false-negative rates, offering a robust diagnostic approach with AUC values surpassing 90%. It emerges as a versatile and effective model for cervical pap smear image classification, capable of classifying whole slide pap smear images without the need for segmentation techniques, thereby bolstering the framework's reliability.

The study by Zaid Alyafeai et al. Alyafeai and Ghouti (2020) presents a fully automated pipeline for cervix detection and cervical cancer classification from cervigram images, comprising two pre-trained deep learning models. The first model achieves rapid cervix region detection, outperforming state-of-the-art models with an accuracy of 0.68 measured by the intersection of union (IoU). The second model utilizes self-extracted features to classify cervical tumors, demonstrating superior performance in accuracy and speed compared to existing models. The classifier achieves an AUC score of 0.82, enabling it to classify cervix regions 20 times faster. Its accuracy, speed, and lightweight design render it suitable for deployment on mobile phones, offering the potential to reduce cervical cancer deaths in less-developed nations. The authors propose future enhancements focusing on improving cervigram image quality and refining manual labeling of cervical regions of interest (RoI).

Recent studies shown in table 23 have made significant advancements in automating cervical cell segmentation and cervical cancer classification through the utilization of advanced deep learning techniques. For example, Deo et al. (2024) introduced CerviFormer, employing Query-Key-Value Attention and achieving high accuracy of 93.70% and 94.57% on different datasets. In 2023, several studies were conducted on cervical cancer detection and classification using deep learning techniques. Ji et al. (2023) proposed an automated cervical cell segmentation method employing various algorithms such as $Unet\_Resnet34$ and $DeepLabV3Plus\_Resnet50$, achieving accuracy ranging from 60.88% to 80.55%. Tan et al. (2023) explored cervical cancer classification using a range of CNN models including VGG-16 and ResNet-152, achieving accuracies between 71.45% and 87.02%. In 2022, Battula and Chandana (2022) proposed an EfficientNet-based method achieving an accuracy of 99.67%, while Pramanik et al. (2022) introduced a fuzzy distance-based ensemble model achieving an accuracy of 96.96%. Habtemariam et al. (2022) developed a system using MobileNetV2-YOLOv3 achieving an accuracy of 96.84%. Dhawan et al. (2021) and Alsalatie, Alquran, Mustafa, Mohd Yacob and Ali Alayed (2022) also presented methods with accuracy of 96.1% and 99.6%, respectively. Finally, in 2021, Chandran et al. (2021) proposed an ensemble deep learning network with an accuracy of 92.30% and 73.30% using CYENET and VGG19(TL) on colposcopy images. The summary of the provided information reveals a comprehensive exploration of deep-learning methodologies for automated cervical cancer detection and classification. Various studies conducted between 2021 and 2024 have employed a range of deep learning architectures, including CNNs, ensemble methods, and attention mechanisms, to achieve high accuracy in this critical area of medical image analysis. Notable contributions include the introduction of novel approaches such as CerviFormer by Deo et al. (2024) which utilizes cross-attention and latent transformer mechanisms, as well as the EfficientNet-based method proposed by Battula and Chandana (2022), achieving an exceptional accuracy of 99.67%. Furthermore, studies by Ji et al. (2023) and Tan et al. (2023) have explored deep ensemble learning and convolutional neural network models, respectively, achieving accuracy ranging from 60.88% to 87.02%. These developments underscore the potential of deep learning methods to transform cervical cancer screening and diagnosis, presenting promising pathways for improved healthcare outcomes via enhanced accuracy and efficiency in detection.

Table 2

Comparison of State-of-an-Art Deep Learning Based Cervical Cancer Classification

Year	Author Name	Algorithm Used	Accuracy (%)
2024	Deo et al. (2024)	Query-Key-Value(QKV) Attention:Multi-Layer Perceptrons Neural Networks	Sipakmed : 93.70 Herlev : 94.57
2023	Ji et al. (2023)	Unet_resnet34	74.11
		Unet_densenet121	75.06
		UnetPlusPlus_resnet34	80.55
		UnetPlusPlus_densenet121	77.31
		DeepLabV3_resnet34	60.88
		DeepLabV3_resnet50	64.19
		DeepLabV3Plus_resnet34	67.21
		DeepLabV3Plus_resnet50	73.53
2023	Tan et al. (2023)	Baseline CNN	83.82
		VGG-16	82.29
		VGG-19	83.82
		Densenet-121	84.73
		Densenet-169	85.26
		Densenet-201	87.02
		ResNet-50	74.45
		ResNet-101	74.96
		ResNet-152	71.45
		Inception	84.86
		Xception	86.72
		MobileNet	86.11
		MobileNet-v2	83.45
2022	Battula and Chandana (2022)	CYENET	92.30
		DenseNet121	72.42
		DenseNet169	69.79
		SVM	63.27
		Proposed Architecture	99.67
2022	Pramanik et al. (2022)	Inception V3	95.30
		MobileNet V2	93.92
		Inception ResNet V2	96.44
		Proposed ensemble technique (above listed model)	96.96
2022	Habtemariam et al. (2022)	MobileNetV2-YOLOv3	96.84
2022	Alsalatie et al. (2022)	Modified CNN	99.6
2022	Huang and Feng (2022)	Basic CNN	95.19
2021	Dhawan et al. (2021)	InceptionV3 ResNet50 VGG19	96.1
2021	Chandran et al. (2021)	CYENET	92.30
		VGG19(TL)	73.30

3. Methodology

A schematic view of the Proposed method is shown in the figure 3.

- The proposed system is trained using extensive data to determine whether patients have cervical cancer. This involves creating a predictive model based on the available data.

- The dataset's labels are divided into two groups according to cancer characteristics, and a deep learning algorithm is employed to construct a model for predicting cancer presence.

- The methodology consists of seven steps as shown in figure 3:

(1) Image Acquisition:
- Utilizes the Herlev dataset for cervical cell images.

(2) Preprocessing Steps:
- Image Normalization: Standardizes pixel values for consistency.
- Image Enhancement: Involve the application of a combination of CLAHE and Median Filtering along with NLM filter to enhance image quality.

(3) Image Augmentation:
- Applies augmentation: Regeneration of multiple images from single image techniques on pre-processed images.

(4) Transfer Learning:
- Adapts a pre-trained model on a large dataset for cervical cell characteristics.

(5) Network Layer Ensemble:
- Integrates multiple deep learning architectures including ResNet50v2, Resnet101, Densenet121, DenseNet169, XceptionNet, and InceptionResnetv2.
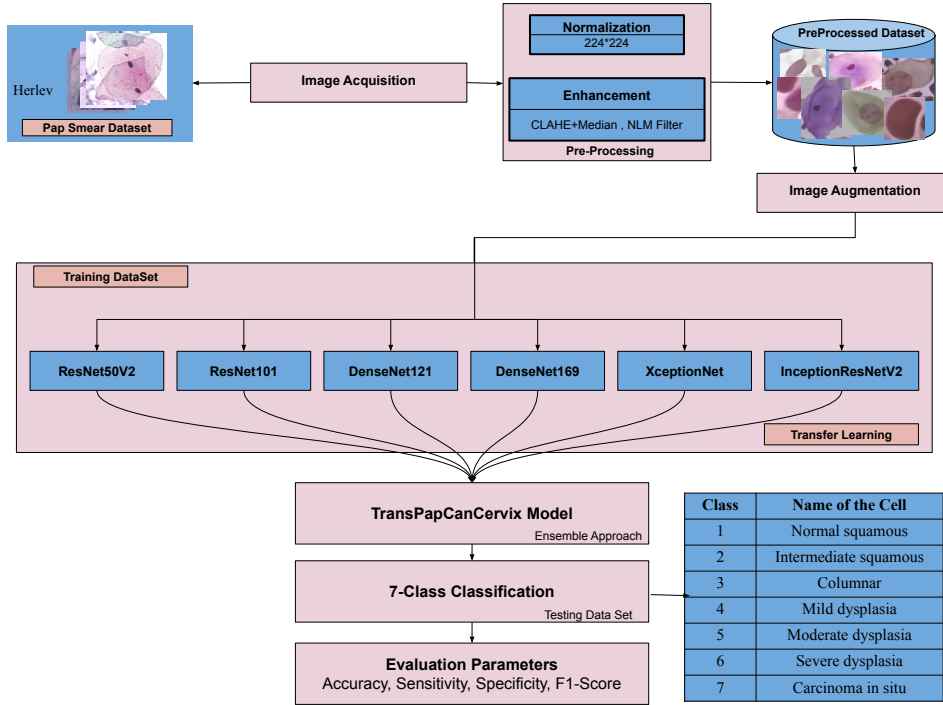
Figure 3: TransPapCanCervix Model Proposed Architecture

(6) Ensemble Classifier:
- Combines predictions from individual models to enhance the model's capacity to detect complex patterns.

(7) Output:
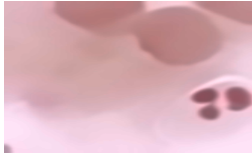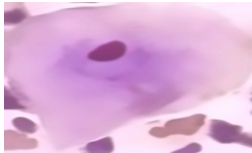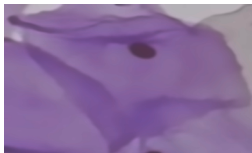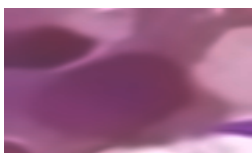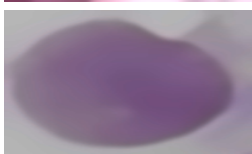- Provides predictions for normal and abnormal cervical cells.

This systematic approach, starting from image acquisition and including diverse preprocessing, augmentation, transfer learning, and ensemble techniques, aims to improve the accuracy and resilience of cervical cell classification Alyafeai and Ghouti (2020). The model's outcomes are highly valuable for disease diagnosis in medical image analysis.

3.1. Pre-training Phase
3.1.1. Datasets
During the image acquisition phase, the study utilized the Herlev dataset, a publicly accessible resource. This dataset, sourced from Herlev University Hospital in Denmark, consists of cervical cell images captured using a digital camera and microscope with a resolution of 0.02 µm/pixel by skilled cyto-technicians and doctors (Jantzen and Dounias, 2006) Shanthi, Faruqi, Hareesha and Kudva (2019); Singh and Goyal (2020). The tabular data shown in table 3 outlines the classification of cervical cell dataset types, categorizing them into normal and abnormal categories based on their characteristics Bhavsar and Shrimali (2023). In the normal category, three types of cells are identified: normal squamous, intermediate squamous, and columnar, totaling 242 cells. On the other hand, the abnormal category consists of four types: mild dysplasia, moderate dysplasia, severe dysplasia, and carcinoma in situ, totaling 675 cells. In total, there are 917 cell images analyzed Win, Kitjaidure, Paing and Hamamoto (2019). This categorization offers essential information on the distribution of cervical cell types, which is pivotal for comprehending and diagnosing cervical abnormalities, thereby advancing strategies for detecting and treating cervical cancer. The graphical representation is shown in figure 4 and figure 5. Under the category of 'Normal', there are three subtypes: 'Normal squamous' with 74 cells, 'Intermediate squamous' with 70 cells, and 'Columnar' with 98 cells. On the other hand, the 'Abnormal' category encompasses four subtypes: 'Mild dysplasia' with 182 cells, 'Moderate dysplasia' with 146 cells, 'Severe dysplasia' with 197 cells, and 'Carcinoma in situ' with 150 cells. These counts provide a detailed breakdown of the dataset, facilitating a comprehensive analysis of cell types and their occurrences within the sample set.

Table 3
Single Cell Herlev Dataset for Pap Smear Images

Class	Cell Type (Category)	Name of the Cancer Cell	Sample	Total Counts
Class 1	Normal	Normal Squamous		74 cells
Class 2	Normal	Intermediate Squamous		70 cells
Class 3	Normal	Columnar		98 cells
Class 4	Abnormal	Mild Dysplasia		182 cells
Class 5	Abnormal	Moderate Dysplasia		146 cells
Class 6	Abnormal	Severe Dysplasia		197 cells
Class 7	Abnormal	Carcinoma-In-Situ		150 cells
Total No. Of Images				**917**

3.1.2. Image Normalization

During the Image Normalization stage, the initial cervical cell images extracted from the dataset undergo resizing to achieve a standardized dimension of 224 x 224 pixels Dhawan et al. (2021). This resizing step is essential for maintaining consistency in the input data, guaranteeing that all images conform to a uniform format for subsequent processing. Normalization is pivotal in addressing image size discrepancies, facilitating smooth integration into DL models that necessitate uniform input dimensions. The resizing of images to 224 x 224 pixels not only adheres to the specifications of widely used pre-trained models but also improves computational efficiency during subsequent training and inference phases.

Pap Smear Image distribution in Herlev Dataset
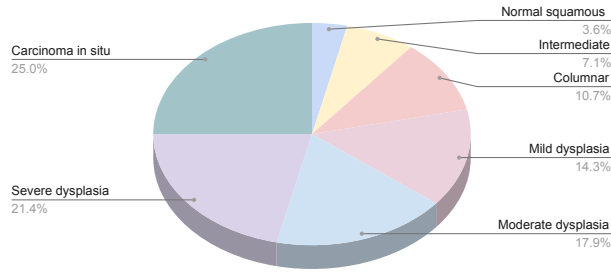
Figure 4: Pap Smear Image Categorical Distribution Graph
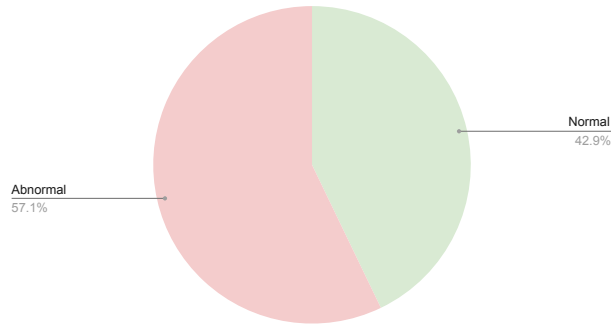
Count of Cell Type (Category)

Figure 5: Normal vs. Abnormal Pap Smear Image Cell Distribution Graph

3.1.3. Image Enhancement

The Image Enhancement stage is crucial for improving the quality of cervical cell images sourced from either the Herlev dataset or Pap smear images Bhavsar and Shrimali (2024). This phase utilizes two main methods: Non-Local Means (NLM) filtering Gayathri and Kumar (2021) and a blend of Contrast CLAHE with median filtering to enhance interpretability and feature representation Win et al. (2019).

(1) CLAHE and Median Filter:

- CLAHE enhances local contrast in images by ensuring uniform contrast distribution across different regions, thus preventing excessive noise amplification.
- Median Filtering is subsequently utilized to diminish noise and unwanted artifacts further, preserving edges while reducing the impact of outliers.

(2) Non-Local Means (NLM) Filter:

- The NLM filter refines images by exploiting non-local similarities, effectively denoising and preserving significant structures.
- By evaluating similarities between image patches, the NLM filter enhances the clarity of cervical cell features, particularly in areas with subtle variations and fine details.

These image enhancement techniques collectively refine and enrich the dataset, ensuring high-quality inputs for subsequent model stages. The objective is to furnish the neural network with superior inputs, facilitating enhanced learning and characterization of both normal and abnormal cervical cell patterns throughout the training process.

3.1.4. Image Augmentation

In the Image Augmentation phase, the dataset undergoes enrichment through diverse transformations applied to cervical cell images Alsalatie et al. (2022); Nambu et al. (2022); Rahaman et al. (2021). Utilizing both images enhanced by the Non-Local Means (NLM) filter from the previous step and those processed with Clahe and Median filters, further augmentation techniques are implemented. NLM augmentation introduces variability via random perturbations, enriching the dataset with augmented versions of NLM-filtered images. Clahe and Median augmentation adds diversity by applying random transformations to images enhanced using Clahe and Median filters. These approaches are essential for strengthening the model's resilience, guarding against overfitting, and ensuring its ability to generalize to unseen data. By integrating variations introduced by preprocessing techniques, the augmented dataset enhances the proposed architecture's adaptability and enhances its capacity to recognize intricate patterns in squamous cell images. The image augmentation method implemented for our proposed framework is illustrated in Figure 6.

```
datagen= ImageDataGenerator(
    rotation_range=45,      # Rotate the image between -45 to 45 degrees
    width_shift_range=0.2,  # Shift the width of the image by up to 20%
    height_shift_range=0.2, # Shift the height of the image by up to 20%
    shear_range=0.2,        # Apply shear transformation with a shear factor of 0.2
    zoom_range=0.2,         # Zoom the image by up to 20%
    horizontal_flip=True,   # Flip the image horizontally
    vertical_flip=True,     # Flip the image vertically
    fill_mode='nearest',
    rescale=1./255          #normalize pixel value
    # Add more augmentation parameters here
)
```

Figure 6: Pre-Processed Pap Smear Image Augmentation Approach

3.2. Transfer Learning Phase

Transfer learning, a fundamental technique in deep learning, involves utilizing a pre-trained neural network model on ImageNet Dataset as a starting point for a new task instead of commencing training from scratch Alyafeai and Ghouti (2020); Bhavsar and Shrimali (2023). By utilizing the knowledge acquired through extensive training on a sizable dataset for a similar task, this method enables quicker convergence and frequently produces superior results in contrast to building a model entirely from scratch. The process of transfer learning using popular network architectures Dhawan et al. (2021); Hussain et al. (2020); Wang et al. (2022) like DenseNet, ResNet, and Xception involves several key steps:

(1) Pre-trained Model Selection:
- Begin by selecting a pre-trained model that has been trained on a large dataset. DenseNet, ResNet, and XceptionNet are widely recognized deep learning architectures known for their effectiveness.

(2) Model Initialization:
- Load the selected pre-trained model along with its weights. Remove the top (output) layers, as they are specific to the original task the model was trained on.

(3) Customize the Top Layers:
- Customize the model by adding new layers tailored to the target task. These layers typically consist of a combination of fully connected (Dense) layers, along with other types of layers like dropout or normalization layers.

(4) Freeze Base Layers:
- Decide whether to unfreeze some of the lower layers (closer to the input) of the pre-trained model. This allows these layers to be updated during training, with their weights being changed accordingly.

(5) Compile the Model:

- Define the loss function, optimizer, and metrics suitable for the specific task at hand. This is typically done using functions provided by deep learning libraries like TensorFlow or PyTorch.

(6) Training:

- Train proposed model on target dataset. Since lower layers are being updated, this phase mainly focuses on training the newly added layers.

(7) Fine-tuning:

- Unfreeze certain initial layers and proceed with training using a reduced learning rate. This enables the model to adjust to the nuances of the dataset.

(8) Evaluation:

- Assess the model's performance by evaluating it on a validation or test set.

(9) Deployment and Inference:

- After the model has been trained and evaluated satisfactorily, it can be deployed for making predictions on new, unseen data.

3.3. Deep Learning Model Architectures

3.3.1. Architecture of ResNet50V2

ResNet50v2, an advanced CNN architecture, is renowned for its utilization of residual blocks, which are instrumental in capturing intricate features. Each residual block within ResNet50v2 is composed of convolutional layers, followed by batch normalization and ReLU activation functions. A distinctive feature illustrated in Figure 7 is the integration of shortcut connections. These connections facilitate deeper feature learning and alleviate overfitting by adding the output of each convolutional layer to the block's input. ResNet50v2 is organized into five stages, each containing a distinct set of residual blocks:

(1) This stage starts with a 7x7 kernel convolutional layer with a stride of 2, followed by a max-pooling layer with a 3x3 kernel and stride 2.

(2) Next phase comprises four residual blocks with a 3x3 kernel and stride 1, succeeded by a residual block with a 1x1 kernel and stride 1.

(3) following stage consists of consist six residual blocks with a 3x3 kernel and stride 1, followed by a residual block with a 1x1 kernel and stride 1.

(4) Further phase incorporates three residual blocks with a 3x3 kernel and stride 1, succeeded by a residual block with a 1x1 kernel and stride 1.

(5) Next stage Comprises a single residual block with a 3x3 kernel and stride 1, followed by another residual block with a 1x1 kernel and stride 1.

(6) Following the last residual block, the output is directed to a global average pooling layer, which computes the average of each channel in the previous layer's output, resulting in a 1x1x2048 tensor.

(7) The output of the global average pooling layer is then passed to a fully connected (FC) layer, with the number of nodes matching the classes in the classification task.

(8) Finally, a softmax layer is employed to process the FC layer's output, generating a probability distribution over the classes in the classification task.

3.3.2. ResNet101

The ResNet101 model as shown in figure 8 is a convolutional neural network (CNN) recognized for its residual network architecture, incorporating residual blocks to facilitate complex feature learning. Residual blocks within ResNet101 consist of a series of convolutional layers followed by batch normalization and ReLU activation functions, with each convolutional layer's output added to the block's input via a shortcut connection to enable deeper feature extraction and prevent overfitting. Structurally, ResNet101 comprises five stages, each featuring a residual block:

(1) starts with a 7x7 kernel convolutional layer and a subsequent 3x3 max pooling layer, resulting in a 112x112 feature map.

(2) Consists of four residual blocks with a 3x3 kernel and stride 1, followed by three residual blocks with a 1x1 kernel and stride 1, yielding a 28x28 feature map.
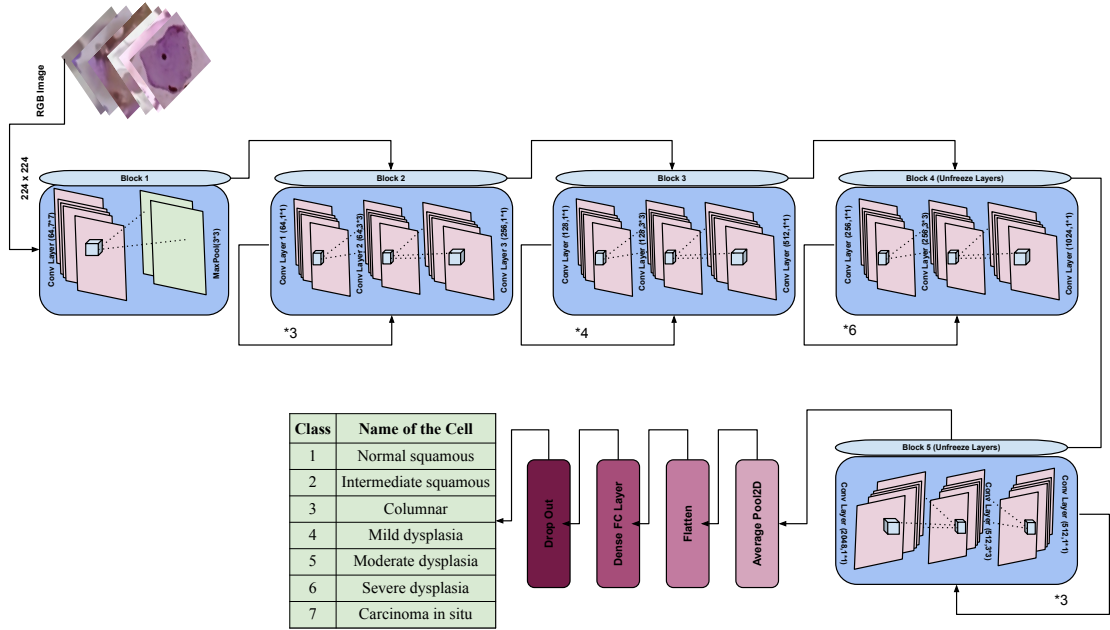
Figure 7: ResNet50V2 Network Layer Architecture

(3) Contains six residual blocks followed by three residual blocks, reducing the feature map to 14x14.

(4) Comprises three residual blocks followed by three residual blocks, further downsizing the feature map to 7x7. It includes a residual block followed by three residual blocks, maintaining a 7x7 feature map.

(5) Following the last residual block, the output is directed to a global average pooling layer, which computes the average of each channel, resulting in a 1x1x2048 tensor. Subsequently, a fully connected (FC) layer is applied, with the number of nodes equivalent to the classification task's class count.

(6) Finally, the output layer, serving as a softmax layer, generates a probability distribution over the classes, completing the ResNet101 architecture. Additionally, Dropout regularization is employed to prevent overfitting within the model.

3.3.3. DenseNet121

DenseNet121 shown in figure 9, is a convolutional neural network (CNN) introduced by Gao Huang, Zhuang Liu, Kilian Weinberger, and Laurens van der Maaten in 2016 Zhou, Zhu, Li, Yuan, Zheng and Liu (2022), is renowned for its dense connectivity, characterized by intricate interconnections among all layers, which promotes feature reuse and facilitates the effective extraction of complex features. Comprising four dense blocks, each followed by a transition layer, the model integrates convolutional layers within its dense blocks, each accompanied by batch normalization with ReLU activation functions. Transition layers incorporate both pooling and convolutional operations to manage feature map dimensions. Subsequently, the model's architecture channels the output of the last transition layer into a global average pooling layer, followed by a fully connected (FC) layer, ultimately leading to an output layer with nodes corresponding to the number of classes in the classification task. Each layer in the DenseNet121 architecture is meticulously crafted:
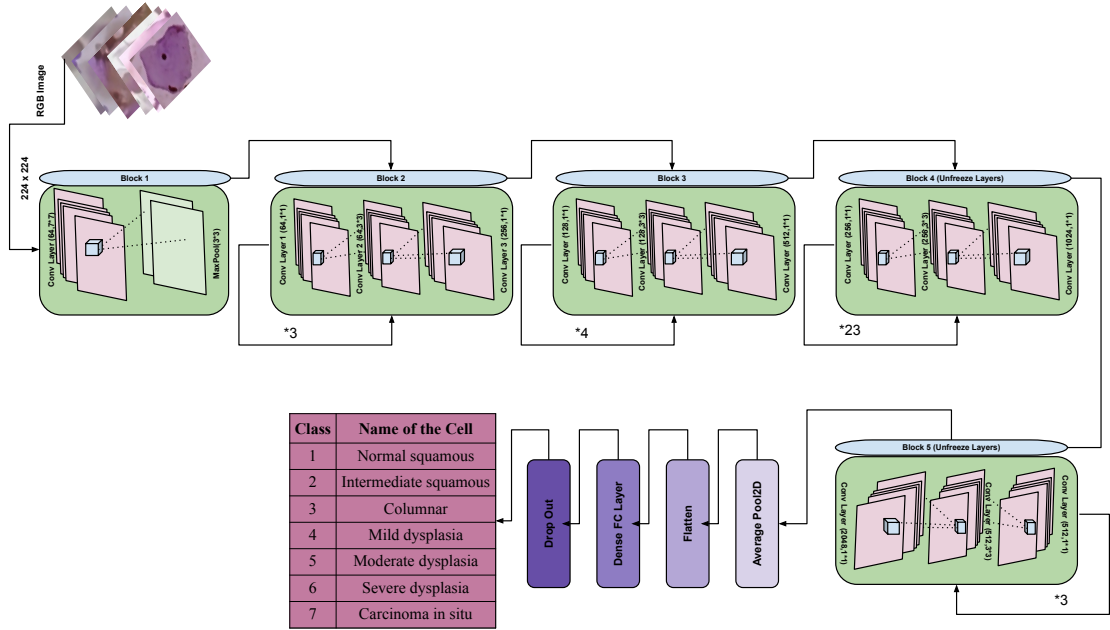
(1) Input layer: Receives 224x224 RGB images.

Figure 8: ResNet101 Network Layer Architecture

(2) Convolution layer (112, 7x7, stride 2): Executes convolution with a 7x7 kernel and stride 2, yielding a 112x112 feature map.

(3) Max pooling layer (56, 3x3, stride 2): Performs max pooling, producing a 56x56 feature map.

(4) Dense blocks: Contain convolutional layers followed by batch normalization with ReLU activation, promoting feature concatenation.

(5) Transition layers: Employ pooling and convolutional operations to regulate feature map dimensions.

(6) Global average pooling layer: Computes the average of each channel in the previous layer's output, generating a 1x1x1024 tensor.

(7) FC layer: Fully connected layer whose node count aligns with the classification task's class count.

(8) Output layer: Softmax layer offering a probability distribution over classification classes.

3.3.4. DenseNet169

The DenseNet169, a convolutional neural network (CNN) devised by Gao Huang, Zhuang Liu, Kilian Weinberger, and Laurens van der Maaten in 2016 Guan, Zhang, Sethares, Kijowski and Liu (2021), stands out for its dense connectivity, where every layer is directly connected to others, fostering feature reuse and effective learning of intricate features. This model is as shown in figure 10 and comprises four dense blocks, each succeeded by a transition layer, embodying a series of convolutional layers, coupled with batch normalization and ReLU activation functions. Transition layers integrate a pooling layer followed by a convolutional layer to manage feature map dimensions. The final transition layer feeds into a fully connected (FC) layer, followed by an output layer configured to accommodate the number of classes in the classification task. Detailed insights into each layer of the DenseNet169 architecture reveal its robust design:
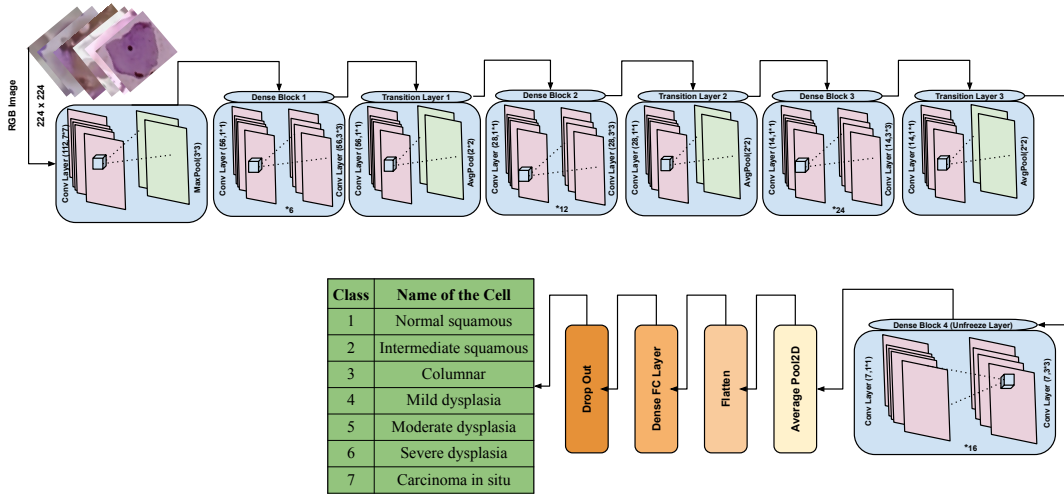
Figure 9: DenseNet121 Network Layer Architecture

(1) Input layer: Receives 224x224 RGB images.

(2) Convolution layer (112x112, 7x7, stride 2): Executes convolution with a 7x7 kernel and stride 2, producing a 112x112 feature map.

(3) Pooling layer (56x56, 3x3, stride 2): Performs max pooling, yielding a 56x56 feature map. Dense blocks: Consists of convolutional layers followed by batch normalization and ReLU activation, facilitating feature concatenation.

(4) Transition layers: Introduce pooling and convolutional operations to regulate feature map dimensions. FC layer: Fully connected layer whose node count aligns with the classification task's class count. Output layer: The Softmax layer provides a probability distribution over classification classes.

3.3.5. XceptionNet

XceptionNet, introduced by François Chollet in 2016 Shaheed, Mao, Qureshi, Kumar, Hussain, Ullah and Zhang (2022), is a convolutional neural network (CNN) architecture inspired by Inception and is as shown in figure 11. It utilizes depthwise separable convolutions instead of traditional convolutions, optimizing parameter efficiency. Xception employs depth-wise separable convolutions, a technique that significantly reduces trainable parameters in the CNN while preserving accuracy. This makes XceptionNet well-suited for resource-constrained environments like mobile and embedded applications. XceptionNet consists of 36 residual modules, each incorporating a depth-wise separable convolution followed by a point-wise convolution. These modules are organized into entry flow, middle flow, and exit flow segments as shown:

(1) Entry Flow: The entry flow comprises two convolutional layers, each followed by batch normalization and ReLU activation. The first layer employs a 3x3 kernel and stride 2, while the second uses a 3x3 kernel with stride 1.
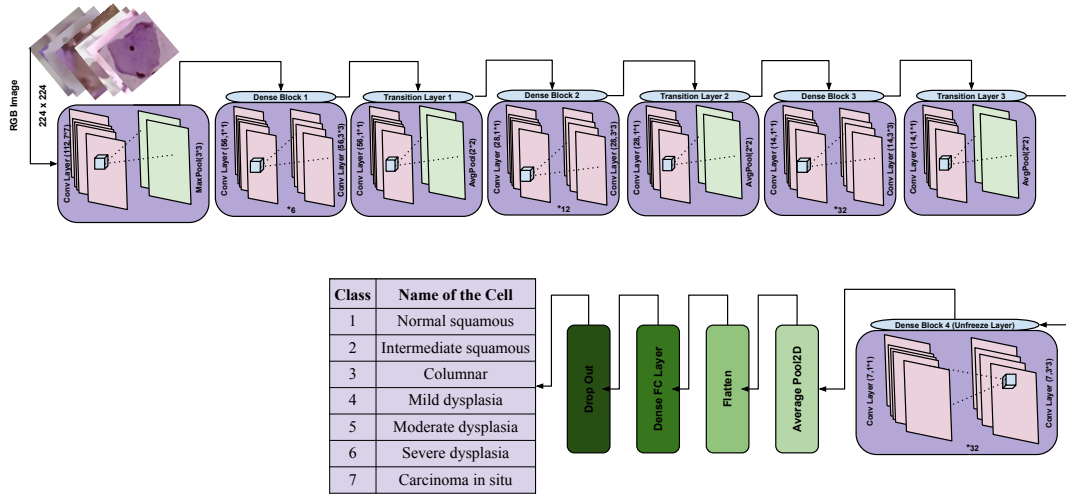
Figure 10: DenseNet169 Network Layer Architecture

(2) Middle Flow: The middle flow includes 36 residual modules, each featuring a depth-wise separable convolution with a 3x3 kernel and stride 1, followed by a point-wise convolution with a 1x1 kernel and stride 1. Shortcut connections are used to add the output of each module to its input.

(3) mirroring the structure of the entry flow, the exit flow comprises two convolutional layers featuring batch normalization with ReLU activation. The initial layer utilizes a 3x3 kernel with a stride of 1, while the subsequent layer employs a 1x1 kernel with a stride of 1.

(4) Global Average Pooling Layer: Following the exit flow, a global average pooling layer computes the average of each channel in the output, resulting in a 1x1x2048 tensor.

(5) Fully Connected (FC) Layer: The output of the global average pooling layer is fed into a fully connected layer, with the number of nodes equal to the number of classes in the classification task.

(6) Output Layer: Finally, a softmax output layer is applied to the FC layer's output, providing a probability distribution over the classes in the classification task.

3.3.6. InceptionResNetV2

The InceptionResNetv2 model, introduced by Szegedy et al. in 2016 Siciarz and McCurdy (2022), is a CNN combining Inception architecture with residual connections. Inception architecture as shown in figure 12 employs parallel convolution paths with varying filter sizes to enable feature learning at multiple scales, while residual connections aid in learning deeper features without overfitting. The model comprises two main components: The Stem block and the Inception block.

(1) The stem block reduces the input image size. The stem block begins with a 7x7 convolutional layer and is followed by a 3x3 max pooling layer to further reduce input dimensionality.
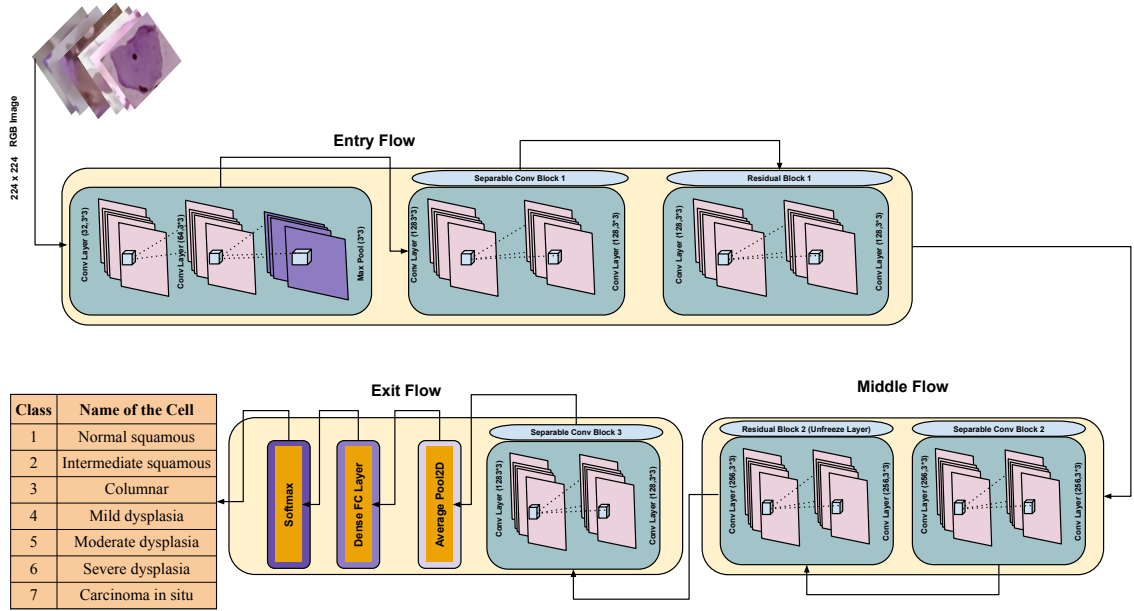
Figure 11: XceptionNet Network Layer Architecture

(2) The Inception blocks extract complex features. Inception blocks consist of different convolution paths, including 1x1, 3x3, and 5x5 convolutions, contributing to feature extraction.

(3) The architecture also includes Reduction blocks, containing max pooling layers and convolution paths, to reduce dimensionality.

(4) The model concludes with a global average pooling layer, followed by a fully connected layer with nodes corresponding to the classification task's class count.

(5) Finally, a softmax layer produces a probability distribution across classification classes based on the FC layer's output.

(6) Each layer within the InceptionResNetv2 architecture serves a specific purpose in feature extraction and classification.

4. Experiments

4.1. Experimental configuration

The experimental settings for this study encompassed specific hardware, software, and programming tools. The hardware configuration included an Intel Core i5-10,700 CPU, 16 GB of memory, and Nvidia GTX 3050 GPUs. On the software side, Ubuntu 20.04 served as the operating system, with CUDA Version 11.3 and Anaconda Version 4.10 installed. In terms of programming, Python Version 3.8 was utilized alongside the TensorFlow deep learning framework. Various libraries were employed for data manipulation, including NumPy, SciPy, Scikit-learn, and Pandas. For data augmentation, the Albumentations library was employed, while the *segmentation_models_tensorflow* library facilitated deep learning model implementation. Additionally, the progress bar functionality was incorporated
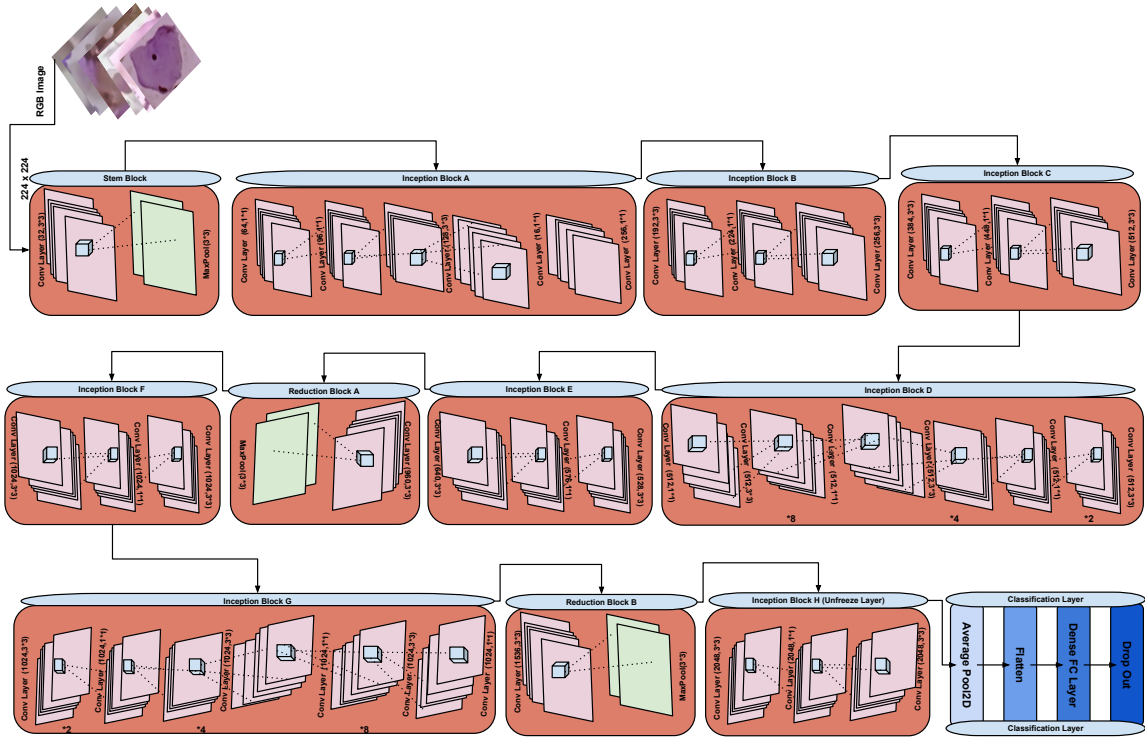
Figure 12: InceptionResNetv2 Network Layer Architecture

using the Tqdm library. These carefully selected components formed the foundation for conducting experiments and analysis in the study.

4.2. PreProcessing

In the preprocessing phase, the cervical cell images are resized to a standardized dimension of 224 x 224 pixels, ensuring uniformity in input data for further processing. This step is crucial for seamless integration into deep learning models, aligning with the requirements of pre-trained models and enhancing computational efficiency. Meanwhile, in the Image Enhancement phase, techniques like Non-Local Means (NLM) filtering and a combination of CLAHE with Median filter are employed to improve image quality Gayathri and Kumar (2021); Win et al. (2019). CLAHE ensures uniform contrast distribution across different regions, while Median Filtering suppresses noise and preserves edges. The NLM filter enhances image clarity by leveraging non-local similarities, particularly beneficial in capturing subtle variations. These enhancements contribute to a more informative dataset for subsequent model stages. In the Image Augmentation phase, the dataset is enriched with diverse transformations, including NLM-augmented CLAHE and Median-augmented images, to enhance model robustness and prevent overfitting. These approaches guarantee the model's flexibility and enhance its capability to recognize intricate patterns in cervical cell images, thereby improving disease diagnosis accuracy.

4.3. Implementation Details

The experimental settings for this study entail a hardware configuration comprising an Intel Core i5-10,700 CPU, 16 GB memory, and Nvidia GTX 3050 GPUs. The software environment includes Ubuntu 20.04 as the operating system, CUDA Version 11.3, and Anaconda Version 4.10. Python 3.8 serves as the programming language, with

Tensorflow as the deep learning framework. Data manipulation relies on libraries such as NumPy, SciPy, Scikit-learn, and Pandas, while Albumentations is utilized for data augmentation. For deep learning models, segmentation models, TensorFlow is employed, and progress tracking is facilitated by the Tqdm library. These settings form the basis for conducting rigorous experiments in the domain of deep learning. The implementation details of the proposed system for cervical cancer classification involve a comprehensive methodology spanning seven key steps. Initially, the system acquires cervical cell images utilizing the Herlev dataset. Subsequently, a series of preprocessing steps are applied, including image normalization for standardizing pixel values and image enhancement techniques to improve image quality. Augmentation techniques are then employed, incorporating CLAHE, Median, and Non-Local Means (NLM) augmentation to diversify the dataset. Transfer learning is utilized by adapting a pre-trained model trained on the Imagenet dataset to extract features specific to cervical cell characteristics. A network layer ensemble approach is adopted, integrating multiple deep learning architectures including ResNet50v2, Resnet101, Densenet121, DenseNet169, XceptionNet, and InceptionResnetv2 Tan et al. (2023). Subsequently, an ensemble classifier combines predictions from individual models to enhance the proposed architecture's ability to capture intricate patterns. Finally, the output of the system provides predictions distinguishing between normal and abnormal cervical cells Bhavsar and Shrimali (2023, 2024). This methodical process, spanning from image capture to ensemble methods, seeks to substantially enhance the precision and resilience of cervical cell classification, thereby streamlining disease diagnosis within the realm of medical image analysis.

4.4. Evaluation Metrics

The effectiveness of the classification models across the seven classes was thoroughly evaluated by utilizing a confusion matrix, which provided a detailed analysis of predictions within various categories. Alongside conventional measures like accuracy, precision, and recall, we utilized more sophisticated metrics to obtain a more nuanced understanding of the models' performance.

(1) Weighted Average Metrics: The weighted average precision, recall, and F1 score are computed to address potential class imbalances, providing a consolidated assessment of the model's performance in correctly classifying instances across all classes. Specifically, the metrics of weighted average $precision(P_{\text{weighted}})$, $Recall(R_{\text{weighted}})$, $F1-score(F1_{\text{weighted}})$ and are employed for this evaluation as shown in equations 1, 2, and 3 .

$$P_{\text{weighted}} = \frac{\sum_{i=1}^{7} w_i \cdot P_i}{\sum_{i=1}^{7} w_i} \tag{1}$$

where:

- P_{weighted} represents the weighted average precision.
- w_i represents the weight associated with class i.
- P_i represents the precision of class i.

$$R_{\text{weighted}} = \frac{\sum_{i=1}^{7} w_i \cdot R_i}{\sum_{i=1}^{7} w_i} \tag{2}$$

where:

- R_{weighted} represents the weighted average recall.
- w_i represents the weight associated with class i.
- R_i represents the recall of class i.

$$F1_{\text{weighted}} = \frac{\sum_{i=1}^{7} w_i \cdot F1_i}{\sum_{i=1}^{7} w_i} \tag{3}$$

where:

- $F1_{\text{weighted}}$ represents the weighted average F1-score.
- w_i represents the weight associated with class i.

- $F1_i$ represents the F1 score of class i.

(2) Macro-Averaged Accuracy: To gauge overall model performance, we computed the macro-averaged accuracy (A_{macro}) providing an equal-weighted average of accuracy across all classes:

$$A_{\mathrm{macro}} = \frac{\sum_{i=1}^{7} A_i}{7} \tag{4}$$

where:

- A_{macro} represents equal weighted average accuracy.
- A_i is the accuracy for each class.

Figure 13 showcases a graph demonstrating the precision scores for cervical cancer cell classification across several deep-learning network model architectures. Notably, ResNet50V2 and ResNet101 stand out with impressively high precision scores of 0.9838 and 0.9817, respectively, indicating their remarkable accuracy in correctly identifying cervical cell classes. Following closely, DenseNet169 and XceptionNet also demonstrate strong performance, achieving precision scores of 0.9506 and 0.9636, respectively. Although slightly lower, DenseNet121 and InceptionResNet50V2 still exhibit respectable precision scores of 0.9118 and 0.9479, respectively. On average, these models attain a macro-averaged precision score of 0.9566, underscoring their collective effectiveness in cervical cancer cell classification. These results underscore the reliability and robustness of deep learning architectures in accurately categorizing cervical cell images across various classes, offering promising prospects for disease diagnosis and treatment.
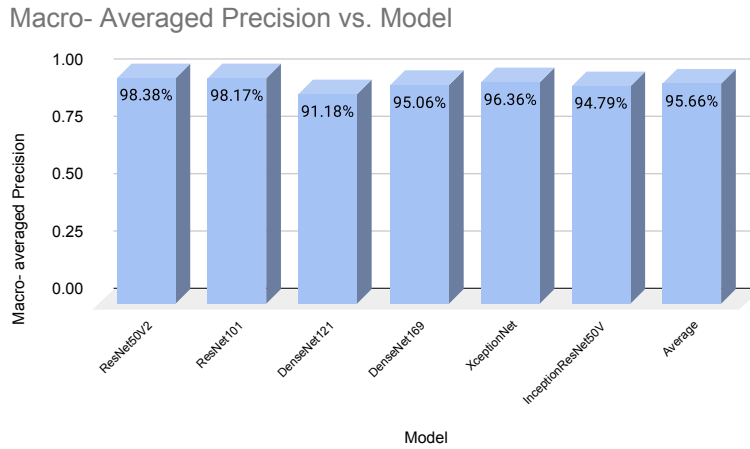
Figure 13: Macro- Averaged Precision For 7-Class Classification

The graphical data displayed in figure 14 illustrates the recall performance of diverse deep learning network model architectures in the context of 7-class cervical cancer cell classification. Remarkably, ResNet50V2 and ResNet101 exhibit exceptional recall scores of 0.9836 and 0.9817, respectively, showcasing their proficiency in accurately retrieving instances of various cervical cell classes. DenseNet169 and XceptionNet also showcase noticeable recall performance, achieving scores of 0.949 and 0.963, respectively. While slightly lower, DenseNet121 and InceptionResNet50V2 still showcase respectable recall rates of 0.9097 and 0.9471, respectively. On average, these models attain a macro-averaged recall score of 0.9557, underscoring their collective efficacy in identifying cervical cancer cell classes accurately. These results highlight the robustness and reliability of deep learning approaches in cervical cancer cell classification, suggesting their potential utility in clinical diagnostics and disease management.

The graphical representation shown in figure 15 showcases the performance of different deep learning network model architectures concerning the F1-score in the context of 7-class cervical cancer cell classification. Notably, ResNet50V2 and ResNet101 exhibit exceptional F1-scores of 0.9837 and 0.9817, respectively, highlighting their

Macro- Averaged Recall vs. Model
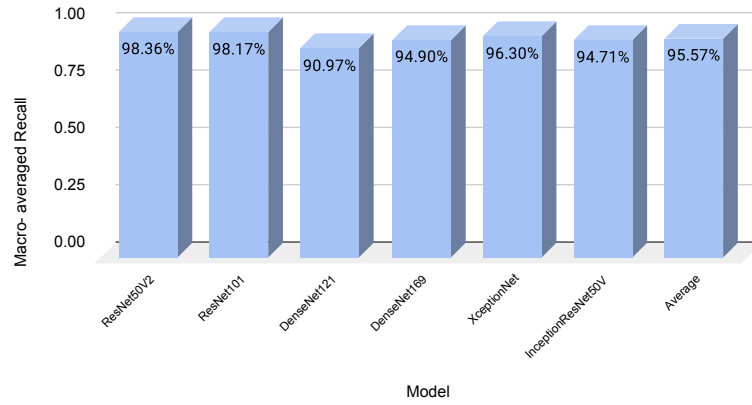
Figure 14: Macro- Averaged Recall For 7-Class Classification

effectiveness in achieving a balance between precision and recall for various cervical cell classes. DenseNet169 and XceptionNet also represent high performance, with F1-scores of 0.9492 and 0.9628 respectively, indicating their robustness in accurately identifying cervical cancer cell types. Although slightly lower, DenseNet121 and InceptionResNet50V2 still showcase respectable F1-scores of 0.9095 and 0.947, respectively. On average, these models achieve a macro-averaged F1-score of 0.9557, underscoring their collective efficacy in cervical cancer cell classification. These results highlight the reliability and utility of deep learning approaches in clinical diagnostics and disease management for cervical cancer. The graph shown in figure 16 illustrates the performance of different deep

Macro- Averaged F1-score vs. Model
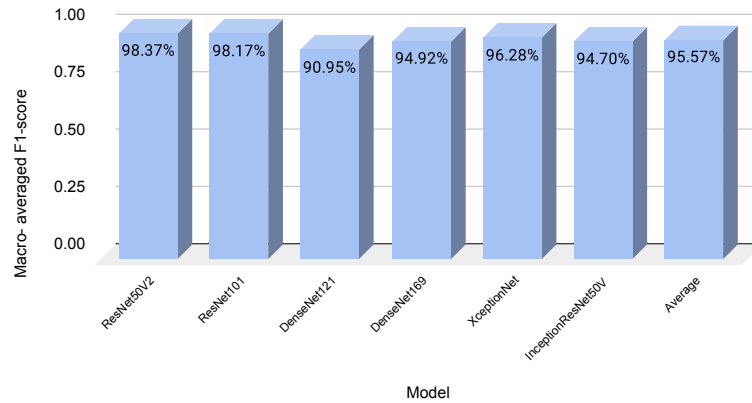
Figure 15: Macro- Averaged F1-Score For 7-Class Classification

learning network model architectures concerning accuracy in the domain of 7-class cervical cancer cell classification. Notably, ResNet101 emerges as the top performer with an exceptional macro-averaged accuracy of 0.9965, closely followed by ResNet50V2 and DenseNet121, achieving accuracies of 0.9948 and 0.9943, respectively. XceptionNet and InceptionResNet50V2 also demonstrate strong accuracy, both achieving a score of 0.9859. Although slightly lower, DenseNet169 still achieves a respectable accuracy of 0.9859. On average, these models achieve a macro-averaged accuracy of 0.9912, highlighting their collective efficacy in accurately classifying cervical cancer cell types. These

results underscore the reliability and utility of deep learning approaches in medical diagnostics, particularly in the context of cervical cancer detection and classification
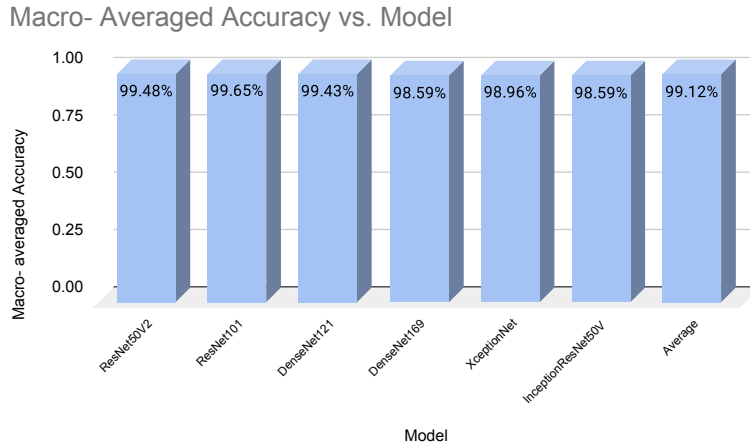
Figure 16: Macro- Averaged Accuracy For 7-Class Classification
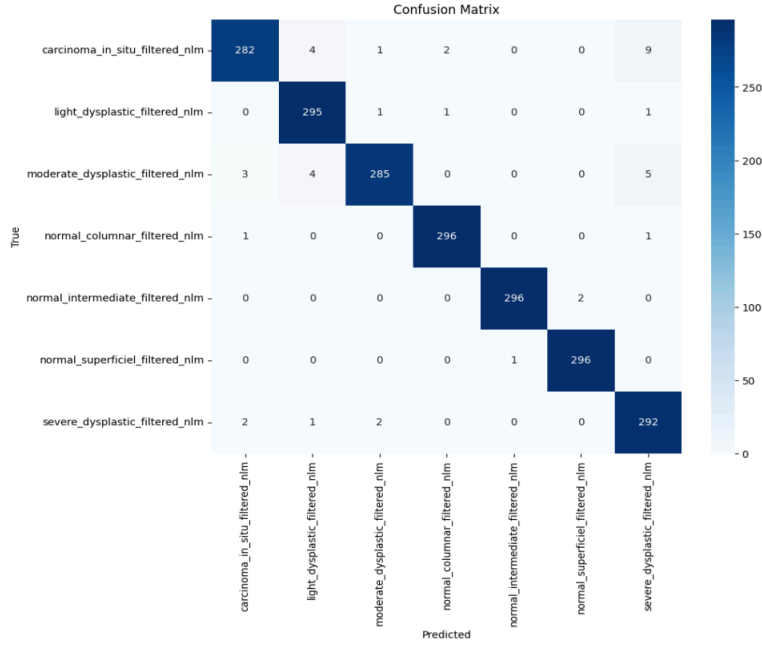
4.5. Confusion Matrix

A confusion matrix for the classification of cervical cancer cells into 7 classes offers a comprehensive breakdown of the model's predictive accuracy for each class. It comprises rows and columns that respectively denote the predicted and actual classes. Each cell in the matrix contains the number of instances where predictions match or mismatch the ground truth for a particular class. In a 7-class classification task for cervical cancer cells, the confusion matrix could reveal how many instances of each class were correctly classified as TP (true positives), incorrectly classified as another class FN (false negatives), or incorrectly classified as the given class FP (false positives). This information helps to assess the model's accuracy, precision, recall, and F1-score for each class, enabling insights into areas where the model may perform well or require improvement. Overall, the confusion matrix provides a comprehensive overview of the model's classification performance across all classes, aiding in the evaluation and refinement of the classification system.

4.5.1. Confusion Matrix for ResNet50V2

The confusion matrix for ResNet50V2, as depicted in the provided figure 17, offers an insightful examination of its classification performance across seven distinct classes. Each cell within this matrix delineates the count of instances attributed to specific categories, including True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). In the case of ResNet50V2, the analysis reveals 295 instances correctly identified as True Positives (TP), 282 instances accurately labeled as True Negatives (TN), only 4 instances erroneously classified as False Positives (FP), and notably, no instances misclassified as False Negatives (FN). This comprehensive breakdown thoroughly evaluates the model's efficacy in discerning between different classes. The varying intensities observed within the matrix visually illustrate the distribution of instances across categories, thereby aiding in pinpointing both the strengths and weaknesses inherent in Model A's predictions. Such a visual representation serves as a valuable tool in refining and optimizing the model's performance across diverse classification tasks.

4.5.2. Confusion Matrix for ResNet101

The confusion matrix for ResNet101, as illustrated in the accompanying figure 18, offers a comprehensive evaluation of its classification proficiency across seven distinct classes. Within this matrix, each cell denotes the number of instances assigned to specific categories, encompassing TP, TN, FP, and FN. For ResNet101, the breakdown reveals 292 instances correctly identified as true positives (TP) and an equal number of instances accurately labeled as true negatives (TN), with no false positives (FP) recorded. However, there were 4 instances erroneously classified as false negatives (FN). The varying intensities within the matrix provide a visual depiction of the distribution of

Figure 17: Confusion matrix for ResNet50V2

instances across categories, facilitating an assessment of Model A's performance across individual classes. This visual representation serves as a valuable tool for pinpointing both the model's strengths and weaknesses in making predictions across diverse classification tasks.

4.5.3. Confusion Matrix for DenseNet121

The confusion matrix for DenseNet121, as depicted in the provided figure 19, offers an insightful analysis of its classification accuracy across seven distinct classes. Each cell within this matrix signifies the number of instances categorized as true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Specifically, DenseNet121 achieved 253 instances correctly identified as true positives (TP) and 280 instances accurately labeled as true negatives (TN). However, the model encountered 7 instances incorrectly classified as false positives (FP) and 8 instances erroneously labeled as false negatives (FN). The varying intensities observed within the matrix enable a visual assessment of the distribution of instances across categories, facilitating the identification of both strengths and weaknesses in the model's predictions. This visual representation serves as a valuable tool in refining and enhancing the model's performance across diverse classification tasks.

4.5.4. Confusion Matrix for DenseNet169

The confusion matrix for DenseNet169, depicted in the accompanying figure 20, offers a comprehensive insight into its classification performance across seven distinct classes. Each cell within this matrix delineates the count of instances categorized as true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Specifically, DenseNet169 demonstrated proficiency by correctly identifying 268 instances as true positives (TP) and accurately classifying 290 instances as true negatives (TN). Despite this, the model encountered only one instance erroneously labeled as a false positive (FP) and four instances incorrectly classified as false negatives (FN). The varying intensities observed within the matrix provide a visual assessment of the distribution of instances across categories, facilitating the identification of both strengths and weaknesses in the model's predictions. Such visual representation proves invaluable in refining and optimizing the model's performance across diverse classification tasks.
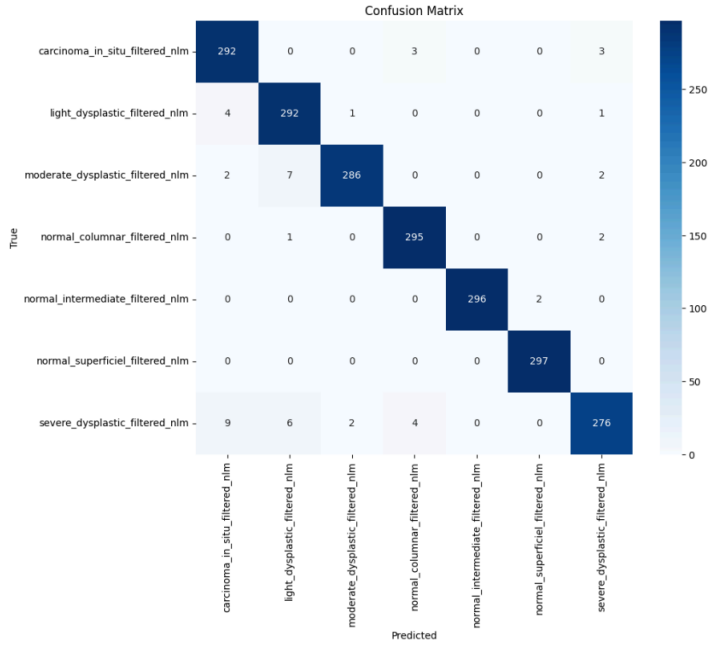
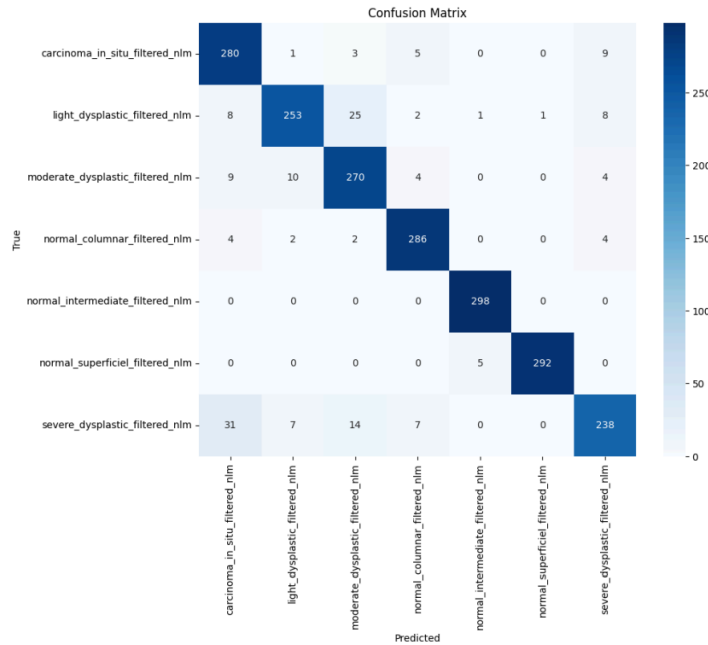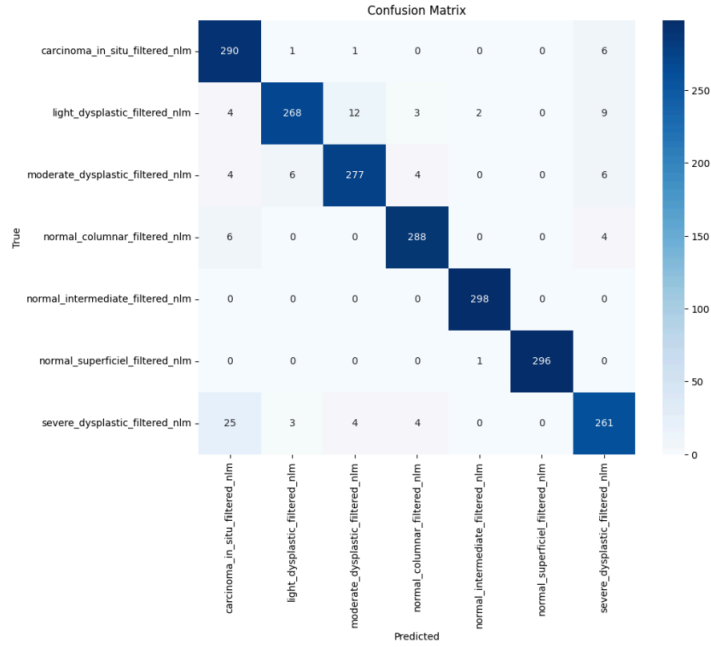**Figure 18:** Confusion matrix for ResNet101

Figure 19: Confusion matrix for DenseNet121

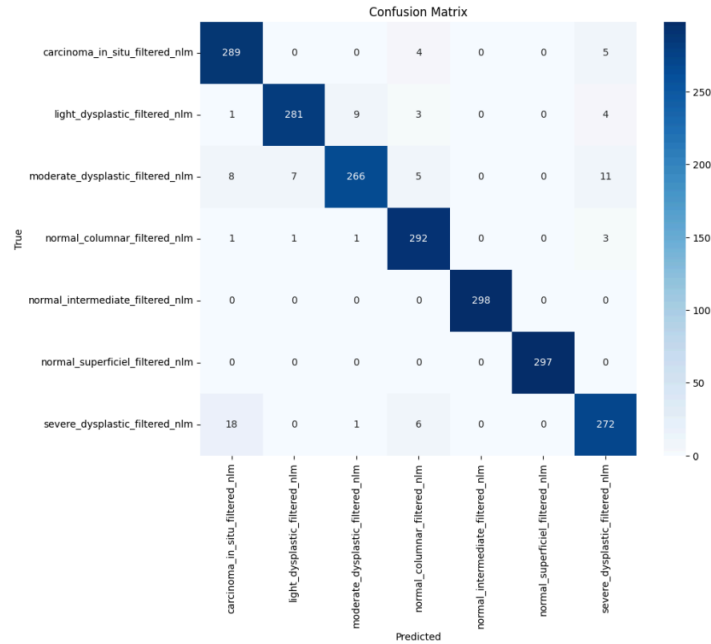### 4.5.5. Confusion Matrix Of XceptionNet

The confusion matrix for XceptionNet, as illustrated in the provided figure 21, offers a detailed examination of its classification accuracy across seven distinct classes. Each cell within this matrix indicates the count of instances attributed to specific categories. Notably, XceptionNet achieved 281 instances correctly identified as true positives (TP) and 289 instances accurately labeled as true negatives (TN). Impressively, the model encountered no instances

Figure 20: Confusion matrix for DenseNet169

incorrectly classified as false positives (FP), highlighting its precision in distinguishing between classes. However, there was a single instance erroneously labeled as a false negative (FN). The varying intensities observed within the matrix allow for a visual assessment of the distribution of instances across categories, aiding in the identification of both strengths and weaknesses in the model's predictions. This visual representation serves as a valuable tool in evaluating and refining the model's performance across diverse classification tasks.

Figure 21: Confusion matrix for XceptionNet

4.5.6. InceptionResNetV2

The confusion matrix for InceptionResNetV2, as depicted in the provided figure, offers a comprehensive overview of its classification performance across seven distinct classes. Each cell within the matrix represents the count of instances assigned to specific categories. In the case of InceptionResNetV2, the analysis reveals 279 instances correctly identified as true positives (TP), alongside 289 instances accurately labeled as true negatives (TN). Despite its overall proficiency, the model encountered a single instance misclassified as a false positive (FP) and three instances erroneously labeled as false negatives (FN). The varying intensities observed within the matrix allow for a visual assessment of the distribution of instances across categories, facilitating the identification of both strengths and weaknesses in the model's predictions. This visual representation serves as a valuable tool in refining and optimizing the model's performance across diverse classification tasks. The summary of the provided graph displayed in figure
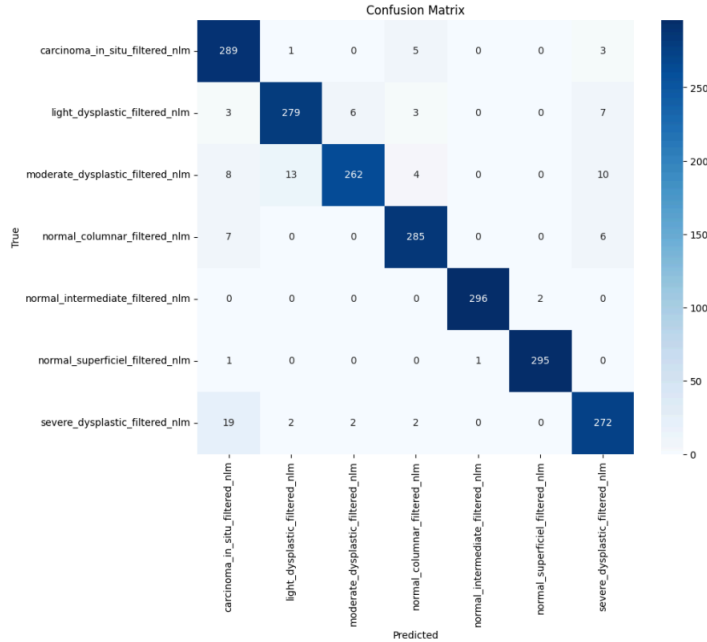
**Figure 22:** Confusion matrix for InceptionResNetV2

23 showcases the classification performance of several models, detailing their TP, TN, FP, and FN counts. Across the models, variations in results are evident. For instance, ResNet50V2 and ResNet101 exhibited high TP and TN rates with minimal FP and FN instances. DenseNet121 and DenseNet169 encountered higher FP and FN counts, indicating some misclassifications. XceptionNet demonstrated excellent performance with a low FP count but encountered a single FN instance. InceptionResNetV2 also performed well overall but had a slightly higher FP count and encountered three FN instances. These insights aid in understanding the efficacy and limitations of each model in accurately classifying instances, providing valuable information for further analysis and optimization.

5. Discussion

The graphical display shown in figure 24 provides insights into the performance metrics of various deep learning network models in the task of 7-class cervical cancer cell classification. ResNet101 emerges as the standout performer, showcasing high precision, recall, F1-score, and accuracy (0.9817, 0.9817, 0.9817, and 0.9965, respectively). ResNet50V2 also exhibits strong performance, particularly excelling in precision and recall (0.9838 and 0.9836) while maintaining an impressive F1-score and accuracy (0.9837 and 0.9948). Though DenseNet121 and DenseNet169 show slightly lower precision and recall compared to ResNet models, they maintain notable accuracy. XceptionNet consistently performs well across all metrics, with InceptionResNet50V2 showing similar performance but with slightly lower scores. In summary, these findings highlight the effectiveness of deep learning models in precisely
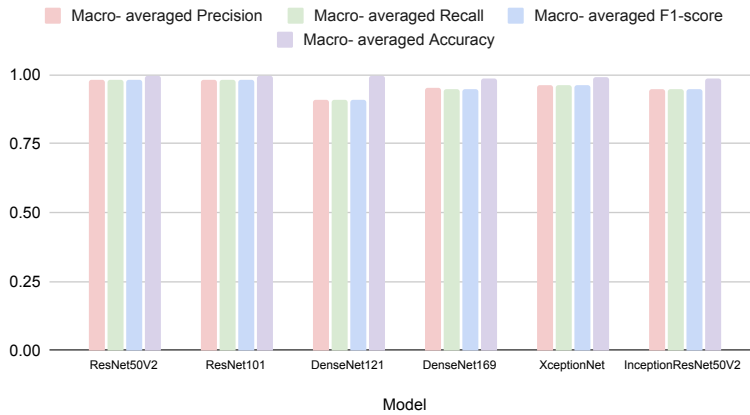
Classification Performance from Confusion Matrix
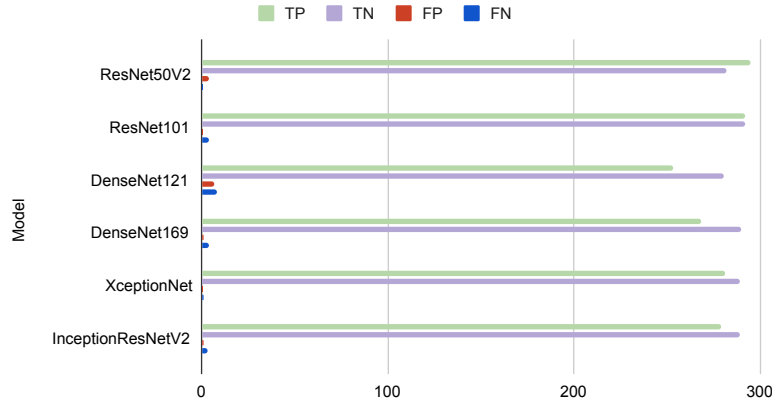
Figure 23: Classification Performance from Confusion Matrix

classifying various types of cervical cancer cells, with ResNet101 and ResNet50V2 emerging as the most notable performers.

Figure 24: Performance metrics for Selected DL model Architectures for 7 Class Classification

The accuracy graph helps us see how well our model is learning during training. Usually, it gets better at predicting the training data over time. But if it gets too good at the training data, it might not do well with new, unseen data. That's what we call overfitting. We look at how close the accuracy of the training data compared to new data. A small difference means our model is doing well on new data too. If the accuracy of new data starts to go down, it means our model might be too focused on the training data. This could happen if we train it for too long or if we don't have enough different examples in our training data. If the accuracy stops improving, it means our model has probably learned as much as it can. Continuing to train it won't make it better.

The accuracy graphs provided for ResNet50V2, ResNet101, DenseNet121, DenseNet169, XceptionNet, and InceptionResNetV2 models shown in table 4 demonstrate their learning capabilities and generalization to unseen data. In all cases, both training and validation accuracies increase over time, indicating effective learning. Although validation accuracies are slightly lower than training accuracies, the small gap suggests minimal overfitting to the training data. Furthermore, the presence of plateaus in performance indicates that continued training is unlikely to

significantly improve model performance. Overall, these models exhibit robust performance and effective learning on the provided tasks. Additionally, the accuracy graph of InceptionResNetV2 indicates slight overfitting to the training data, but the model still demonstrates good generalization capabilities, making it a suitable choice for the given task.

The loss graph, depicted as a line graph, illustrates the model's loss on both training and validation datasets over time. Loss measures how accurately the model predicts outputs, with lower values indicating better performance. This graph serves as a valuable tool for monitoring the training process, aiding in the identification of potential issues such as overfitting and underfitting. Overfitting arises when the model excessively learns the training data, leading to poor generalization to new data, evident as a large gap between training and validation losses on the graph. Conversely, underfitting occurs when the model inadequately learns the training data, resulting in high losses for both training and validation sets. A desirable loss graph exhibits decreasing trends for both training and validation losses. Eventually, the validation loss should surpass the training loss, indicating effective learning of the training data and robust generalization to new data.

The loss graphs for ResNet50V2, ResNet101, DenseNet121, DenseNet169, XceptionNet, and InceptionResNetV2 models shown in table 5 illustrate the training and validation loss trends over epochs. In all cases, both training and validation losses decrease with increasing epochs, indicating effective learning and generalization to new data. Notably, the training loss consistently remains lower than the validation loss, suggesting minimal overfitting to the training data. Overall, these graphs signify that the models are learning well and generalizing effectively to validation data. However, for ResNet101 and InceptionResNetV2, slight deviations in validation loss trends hint at potential overfitting, necessitating careful monitoring to ensure optimal model performance.

6. Conclusion and Future Work

This study showcases notable advancements in medical image analysis, utilizing transfer learning techniques with prominent models such as ResNet50V2, ResNet101, DenseNet121, DenseNet169, XceptionNet, and InceptionRes-Netv2. By meticulously preprocessing the Herlev dataset and employing augmentation methods, we prepared it for robust classification. A key milestone of this endeavor was the application of ensemble techniques, consolidating the strengths of individual models to create a highly accurate predictive tool, each achieving over 97% accuracy. This achievement holds great promise for clinical applications, offering a reliable means of distinguishing normal and abnormal cellular images, particularly in the context of Squamous Cell Carcinoma for cervical cancer. The success of this project underscores the potential for similar approaches to transform medical diagnostics, potentially leading to earlier cancer detection. Moreover, the methodology and techniques employed serve as a valuable blueprint for future research in deep learning-based medical image analysis, marking a significant advancement in leveraging technology for healthcare improvement. This study has demonstrated remarkable progress in the Healthcare sector through the application of advanced deep-learning techniques. Leveraging transfer learning and ensemble methods, we achieved exceptional accuracy in classifying cellular images from the Herlev dataset.

Generating an Internet Of Things (IOT) driven tool and Collaborations with medical professionals and institutions can facilitate the translation of research findings into clinical practice, ultimately improving patient outcomes.
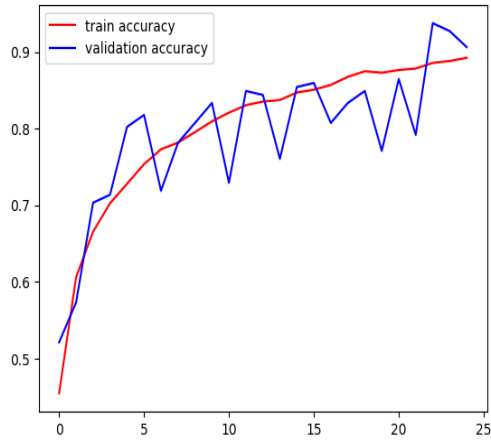
Table 4

Deep Learning Network Model Architecture Accuracy Graph: (A) ResNet50v2 Accuracy graph, (B) ResNet101 Accuracy graph, (C) DenseNet121 Accuracy graph, (D) DenseNet169 Accuracy graph, (E) XceptionNet Accuracy Graph, (F) InceptionResNetv2 Accuracy graph
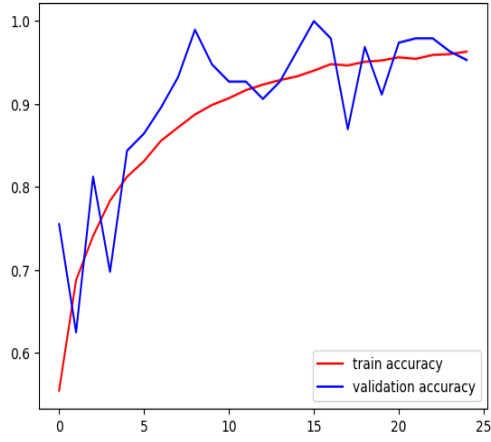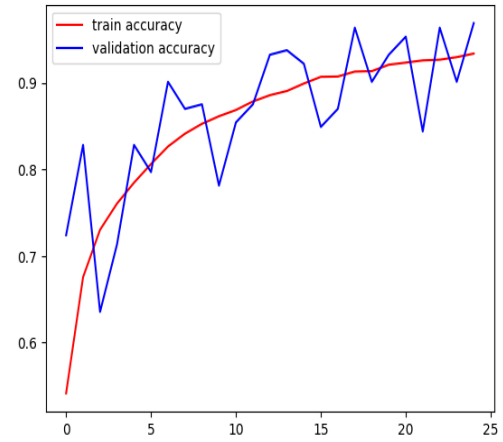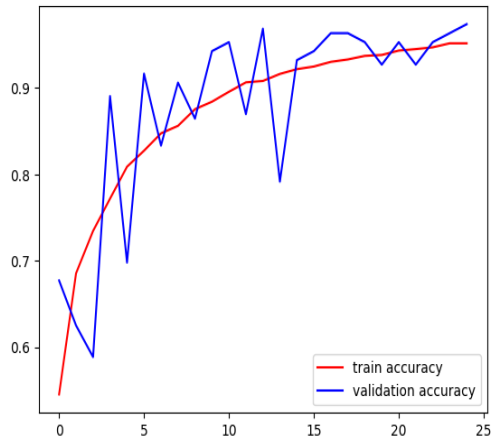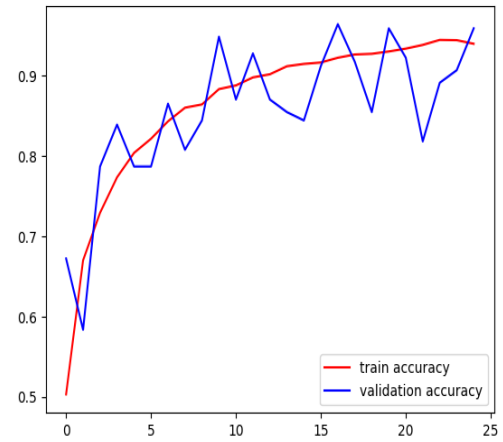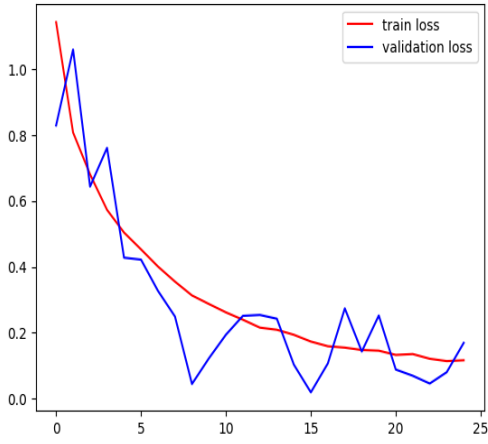

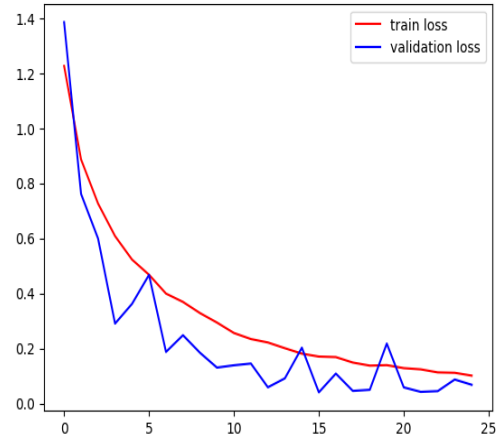(A)
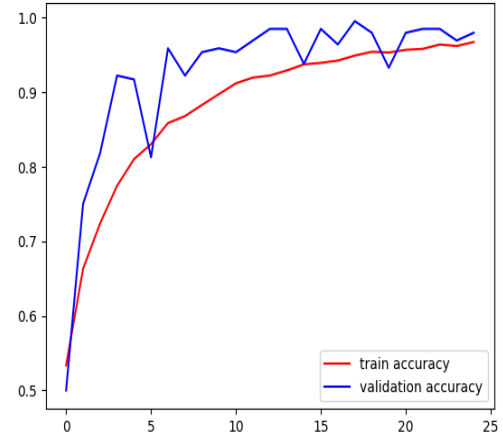
(B)
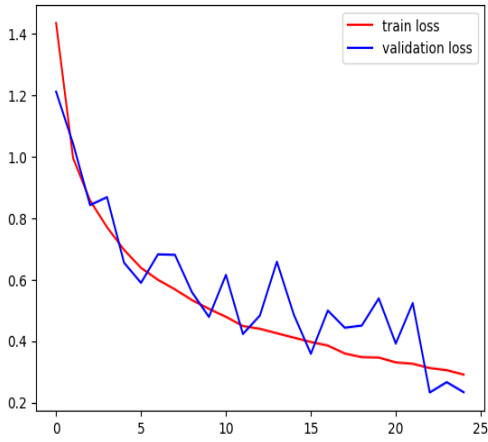

(C)

(D)

(E)

(F)

Table 5
Deep Learning Network Model Architecture Loss Graph: (A) ResNet50v2 Loss graph, (B) ResNet101 Loss graph, (C) DenseNet121 Loss graph, (D) DenseNet169 Loss graph, (E) XceptionNet Loss Graph, (F) InceptionResNetv2 Loss graph
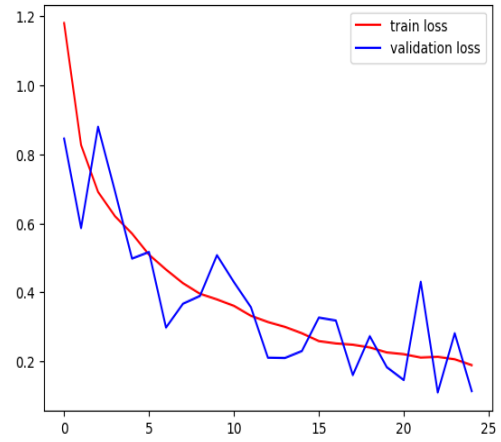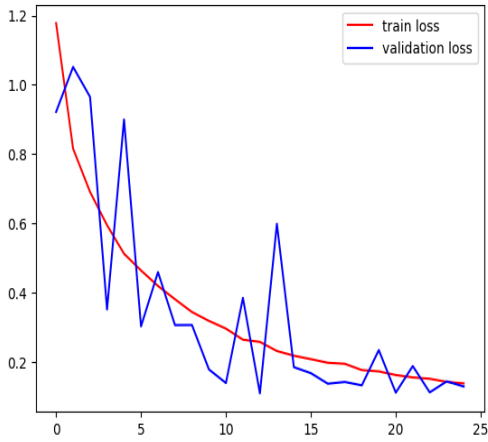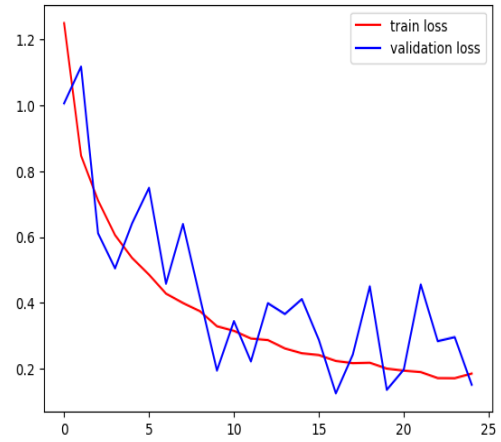

(A)

(B)

(C)

(D)

(E)

(F)

References

Allahqoli, L., Laganà, A.S., Mazidimoradi, A., Salehiniya, H., Günther, V., Chiantera, V., Karimi Goghari, S., Ghiasvand, M.M., Rahmani, A., Momenimovahed, Z., et al., 2022. Diagnosis of cervical cancer and pre-cancerous lesions by artificial intelligence: a systematic review. Diagnostics 12, 2771.

Alquran, H., Mustafa, W.A., Qasmieh, I.A., Yacob, Y.M., Alsalatie, M., Al-Issa, Y., Alqudah, A.M., 2022. Cervical cancer classification using combined machine learning and deep learning approach. Comput. Mater. Contin 72, 5117–5134.

Alsalatie, M., Alquran, H., Mustafa, W.A., Mohd Yacob, Y., Ali Alayed, A., 2022. Analysis of cytology pap smear images based on ensemble deep learning approach. Diagnostics 12, 2756.

Alyafeai, Z., Ghouti, L., 2020. A fully-automated deep learning pipeline for cervical cancer classification. Expert Systems with Applications 141, 112951.

Arifianto, D., Agoes, A.S., 2021. Cervical cancer image classification using cnn transfer learning, in: 2nd International Seminar of Science and Applied Technology (ISSAT 2021), Atlantis Press. pp. 145–149.

Arora, M., Dhawan, S., Singh, K., 2021. Deep learning in health care: Automatic cervix image classification using convolutional neural network, in: Mobile Radio Communications and 5G Networks: Proceedings of MRCN 2020, Springer. pp. 145–151.

Arora, M., Dhawan, S., Singh, K., 2023. Cervical cancer diagnosis and prediction: An application of machine learning techniques, in: Computational Intelligence in Analytics and Information Systems. Apple Academic Press, pp. 213–226.

Attallah, O., 2023a. Cercan· net: Cervical cancer classification model via multi-layer feature ensembles of lightweight cnns and transfer learning. Expert Systems with Applications 229, 120624.

Attallah, O., 2023b. Cervical cancer diagnosis based on multi-domain features using deep learning enhanced by handcrafted descriptors. Applied Sciences 13, 1916.

Battula, K.P., Chandana, B.S., 2022. Deep learning based cervical cancer classification and segmentation from pap smears images using an efficientnet. International Journal of Advanced Computer Science and Applications 13.

Bhatt, A.R., Ganatra, A., Kotecha, K., 2021. Cervical cancer detection in pap smear whole slide images using convnet with transfer learning and progressive resizing. PeerJ Computer Science 7, e348.

Bhavsar, B., Shrimali, B., 2023. Neural network based diagnostic approach for cervical cancer cell stratification from pap smear images, in: 2023 IEEE 11th Region 10 Humanitarian Technology Conference (R10-HTC), IEEE. pp. 699–703.

Bhavsar, B., Shrimali, B., 2024. Meticulous review: Cutting-edge cervix cancer stratification using image processing and machine learning. International Journal of Computing and Digital Systems 15, 1343–1358.

Bruni, L., Albero, G., Serrano, B., Mena, M., Collado, J., Gómez, D., Muñoz, J., Bosch, F., Sanjosé, S., 2023. Ico/iarc information centre on hpv and cancer. Human Papillomavirus and Related Diseases in the World. Summary Report .

Chandran, V., Sumithra, M., Karthick, A., George, T., Deivakani, M., Elakkiya, B., Subramaniam, U., Manoharan, S., et al., 2021. Diagnosis of cervical cancer based on ensemble deep learning network using colposcopy images. BioMed Research International 2021.

Chen, W., Li, X., Gao, L., Shen, W., 2020. Improving computer-aided cervical cells classification using transfer learning based snapshot ensemble. Applied Sciences 10, 7292.

Deo, B.S., Pal, M., Panigrahi, P.K., Pradhan, A., 2024. Cerviformer: A pap smear-based cervical cancer classification method using cross-attention and latent transformer. International Journal of Imaging Systems and Technology 34, e23043.

Devi, S., Gaikwad, S.R., Harikrishnan, R., 2023. Prediction and detection of cervical malignancy using machine learning models. Asian Pacific journal of cancer prevention: APJCP 24, 1419.

Dhawan, S., Singh, K., Arora, M., 2021. Cervix image classification for prognosis of cervical cancer using deep neural network with transfer learning. EAI Endorsed Transactions on Pervasive Health and Technology 7.

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

Gabutti, G., Ogliastro, M., 2024. History of human papilloma virus vaccination with a focus on italy. Women 4, 42–52.

Gayathri, R., Kumar, M.M., 2021. Medical image restoration using non-local means algorithm. Annals of the Romanian Society for Cell Biology , 2307–2314.

Ghoneim, A., Muhammad, G., Hossain, M.S., 2020. Cervical cancer classification using convolutional neural networks and extreme learning machines. Future Generation Computer Systems 102, 643–649.

Guan, B., Zhang, J., Sethares, W.A., Kijowski, R., Liu, F., 2021. Spectral domain convolutional neural network, in: ICASSP 2021-2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE. pp. 2795–2799.

Habtemariam, L.W., Zewde, E.T., Simegn, G.L., 2022. Cervix type and cervical cancer classification system using deep learning techniques. Medical Devices: Evidence and Research , 163–176.

Hemalatha, K., Vetriselvi, V., 2022. Deep learning based classification of cervical cancer using transfer learning, in: 2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC), IEEE. pp. 134–139.

Huang, W., Feng, H., 2022. Chronic cervicitis and cervical cancer detection based on deep learning of colposcopy images toward translational pharmacology. Frontiers in Pharmacology 13, 911962.

Hull, R., Mbele, M., Makhafola, T., Hicks, C., Wang, S.M., Reis, R.M., Mehrotra, R., Mkhize-Kwitshana, Z., Kibiki, G., Bates, D.O., et al., 2020. Cervical cancer in low and middle-income countries. Oncology letters 20, 2058–2074.

Hussain, E., Mahanta, L.B., Das, C.R., Talukdar, R.K., 2020. A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network. Tissue and Cell 65, 101347.

Ji, J., Zhang, W., Dong, Y., Lin, R., Geng, Y., Hong, L., 2023. Automated cervical cell segmentation using deep ensemble learning. BMC Medical Imaging 23, 137.

Kalbhor, M.M., Shinde, S.V., 2023. Cervical cancer diagnosis using convolution neural network: feature learning and transfer learning approaches. Soft Computing , 1–11.

Karani, H., Gangurde, A., Dhumal, G., Gautam, W., Hiran, S., Marathe, A., 2022. Comparison of performance of machine learning algorithms for cervical cancer classification, in: 2022 Second International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), IEEE. pp. 1–7.

Khamparia, A., Gupta, D., de Albuquerque, V.H.C., Sangaiah, A.K., Jhaveri, R.H., 2020. Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning. The Journal of Supercomputing 76, 8590–8608.

Khan, A., Han, S., Ilyas, N., Lee, Y.M., Lee, B., 2023. Cervixformer: A multi-scale swin transformer-based cervical pap-smear wsi classification framework. Computer Methods and Programs in Biomedicine 240, 107718.

Khobragade, V., Jain, N., Sisodia, D.S., 2020. Deep transfer learning model for automated screening of cervical cancer cells using multi-cell images, in: Applied Informatics: Third International Conference, ICAI 2020, Ota, Nigeria, October 29–31, 2020, Proceedings 3, Springer. pp. 409–419.

Kruczkowski, M., Drabik-Kruczkowska, A., Marciniak, A., Tarczewska, M., Kosowska, M., Szczerska, M., 2022. Predictions of cervical cancer identification by photonic method combined with machine learning. Scientific Reports 12, 3762.

Li, X., Xu, Z., Shen, X., Zhou, Y., Xiao, B., Li, T.Q., 2021. Detection of cervical cancer cells in whole slide images using deformable and global context aware faster rcnn-fpn. Current Oncology 28, 3585–3601.

Lilhore, U.K., Poongodi, M., Kaur, A., Simaiya, S., Algarni, A.D., Elmannai, H., Vijayakumar, V., Tunze, G.B., Hamdi, M., 2022. Hybrid model for detection of cervical cancer using causal analysis and machine learning techniques. Computational and Mathematical Methods in Medicine 2022.

Lu, J., Song, E., Ghoneim, A., Alrashoud, M., 2020. Machine learning for assisting cervical cancer diagnosis: An ensemble approach. Future Generation Computer Systems 106, 199–205.

Nambu, Y., Mariya, T., Shinkai, S., Umemoto, M., Asanuma, H., Sato, I., Hirohashi, Y., Torigoe, T., Fujino, Y., Saito, T., 2022. A screening assistance system for cervical cytology of squamous cell atypia based on a two-step combined cnn algorithm with label smoothing. Cancer Medicine 11, 520–529.

Park, Y.R., Kim, Y.J., Ju, W., Nam, K., Kim, S., Kim, K.G., 2021. Comparison of machine and deep learning for the classification of cervical cancer based on cervicography images. Scientific Reports 11, 16143.

Paul, P.R., Bhowmik, M.K., Bhattacharjee, D., 2015. Automated cervical cancer detection using pap smear images, in: Proceedings of Fourth International Conference on Soft Computing for Problem Solving: SocProS 2014, Volume 1, Springer. pp. 267–278.

Plissiti, M.E., Dimitrakopoulos, P., Sfikas, G., Nikou, C., Krikoni, O., Charchanti, A., 2018. Sipakmed: A new dataset for feature and image based classification of normal and pathological cervical cells in pap smear images, in: 2018 25th IEEE international conference on image processing (ICIP), IEEE. pp. 3144–3148.

Pramanik, R., Biswas, M., Sen, S., de Souza Júnior, L.A., Papa, J.P., Sarkar, R., 2022. A fuzzy distance-based ensemble of deep models for cervical cancer detection. Computer Methods and Programs in Biomedicine 219, 106776.

Priya, S., Karthikeyan, N., Palanikkumar, D., 2023. Pre screening of cervical cancer through gradient boosting ensemble learning method. Intelligent Automation & Soft Computing 35.

Rahaman, M.M., Li, C., Yao, Y., Kulwa, F., Wu, X., Li, X., Wang, Q., 2021. Deepcervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques. Computers in Biology and Medicine 136, 104649.

Rahimi, M., Akbari, A., Asadi, F., Emami, H., 2023. Cervical cancer survival prediction by machine learning algorithms: a systematic review. BMC cancer 23, 341.

Ratul, I.J., Al-Monsur, A., Tabassum, B., Ar-Rafi, A.M., Nishat, M.M., Faisal, F., 2022. Early risk prediction of cervical cancer: A machine learning approach, in: 2022 19th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), IEEE. pp. 1–4.

Shaheed, K., Mao, A., Qureshi, I., Kumar, M., Hussain, S., Ullah, I., Zhang, X., 2022. Ds-cnn: A pre-trained xception model based on depth-wise separable convolutional neural network for finger vein recognition. Expert Systems with Applications 191, 116288.

Shanthi, P., Faruqi, F., Hareesha, K., Kudva, R., 2019. Deep convolution neural network for malignancy detection and classification in microscopic uterine cervix cell images. Asian Pacific journal of cancer prevention: APJCP 20, 3447.

Siciarz, P., McCurdy, B., 2022. U-net architecture with embedded inception-resnet-v2 image encoding modules for automatic segmentation of organs-at-risk in head and neck cancer radiation therapy based on computed tomography scans. Physics in Medicine & Biology 67, 115007.

Singh, M., Jha, R.P., Shri, N., Bhattacharyya, K., Patel, P., Dhamnetiya, D., 2022. Secular trends in incidence and mortality of cervical cancer in india and its states, 1990-2019: data from the global burden of disease 2019 study. BMC cancer 22, 149.

Singh, S.K., Goyal, A., 2020. Performance analysis of machine learning algorithms for cervical cancer detection. International Journal of Healthcare Information Systems and Informatics (IJHISI) 15, 1–21.

Stelzle, D., Tanaka, L.F., Lee, K.K., Khalil, A.I., Baussano, I., Shah, A.S., McAllister, D.A., Gottlieb, S.L., Klug, S.J., Winkler, A.S., et al., 2021. Estimates of the global burden of cervical cancer associated with hiv. The lancet global health 9, e161–e169.

Sundarambal, B., Karthikeyini, C., Bommi, R., Subramanian, S., Jacintha, V., 2022. Comparison of machine learning and deep learning models for cervical cancer prediction, in: 2022 6th International Conference on Devices, Circuits and Systems (ICDCS), IEEE. pp. 135–138.

Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F., 2021. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians 71, 209–249.

Tan, S.L., Selvachandran, G., Ding, W., Paramesran, R., Kotecha, K., 2023. Cervical cancer classification from pap smear images using deep convolutional neural network models. Interdisciplinary Sciences: Computational Life Sciences , 1–23.

Tanimu, J.J., Hamada, M., Hassan, M., Kakudi, H., Abiodun, J.O., 2022. A machine learning method for classification of cervical cancer. Electronics 11, 463.

Tripathi, A., Arora, A., Bhan, A., 2021. Classification of cervical cancer using deep learning algorithm, in: 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE. pp. 1210–1218.

Vargas-Cardona, H.D., Rodriguez-Lopez, M., Arrivillaga, M., Vergara-Sanchez, C., García-Cifuentes, J.P., Bermúdez, P.C., Jaramillo-Botero, A., 2023. Artificial intelligence for cervical cancer screening: Scoping review, 2009–2022. International Journal of Gynecology & Obstetrics .

Wang, W., Tian, Y., Xu, Y., Zhang, X.X., Li, Y.S., Zhao, S.F., Bai, Y.H., 2022. 3cde-net: a cervical cancer cell detection network based on an improved backbone network and multiscale feature fusion. BMC Medical Imaging 22, 130.

Win, K.P., Kitjaidure, Y., Paing, M.P., Hamamoto, K., 2019. Cervical cancer detection and classification from pap smear images, in: Proceedings of the 2019 4th International Conference on Biomedical Imaging, Signal Processing, pp. 47–54.

Xue, P., Wang, J., Qin, D., Yan, H., Qu, Y., Seery, S., Jiang, Y., Qiao, Y., 2022. Deep learning in image-based breast and cervical cancer detection: a systematic review and meta-analysis. NPJ digital medicine 5, 19.

Youneszade, N., Marjani, M., Pei, C.P., 2023. Deep learning in cervical cancer diagnosis: architecture, opportunities, and open research challenges. IEEE Access 11, 6133–6149.

Zhou, Q., Zhu, W., Li, F., Yuan, M., Zheng, L., Liu, X., 2022. Transfer learning of the resnet-18 and densenet-121 model used to diagnose intracranial hemorrhage in ct scanning. Current Pharmaceutical Design 28, 287–295.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., He, Q., 2020. A comprehensive survey on transfer learning. Proceedings of the IEEE 109, 43–76.

Authors' Biographies

Barkha Bhavsar: She is a Ph.D. candidate in Computer Engineering at LDRP Institute of Technology and Research, Kadi Sarva Vishwavidyalaya, SVKM. Her research focuses on advanced topics in computer engineering, with a particular emphasis on addressing healthcare-related problems. She is dedicated to pushing the boundaries of knowledge and contributing valuable insights to her field, especially in the context of enhancing healthcare through innovative technological solutions. Main areas of interest include Image Processing, Artificial Intelligence, Computer Algorithms and data structures, Machine Learning and Deep Learning. Mail: barkha.bhavsar@gmail.com

Dr. Bela Shrimali: She is currently working as an Assistant Professor in the Computer Science and Engineering Department of the Institute of technology, Nirma university, Gujarat, India. She completed her Ph.D in the Cloud computing domain in the year 2018 and a master's in computer science and engineering from Government engineering college, Gandhinagar-India. She has 17 years of teaching experience. Dr. Shrimali has presented and published many research papers in various conferences and journals of international/national repute. Also, she delivers expert lectures and talks at different institutes. She is a registered PhD guide at Kadi Sarva Vishwavidyalaya- KSV university- Gandhinagar. Also, she is a life time member of ISTE and ACM. She has guided many PG students for research dissertations. A few of her main subjects of interest are Blockchain technology, IoT, Cloud Computing, Wireless Sensor Networks, and Machine learning. Mail: bela.shrimali@gmail.com