DeepCervixNet: An Advanced Deep Learning Approach for Cervical Cancer Classification in Pap Smear Images

Dhani Bux Talpur [1,2*]**, Ali Raza Rang** [1,2]**Ashfaque Khowaja** [3]**, Asadullah Shah** [2]

[1]Department of Information and Computing, University of Sufism and Modern Science, Bhitshah, Pakistan; [2]Kulliyyah of Information and Communication Technology, Islamic International University Malaysia; [3]School of Computer Science and Engineering Central South University Changsha, China

Abstract

Cervical cancer is among the leading causes of female mortality, emphasizing the significance of early detection and treatment to prevent its spread. While Pap smear images are widely used for cervical cancer screening, the manual diagnostic method is time-consuming and prone to error. The research article introduces DeepCervixNet, an innovative automated computerized approach designed for detecting cervical cancer in Pap smear images. In this study, we enhance ResNet101 and DenseNet169, state-of-the-art Convolutional Neural Network (CNN) architectures, by integrating the sequence and excitation (SE) blocks. Subsequently, Ensemble learning is employed to utilize the extracted features and classify the final output. The Harlev dataset was employed to test our model, with Gaussian smoothing and median filtering applied for image enhancement. This resulted in an overall improvement in the performance of the model. DeepCervixNet had an accuracy of 99.89% in cervical cells. The study's findings validate our model's robustness and efficacy, proving its superiority over a majority of current state-of-the-art models used to classify cervical cells, including standard ResNet and DenseNet architectures without SE blocks.

1 Introduction

Cervical cancer (CC) is the seventh most common disease worldwide and the fourth most common cancer in women. It currently ranks as a leading cause of death, particularly in impoverished countries. Main risk factors include early sexual activity, smoking, multiple partners, early pregnancies, a weakened immune system, the use of oral contraceptives for pregnancy prevention, and poor menstrual hygiene. Signs of cervical cancer encompass

abnormal vaginal bleeding, vaginal discharge, and mild pain during sexual activity [1]. Prevention of cervical cancer necessitates early detection and effective treatment.

The Papanicolaou smear, commonly known as the Pap smear test, plays a crucial role in lowering mortality and morbidity rates by detecting cervical cancer before it progresses to a malignant stage [2]. Therefore, a widely employed method for identifying precancerous lesions or cancerous cells in the cervical region is Pap smear testing. A well-performed Pap test can determine the stage of cervical cancer and categorize cells using standard procedures. Analyzing malignant cells through medical image processing and intelligent systems proves to be more cost-effective and time-efficient than traditional methods like Pap smears, colposcopy, and cervicography. Despite offering objective assessments, these procedures cannot replace the subjective evaluation of a specialized physician. Nevertheless, they can provide vital support to doctors, enabling them to make accurate diagnoses.

According to recent studies, computer-aided methods can significantly contribute to the automation of cancer detection. While there have been encouraging findings, it's important to note that there are still existing issues or challenges that need to be fixed with the ways that cervical cancer is currently diagnosed and classified. Firstly, the Pap smear test heavily depends on the manual examination and analysis of microscope images, as indicated by the authors in [3–7] . With almost three million cells that vary in orientations, sizes, and shapes—many of which are overlapping—this manual assessment method has proven to be arduous, costly, extremely time-consuming, and highly prone to human error. Secondly, an increasing body of research has been dedicated to the development of computer-aided diagnostic (CAD) technologies that can automatically classify aberrant cervical cells from cytology specimens, offering a contrast to manual assessment.

Convolutional neural networks (CNNs) have proven to be extremely effective in medical imaging. ResNet101 [8] and DenseNet169 [9] are notable for their powerful architectures and unique strengths. ResNet101, a Residual Network with 101 layers, employs residual learning to effectively train deep networks and integrates a sequence and excitation (SE) block for enhanced feature recalibration, demonstrating success in image classification tasks [10]. DenseNet169, characterized by its dense connectivity, promotes feature reuse and gradient flow. The addition of SE blocks further refines its capabilities, making DenseNet169 an efficient and robust choice for medical image analysis, balancing model complexity and performance [11].

Ensemble learning [12] combines the predictions of multiple models to enhance performance and accuracy. By leveraging the strengths of models like ResNet101 and DenseNet169, it mitigates individual shortcomings and produces more robust predictions. This approach improves generalization and performance, making it popular in machine learning applications, including medical image classification [13].

The following are the primary contributions of our suggested research: 1. Introduced DeepCervixNet, a novel automated approach for cervical cancer detection using enhanced ResNet101 and DenseNet169 architectures integrated with SE blocks. 2. Applied ensemble learning to combine features from both architectures, improving classification performance. 3. Utilized advanced image enhancement techniques, including Gaussian smoothing and median filtering, on the Harlev dataset.

2 Related Work

Deep learning (DL) has already surpassed human experts in modelling highly intricate connections between inputs and outputs, especially in cases where the characteristics are not easily intelligible to humans. P. Wang et al. [14] compressed a transfer learning-based Pap stain image classification model using a pruning strategy, eliminating a subset of neural network connections to reduce complexity and size without significant loss in efficacy. However, training their method was computationally intensive as the pruning was iteratively learned for a set number of epochs. T. Zhang et al. [15], demonstrated that a pre-trained DenseNet, specifically DenseNet-169, outperformed DenseNet-121 in classifying lesion levels in cervical images, suggesting a beneficial correlation be-

tween network depth and sensitivity. Furthermore, SVM classifiers trained with unique features outperformed DenseNet-based classifiers, analyzing 600 images in less than a minute. W. Mousser et al. [16] proposed an Incremental Deep Tree (IDT) structure designed exclusively for biological image classification. The IDT framework combines deep learning techniques with a tree-like structure, allowing for progressive learning by introducing new classes without retraining the entire model. Notably, it retains acceptable accuracy for previously learned lessons while adapting to new ones. The study tests the IDT framework on a variety of biological imaging datasets, including MNIST, BreakHis, LBC (Liquid-Based Cytology), and SIPaKMeD. The IDT framework shows encouraging results, with accuracies of 87%, 92%, 98%, and 93%, respectively. R. Maurya et al. [17] proposed an ensemble approach to classifying cervical cell Pap smear images that takes advantage of both Vision Transformers (ViT) and Convolutional Neural Networks (CNNs). ViT's low inductive bias and competitive performance make it an effective instrument. However, fine-tuning big ViT models is computationally intensive. As an alternative, the authors investigate a transfer learning strategy that employs pre-trained CNN features classified with a Long Short-Term Memory (LSTM) network, a resource-efficient method.

M. Karri et al. [18] proposed a three-phase methodology for cervical cancer detection, involving segmentation, nucleus localization, and classification. Their approach combines a ConvNet model, enhanced HDDT-based segmentation, OSS-SS for information extraction, retraining of ResNet-50 via transfer learning, feature downsampling with t-SNE, and classification using multi-class WKELM with SMCCA. The model outperforms existing approaches by achieving 99.12% accuracy, 99.45% specificity, and 99.25% sensitivity on SIPaKMeD. P. V. Mulmule et al. [19] developed an AI-assisted tool for cervical dysplasia detection using a pixel-based segmentation-to-classification approach. The two-step process involves cell segmentation, utilizing a novel filter-to-feature map approach, and cell classification, considering 163 features. Artificial neural network (ANN), support vector machine (SVM), and random forest (RF) classifiers were employed, with RF showing the highest accuracy (99.07%) in cell segmentation. Overall, ANN achieved 97.5% accuracy in identifying cervical dysplasia, outperforming SVM and RF. K. Ashfaque et al. [20] introduced Cervix Visionator ELM, an innovative automated method for detecting cervical cancer in Pap smear images. Their approach integrates the self-attention mechanism with EfficientNet, an advanced Convolutional Neural Network (CNN) architecture, and Vision Transformer (ViT) models to extract deep-learned features from Pap smear images. These features are then classified using an Extreme Learning Machine (ELM)-based classifier. The model's performance was evaluated using the SIPaKMeD open dataset, achieving an accuracy of 98.89%, a precision of 99.42%, a recall of 97.87%, and an F-measure of 98.76% in cervical cell classification.

3 Materials and Methods

This section offers a comprehensive overview of the dataset, including its pre-processing, augmentation, and image enhancement steps, along with a description of the proposed model.

3.1 Dataset Description

In this study, we utilized the publicly available Harlev dataset [21], consisting of a total of 917 Pap smear images categorized into seven distinct classes representing different types of cervical cells, collected at Herlev University Medical Center using a microscope and digital camera, as shown in Figure 1. The images have a resolution of 0.201 meters per pixel. Standard Pap staining and smear procedures were applied to the samples. The dataset, with ground truth for segmentation and classification, contains single cervical nuclei cells. Figure 2 depicts seven distinct diagnostic categories are presented based on assessments by doctors and two cytotechnicians, enhancing diagnostic precision. Notably, the nuclei of abnormal cells are more prominent than those of healthy cells in this dataset. As we can see in Figure 2 there is a class imbalance between normal and abnormal images.

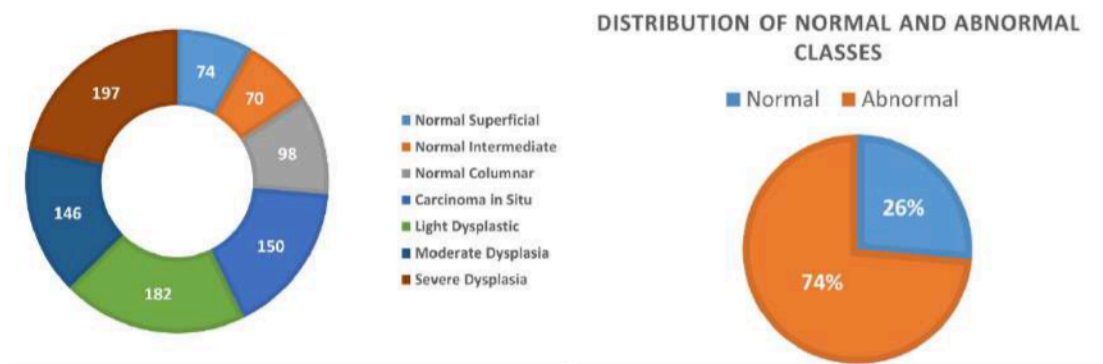**Figure 1.** Images sample of Harlev dataset according to the class

Figure 2. Distribution of Harlev dataset and percentage of normal and abnormal cells

3.2 Data Preprocessing

In the data pre-processing phase, we address the variability in image formats and dimensions within the Harlev dataset. The original images, initially in BMP format and varying in dimensions, are standardized to a uniform size of 224x224 pixels and converted into the widely used JPG format. This standardization facilitates consistent input for our model, ensuring optimal performance. Moreover, to enhance image quality and enable the model to better comprehend the visual information, we apply Gaussian smoothing and Bayesian filtering. Gaussian smoothing helps reduce noise and fine-tune image details, while Bayesian filtering further refines the features present in the images. These pre-processing steps collectively contribute to improved model interpretability and accuracy in subsequent classification tasks. The dataset split into training, validation, and testing parts with the ratio of 70%, 15%, and 15% respectively.

3.3 Proposed Methodology

The proposed framework for classifying cervical cancer cells is a structured approach involving several stages: Pre-processing, Classification using SE-Net architectures, and Ensemble Learning. The overall proposed framework is illustrated in Figure 3.

3.3.1 Sequence and Excitation Network

Sequence and Excitation (SE) blocks consist of two main operations: squeeze and excitation. The squeeze operation reduces spatial dimensions by applying global average pooling and compressing spatial information into channel-wise descriptors. The excitation operation models channel-wise dependencies using a small fully connected neural network. It learns the importance of each channel and generates a channel-wise attention vector through sigmoid or SoftMax activation. Algorithm 1 and Table 1 provide a detailed view of the phases and the

mathematical symbols used in SeNet. Understanding these phases and symbols is essential for comprehending how SeNet functions and appreciating its role in feature extraction and identification within the proposed framework. SeNet's ability to selectively highlight and assign weights to significant features is a crucial aspect of its contribution to the accurate identification of cervical cancer nests in medical images.

Algorithm 1. The fundamental process of SeNet Algorithm

1: **Input:** Feature graph of size $C \times W \times H$, u
2: **Output:** Feature graph of size $C \times W \times H$, x^*
3: $z_c \leftarrow F_{sq}(u_c) = \frac{1}{HW} \sum_{i=1}^{H} \sum_{j=1}^{W} u_c(i,j)$
4: $s \leftarrow F_{ex}(z, W) = \sigma[g(z, W)] = \sigma[W_2 \delta(W_1 z)]$
5: $x_c \leftarrow F_{scale}(u_c, s) = s_c u_c$
6: **return** x_c

Table 1. The description of the mathematical symbols used in SeNet Algorithm

Math Symbols	Description
u_c	Feature Channel
F_{sq}	Squeeze operation
F_{ex}	Excitation operation
z	Output of squeeze operation
W	Weight of channel
δ	Activation function
σ	Normalized function
F_{scale}	The definition of reweight operation
s_c	Output matrix channel for excitation operation

In the SE block, the input feature map u is first passed through the squeeze operation F_{sq}, which performs global average pooling to generate the output z. The excitation operation F_{ex} then takes z and applies a small fully connected neural network to produce the attention vector s. This vector s is used in the reweight operation F_{scale} to scale the input feature map u channel-wise, producing the final output x.
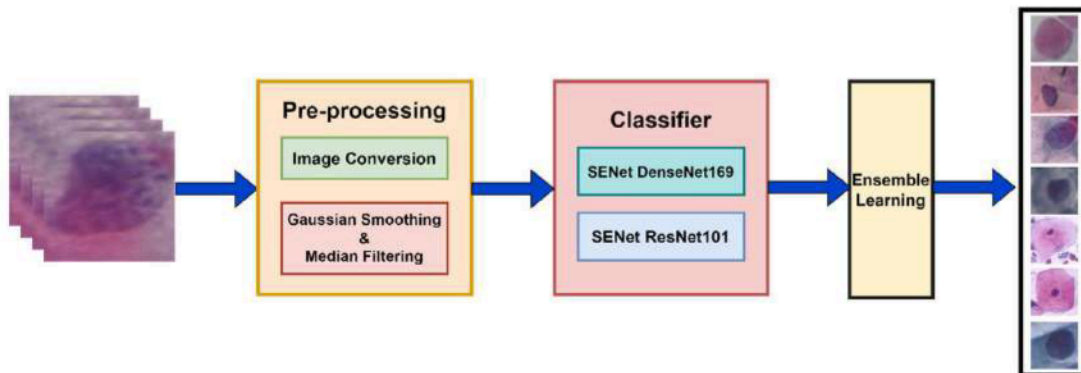

Figure 3. Block diagram of our proposed DeepCervixNet model

3.3.2 ResNet Architecture with SE Blocks

ResNet (Residual Network) is a deep convolutional neural network architecture known for its success in image classification tasks. We enhance the ResNet101 architecture by integrating SE (squeeze-and-excitation) blocks [22] to improve the detection of cervical cancer in Pap smear images. By integrating SE blocks into ResNet101, DeepCervixNet captures channel-wise dependencies and recalibrates feature maps. This enables the network to effectively capture discriminative features for cervical cancer detection, improving classification performance in Pap smear images. Figure 4 shows the architecture of ResNet integrating with SeNet blocks.
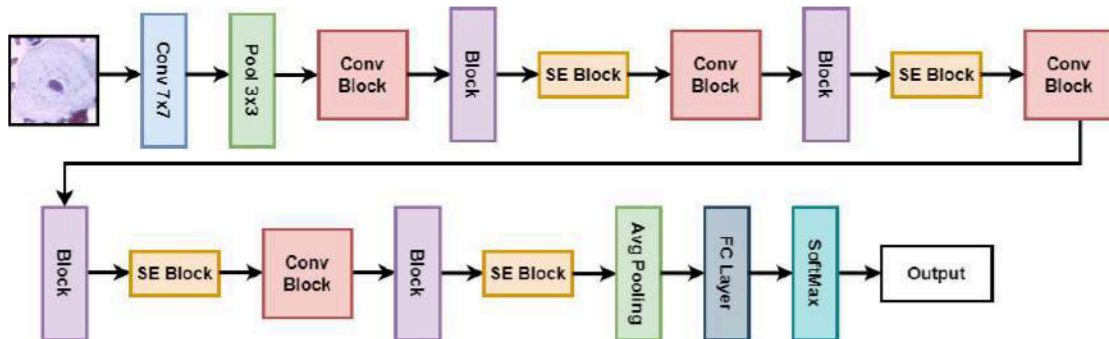

Figure 4. Integration of SE blocks with ResNet101

3.3.3 DenseNet Architecture with SE Blocks

DenseNet (Densely Connected Convolutional Network) is an architecture for deep convolutional neural networks that stresses dense connections between layers. It overcomes the shortcomings of traditional systems, in which information flow might degrade as it passes through numerous levels. DenseNet solves this problem by connecting all layers directly, allowing feature reuse and increasing gradient flow. We improve the DenseNet169 design with SE (squeeze-and-excite) blocks in our methodology. DenseNet169 is a DenseNet variation that consists of 169 layers. DeepCervixNet improves the architecture's discriminative capacity for cervical cancer diagnosis by inserting SE blocks into DenseNet169. This integration takes advantage of the strengths of DenseNet's dense connectivity patterns and the SE blocks' adaptive recalibration, leading to improved classification performance in Pap smear images. Figure 5 shows the architecture of DenseNet integrating with SeNet blocks.

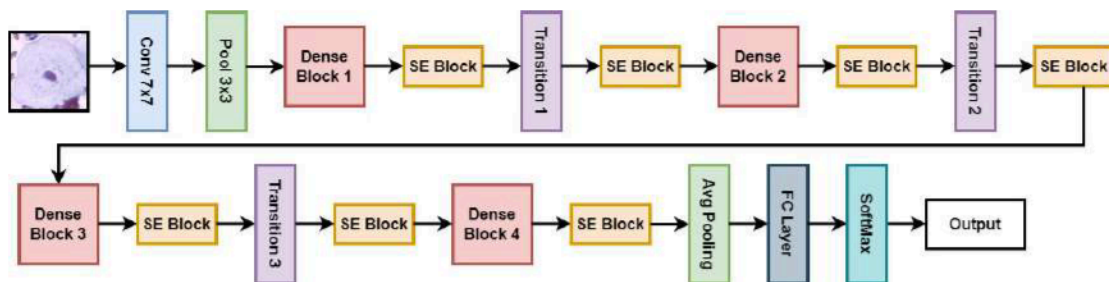**Figure 5.** Integration of SE blocks with DenseNet169

3.3.4 Ensemble Learning

Ensemble learning is employed to leverage the complementary strengths of multiple models. DeepCervixNet utilizes an ensemble of ResNet101 with SE blocks and DenseNet169 with SE blocks. The individual models are trained independently and combined through weighted averaging to obtain the final prediction. Ensemble learning helps

to improve the robustness and generalization performance of DeepCervixNet.

4 Experimental Results Discussion

The experimental hardware platform comprises an Intel Iris Xe with 8GB memory, and a 12th Generation 2.3 GHz Intel Core i7-12700H processor with 16GB RAM. The software platform utilizes Windows 11 (64-bit) as the operating system and Google Colab Pro [23] for training—an efficient cloud-based environment leveraging a GPU for machine learning and deep learning tasks. The experiment employs the PyTorch library, with Python version 3.10.12 facilitating a seamless development environment.

4.1 Experimental Process

In our experiment, a two-phase training and evaluation technique was employed for DeepCervixNet. Initially, ResNet101 and DenseNet169 models, both enhanced with SE blocks for refined feature representation, underwent distinct training periods. Figure 6(a) and Figure 6(b) illustrate the training performance of SeNet ResNet, and SeNet DenseNet169 respectively. The SeNet ResNet and SeNet DenseNet graphs both depict the training accuracy and loss over 200 epochs, highlighting their learning processes and convergence behavior. For SeNet ResNet, training accuracy starts at 0.2 and rises to 1.4, while training loss decreases from 1.2 to near zero, indicating effective learning. Slight fluctuations in accuracy after epoch 175 and the loss reaching zero around epoch 150 suggest learning rate adjustments. Similarly, SeNet DenseNet shows a rapid decrease in training loss and a quick increase in accuracy, with the loss stabilizing at a low value and occasional spikes around epochs 75 and 175. The high and stable accuracy after the initial rise confirms robust model performance. Both graphs underscore the models' effectiveness in achieving high accuracy and low loss, crucial for their application in the research study.
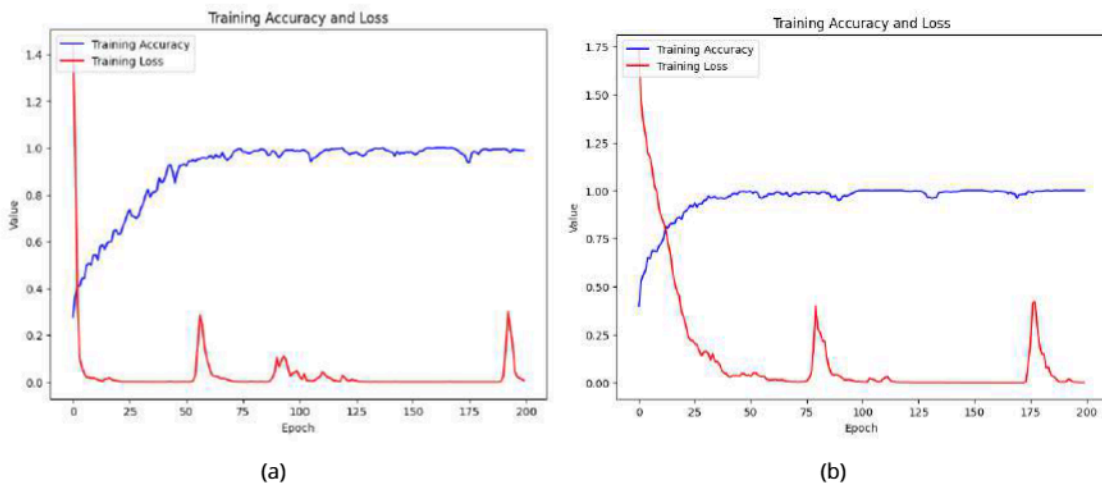

(a)   (b)

Figure 6. Train and loss graph of (a) SeNet ResNet101 (b) SeNet DenseNet169

Figure 7(a) and Figure 7(b) display the confusion matrix for both SeNet ResNet101 and SeNet DenseNet169, facilitating a comprehensive analysis of their performance. This matrix provides a holistic view of classification accuracy and misclassifications, offering valuable insights into detection proficiency. Moreover, the outputs of SeNet ResNet101 and SeNet DenseNet169 were strategically amalgamated during the ensemble learning phase. This collaborative approach significantly enhanced the final output, elevating DeepCervixNet's overall efficacy in cervical cancer diagnosis. The hyperparameters for our base models are detailed in Table 2.
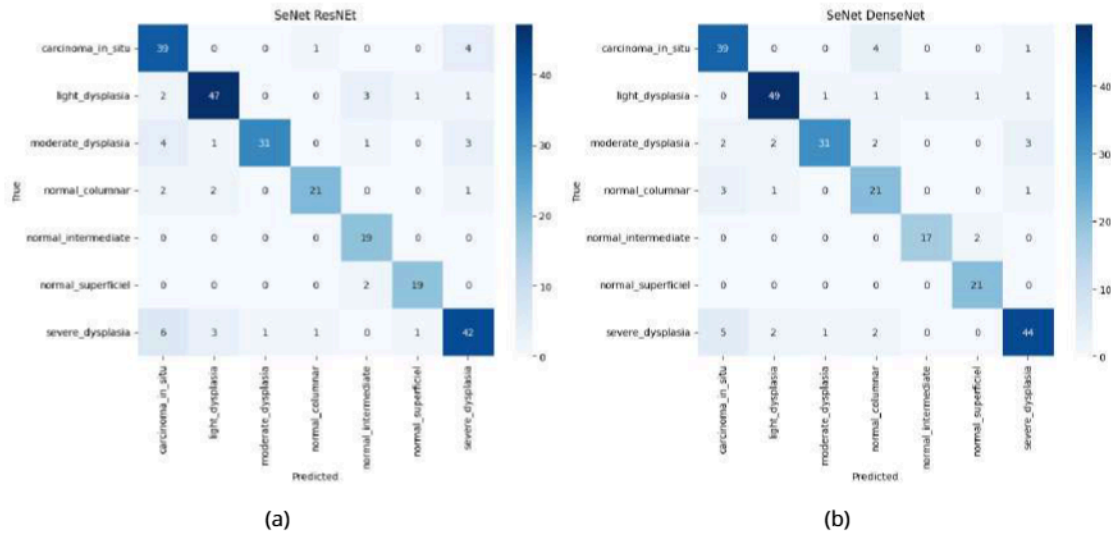
Figure 7. Confusion Matrix (a) SeNet ResNet101 and (b) SeNet DenseNet169

Table 2. Hyperparameters used in our base models

Hyperparameters	SeNet ResNet101	SeNet DenseNet169
Batch size	32	64
Optimizer	Adam	Adam
Learning rate	0.001	0.0001
Epochs	200	200
Input size	224x224x3	224x224x3

4.2 Evaluation Metrics

The confusion matrix depicts real data values vs expected data values. It is the structure of tables that is utilized to assess the efficacy of our deep learning classification model.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{3}$$

$$\text{F1 Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

Where TN = True Negative, TP = True Positive, FN = False Negative, and FP = False Positive.

4.3 Analytical Discussion

The overall accuracy of SeNet DenseNet is 95.60% and the accuracy of SeNet ResNet is 99.12, after using weighted averaging method of ensemble learning and assign 0.6 and 0.4 weights to SeNet ResNet101 and SeNet DenseNet169 respectively. We chose these weights accordingly to the model's individual performance on the dataset. Table 3 presents the base models and final model training accuracy comparison. The SE-ResNet has superior performance over SE-DenseNet in our study can be attributed to several factors. SE-ResNet's architecture leverages residual connections that mitigate the vanishing gradient problem, facilitating the training of deeper networks, which is particularly beneficial for complex image patterns in cervical cancer detection. While SE-DenseNet also uses skip connections, its densely connected layers might introduce redundant features, leading to overfitting with limited training data. Additionally, SE-ResNet's integration of Squeeze-and-Excitation (SE) blocks enhances the network's ability to focus on the most informative features by recalibrating channel-wise feature responses, whereas SE-DenseNet might suffer from excessive parameterization that dilutes the impact of SE blocks. The ensemble method employed in this study significantly boosts performance by combining the strengths of SE-ResNet and SE-DenseNet, reducing variance and bias, and ensuring robust predictions through diverse perspectives on the input data. This approach also mitigates overfitting, as errors from individual models are averaged out, resulting in a more generalized performance, which is reflected in the higher accuracy and F1-scores achieved by the ensemble model. This robust and reliable ensemble method, combined with the architecture advantages of SE-ResNet, offers a powerful tool for early detection and treatment of cervical cancer, underscoring the importance of architecture choice and feature recalibration in medical image analysis. Table 4 presents a comparison of the validation and testing results for the proposed DeepCervixNet model. The validation phase shows an accuracy of 98.18%, precision of 97.57%, recall of 97.22%, and an F1-score of 97.29%. During testing, the model achieves slightly higher metrics, with an accuracy of 98.36%, precision of 97.83%, recall of 97.42%, and an F1-score of 97.57%. These results indicate the model's consistent and robust performance across both validation and testing phases, underscoring its effectiveness in accurately detecting cervical cancer in Pap smear images.

Table 3. Comparison of results of models used in this study

Models	Accuracy
SeNet DenseNet169	95.60%
SeNet ResNet101	99.12%
Proposed Model	99.89%

Table 4. Validation and testing results of proposed model DeepCervixNet

	Validation	Testing
Accuracy	98.18%	98.36%
Precision	97.57%	97.83%
Recall	97.22%	97.42%
F1-score	97.29%	97.57%

Table 5 presents a comparative performance analysis of the proposed framework with other state-of-the-art methods. The proposed method, which employs ensemble learning integrating SE blocks with CNN models,

achieves an impressive accuracy of 99.89% on the Harlev Dataset. This is significantly higher than the accuracies reported for other methods on the same dataset, such as 64.00% by MH Bhuiyan et al. using transfer learning-based CNN and 75.06% by T Albuquerque et al. using ordinal losses. For the SIPaKMeD dataset, the highest accuracy recorded is 98.89% by K. Ashfaque et al. using an Extreme Machine Learning Classifier. Figure 8(a) and Figure 8(b) show the bar graph of actual and classified images of each class for both base models.

Table 5. Performance analysis of the proposed framework with other state-of-the-art methods

| Method	Dataset	Approach	Accuracy
M. H. Bhuiyan et al. [24]	Harlev Dataset	Transfer Learning based CNN	64.00%
T. Albuquerque et al. [25]	Harlev Dataset	Ordinal Losses	75.06%
W. Mousser et al. [16]	SIPaKMeD Dataset	Incremental Learning Method	93.00%
R. Maurya et al. [17]	SIPaKMeD Dataset	CNN-LSTM	95.80%
K. Ashfaque et al. [20]	SIPaKMeD Dataset	Extreme Machine Learning Classifier	98.89%
K. P. Win et al. [26]	SIPaKMeD Dataset	SVM, KNN, boosted trees, bagged trees, and major voting	94.09%
B. S. Deo et al. [27]	SIPaKMeD Dataset	Cross attention and latent transformer	93.70%
Proposed Model	Harlev Dataset	SeNet with CNN and Ensemble Learning	99.89%
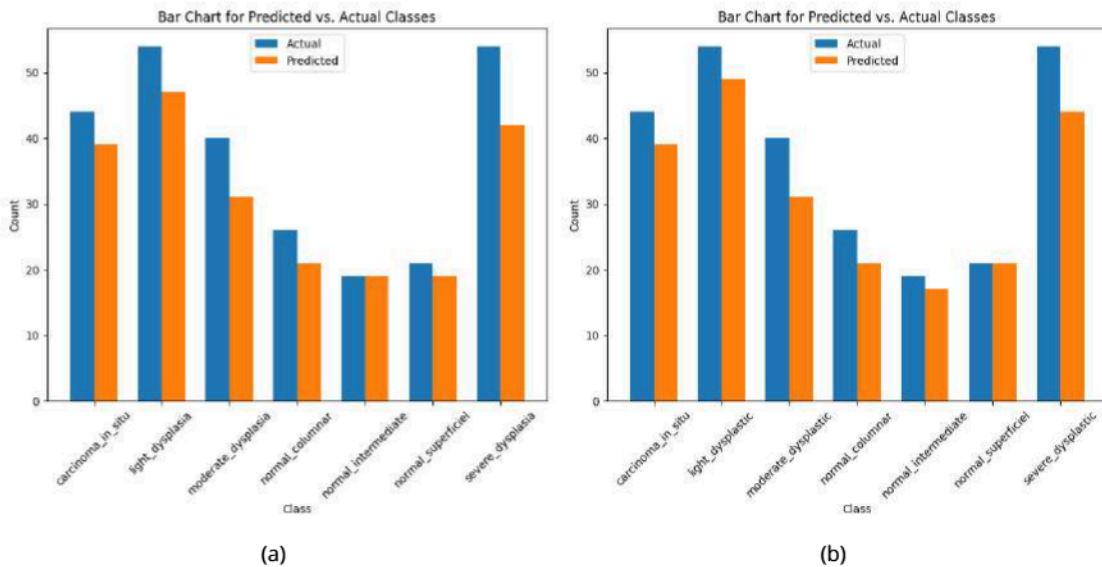
Figure 8. Bar graph for each actual and predicted class (a) SeNet ResNet101 and (b) SeNet DenseNet169

4.4 Ablation Study

We performed an ablation study to obtain more insight into the significance that each ResNet and DenseNet model plays within our suggested framework for classifying cervical cancer. Table 6 summarizes the performance metrics, including accuracy, precision, recall, and F1-score, for each model. The results demonstrate that the ResNet-101 model achieves the highest performance across all metrics, with an accuracy of 93.37%, a precision of 93.30%, a recall of 93.19%, and an F1-score of 93.34%. This indicates a superior ability of ResNet-101 to accurately identify and classify features within the cervical cancer dataset. ResNet-50 and ResNet-152 also performed well, with ResNet-50 achieving an accuracy of 90.83% and ResNet-152 achieving 92.95%. Both models

show a balance in precision, recall, and F1-score, suggesting their robustness in feature extraction and classification tasks. However, they fall short of the performance achieved by ResNet-101. On the other hand, DenseNet variants showed relatively lower performance compared to ResNet models. DenseNet-121 achieved an accuracy of 88.14%. DenseNet-161 and DenseNet-169 showed accuracies of 90.20% and 91.72%, respectively. The lower performance of DenseNet models could be attributed to their increased depth and complexity, which might lead to overfitting or difficulty in optimizing parameters for this specific dataset. Based on these results, we chose ResNet-101 and DenseNet-169 for their good performance and integrated SeNet with these architectures to enhance the overall accuracy of our model. The integration of SeNet allows for better feature extraction and classification by selectively highlighting and assigning weights to significant features.

Table 6. Performance analysis of proposed model DeepCervixNet with other ResNet and DenseNet architectures (in %)

Model	Accuracy	Precision	Recall	F1-score
ResNet-50	90.83	91.21	91.04	90.75
ResNet-101	93.37	93.30	93.19	93.34
ResNet-152	92.95	93.23	92.87	92.97
DenseNet-121	88.14	87.95	87.85	87.90
DenseNet-161	90.20	90.12	89.95	89.93
DenseNet-169	91.72	91.73	91.68	91.65

5 Conclusion and Future Work

The significant impact of cervical cancer on female mortality emphasizes the critical need for effective screening measures. This research introduces DeepCervixNet, an enhanced model combining ResNet101 and DenseNet169 with SE blocks and utilizing Ensemble learning for cervical cancer detection. Tested on the Harlev dataset with Gaussian smoothing, the model achieves a robust 99.89% accuracy, demonstrating superiority over existing methods. Validation metrics further confirm its superiority, with DeepCervixNet outperforming existing methods by achieving a precision of 97.57%, recall of 97.22%, and F1 score of 97.29%. However, several limitations should be noted. The model's performance has been validated only on the Harlev dataset, which may limit its generalizability to other datasets. Additionally, the model's real-time application has not been tested, and the interpretability of the results needs further improvement. Future work includes integrating multi-modal data, conducting extensive clinical validations, implementing the model in real-time settings, and enhancing interpretability. Despite its current limitations, DeepCervixNet marks a significant advancement in automating and improving cervical cancer diagnostics, offering a promising avenue for future developments in early detection and treatment.

Author Contributions

Dhani Bux Talpur: Conceptualization, Methodology, Software, Data curation. **Ali Raza Rang**: Visualization, Investigation. **Ashfaque Khowaja**: Software, Validation, Writing- Original draft preparation. **Asadullah Shah**: Supervision, Writing- Reviewing and Editing.

Compliance with Ethical Standards

It is declare that all authors don't have any conflict of interest. It is also declare that this article does not contain any studies with human participants or animals performed by any of the authors. Furthermore, informed consent was obtained from all individual participants included in the study.

References

[1] T. Šarenac and M. Mikov, "Cervical cancer, different treatments and importance of bile acids as therapeutic agents in this disease," *Frontiers in Pharmacology*, vol. 10, p. 484, 2019.

[2] K. Hemalatha and V. Vetriselvi, "Deep learning based classification of cervical cancer using transfer learning," in *2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC)*, pp. 1–6, 2022.

[3] X. Li *et al.*, "Detection of cervical cancer cells in whole slide images using deformable and global context aware faster rcnn-fpn," *Current Oncology*, vol. 28, no. 5, pp. 3585–3601, 2021.

[4] E. Hussain *et al.*, "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*, vol. 65, p. 101347, 2020.

[5] M. M. Rahaman *et al.*, "Deepcervix: A deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques," *Computers in Biology and Medicine*, vol. 136, p. 104649, 2021.

[6] M. M. Rahaman *et al.*, "A survey for cervical cytopathology image analysis using deep learning," *IEEE Access*, vol. 8, pp. 61687–61710, 2020.

[7] M. Alsalatie *et al.*, "Analysis of cytology pap smear images based on ensemble deep learning approach," *Diagnostics*, vol. 12, no. 11, p. 2756, 2022.

[8] A. Demir, F. Yilmaz, and O. Kose, "Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3," in *2019 medical technologies congress (TIPTEKNO)*, pp. 1–4, 2019.

[9] D. Lv *et al.*, "A cascade network for detecting covid-19 using chest x-rays," *arXiv preprint arXiv:2005.01468*, 2020.

[10] P. Ghosal *et al.*, "Brain tumor classification using resnet-101 based squeeze and excitation deep neural network," in *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*, pp. 1–6, 2019.

[11] G. Ye *et al.*, "Multitask classification of breast cancer pathological images using se-densenet," in *2019 Eleventh International Conference on Advanced Computational Intelligence (ICACI)*, pp. 1–6, 2019.

[12] R. Polikar, *Ensemble learning*, pp. 1–34. 2012.

[13] R. S. Akinbo and O. A. Daramola, *Ensemble machine learning algorithms for prediction and classification of medical images*, pp. 1–15. 2021.

[14] P. Wang *et al.*, "Adaptive pruning of transfer learned deep convolutional neural network for classification of cervical pap smear images," *IEEE Access*, vol. 8, pp. 50674–50683, 2020.

[15] T. Zhang *et al.*, "Cervical precancerous lesions classification using pre-trained densely connected convolutional networks with colposcopy images," *Biomedical Signal Processing and Control*, vol. 55, p. 101566, 2020.

[16] W. Mousser *et al.*, "Idt: An incremental deep tree framework for biological image classification," *Artificial Intelligence in Medicine*, vol. 134, p. 102392, 2022.

[17] R. Maurya, N. N. Pandey, and M. K. Dutta, "Visioncervix: Papanicolaou cervical smears classification using novel cnn-vision ensemble approach," *Biomedical Signal Processing and Control*, vol. 79, p. 104156, 2023.

[18] M. Karri *et al.*, "Multi-class nucleus detection and classification using deep convolutional neural network with enhanced high dimensional dissimilarity translation model on cervical cells," *Biocybernetics and Biomedical Engineering*, vol. 42, no. 3, pp. 797–814, 2022.

[19] P. V. Mulmule, R. D. Kanphade, and D. M. Dhane, "Artificial intelligence-assisted cervical dysplasia detection using papanicolaou smear images," *The Visual Computer*, vol. 39, no. 6, pp. 2381–2392, 2023.

[20] K. Ashfaque, Z. BeiJi, and X. Kui, "Cervix visionator elm: A novel approach to early detection of cervical cancer," in *IEEE*, pp. 1–4, 2024.

[21] J. Jantzen *et al.*, "Pap-smear benchmark data for pattern classification," in *Nature inspired smart information systems (NiSIS 2005)*, pp. 1–9, 2005.

[22] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 7132–7141, 2018.

[23] E. Bisong, *Building machine learning and deep learning models on Google cloud platform*. Springer, 2019.

[24] M. H. Bhuiyan and M. A. R. Arju, "Classification of pap smear image of cervix cell using machine learning techniques and transfer learning-based convolutional neural network architecture and scrutinizing their performances," in *ICT Analysis and Applications: Proceedings of ICT4SD 2022*, pp. 757–769, 2022.

[25] T. Albuquerque, R. Cruz, and J. S. Cardoso, "Ordinal losses for classification of cervical cancer risk," *PeerJ Computer Science*, vol. 7, p. e457, 2021.

[26] K. P. Win *et al.*, "Computer-assisted screening for cervical cancer using digital image processing of pap smear images," *Applied Sciences*, vol. 10, no. 5, p. 1800, 2020.

[27] B. S. Deo *et al.*, "Cerviformer: A pap smear-based cervical cancer classification method using cross-attention and latent transformer," *International Journal of Imaging Systems and Technology*, vol. 34, no. 2, p. e23043, 2024.