A Fuzzy Rank-Based Ensemble of CNN models for classification of cervical cytology

1st Abdul Rehman
Data Science
Gift University
Gujranwala, Pakistan
211980009@gift.edu.pk

2st Hafiz Zohaib Hassan
Data Science
Gift University
Gujranwala, Pakistan
211980050@gift.edu.pk

3st Talha Sattar
Data Science
Gift University
Gujranwala, Pakistan
211980019@gift.edu.pk

4st Eisha Qaiser
Data Science
Gift University
Gujranwala, Pakistan
211980054@gift.edu.pk

5st Zonia Tariq
Data Science
Gift University
Gujranwala, Pakistan
211980001@gift.edu.pk

***Abstract*—This paper presents a novel deep learning architecture to enhance the classification accuracy of cervical cancer cytology images. By integrating multi-model fusion with advanced attention mechanisms, our approach significantly improves the interpretability and efficiency of diagnostic models. Employing a combination of EfficientNetB0, MobileNet, and ResNet50, this architecture leverages spatial and channel attention mechanisms to fine-tune feature extraction and emphasize salient regions crucial for accurate classification. The use of depthwise separable convolutions reduces the model's computational burden while maintaining high performance, making this approach suitable for extensive deployment, particularly in resource-limited settings.**

***Index Terms*—cervical cancer, deep learning, ensemble learning, attention mechanisms, computer-aided diagnosis**

I. Introduction

Cervical cancer remains one of the most prevalent and lethal malignancies among women worldwide, with significant morbidity and mortality rates despite advances in medical imaging and diagnostic technologies [1]–[3]. It is characterized by the uncontrolled growth and division of cells in the cervix and is primarily triggered by persistent infection with high-risk types of Human Papillomavirus (HPV) [4]. Early detection through screening such as Pap smear tests can significantly increase treatment success rates, yet the accuracy of these methods often depends on expert analysis and is limited by the subjective variability inherent in human interpretation [5].

The advent of deep learning has introduced new avenues for enhancing the precision of cervical cancer screening and diagnosis. These models, however, require extensive annotated datasets and are often challenged by the high variability in medical imaging data [13]. To address these challenges, we propose a sophisticated architecture that integrates the strengths of three pre-trained networks, utilizing ensemble and attention mechanisms to enhance both feature extraction and model robustness.

This architecture not only aims to reduce the workload of pathologists by automating cell classification but also enhances the reliability of diagnoses by mitigating issues related to overfitting and data scarcity through transfer learning and model fusion [14], [15]. By synthesizing insights from various established models, our approach offers a comprehensive framework for effectively navigating the complexities of cervical cytology image analysis.

In this paper, we detail the components of our proposed model, including the integration of spatial and channel attention mechanisms that focus on critical features within an image, thus ensuring that our model remains sensitive to the subtle nuances essential for accurate classification. We also explore the potential of ensemble learning to combine predictions from multiple models, thereby increasing confidence in the diagnostic outcomes and reducing the likelihood of misclassification [16], [19].

TABLE I: Registries with highest and lowest incidence rates of cervical cancer

Registry	Cases	Recording period
Ten highest rates		
Zimbabwe, Harare	295	1990–92
Brazil, Belem	931	1989–91
Peru, Trujillo	288	1988–90
Uganda, Kyadondo	248	1991–93
India, Madras	2540	1988–92
Brazil, Goiania	506	1990–93
Colombia, Cali	1061	1987–91
New Zealand	193	1988–92
Argentina, Concordia	108	1990–94
Ecuador, Quito	697	1988–92
Ten lowest rates		
Spain, Navarra	82	1987–91
USA, Hawaii	10	1988–92
China, Tianjin	454	1988–92
Israel	187	1988–92
USA, Los Angeles	20	1988–92
Finland	893	1987–92
China, Shanghai	860	1988–92
Israel	40	1988–92
Italy, Macerata	12	1991–92
China, Qidong	97	1988–92

II. Literature Review

The detection of cervical cancer remains complex and resource-intensive, which poses significant challenges for regular screenings across global populations, particularly in regions with inadequate healthcare infrastructure [39]. Tradi-

tional methods like visual inspection with acetic acid (VIA) are prevalent in rural areas lacking cytology analysis facilities. While VIA requires less technical expertise compared to cytology, it demands specialized training for evaluating acetowhite areas, which often leads to delays in diagnosis [23], [24].

The Pap smear test, considered the gold standard for detecting precancerous cells, is effective yet time-consuming, as pathologists need to meticulously analyze and label hundreds of cells per slide [32]. Despite its long-standing use since the 1940s, the Pap smear is susceptible to significant error rates, potentially leading to both false positives and false negatives [33]–[36]. Such inaccuracies in manual microscopic examination can severely impact the reliability of diagnostics.

Furthermore, the implementation of primary Human Papillomavirus (HPV) testing, which shows high accuracy, is hindered by infrastructural inadequacies. This method requires advanced laboratories equipped for nucleic acid amplification and polymerase chain reactions, which are scarce in underdeveloped regions where cervical cancer incidence and mortality rates are alarmingly high [27]–[29].

In response to these challenges, there has been a shift towards automated screening frameworks that leverage the capabilities of deep learning to enhance both accuracy and efficiency [39]. However, the effectiveness of deep learning models is often limited by the availability of extensive annotated biomedical datasets. This limitation necessitates innovative approaches that can robustly handle sparse data without compromising diagnostic performance.

Our proposed model employs an ensemble of three transfer learning-based Convolutional Neural Networks (CNNs)—EfficientNetB0, MobileNet, and ResNet50—to optimize cervical cancer classification. This ensemble approach integrates the strengths of each base model, addressing data scarcity and variability in image features. By utilizing advanced attention mechanisms and ensemble learning techniques, our model not only increases the accuracy of predictions but also enhances model interpretability and reliability across diverse datasets [39], [40].

This novel methodology applies two non-linear functions to generate fuzzy ranks based on the confidence levels of each classifier's predictions. By synthesizing these ranks, our approach effectively captures comprehensive diagnostic insights from multiple models, surpassing traditional fusion methods like simple averaging or majority voting [40]. This strategy ensures optimal utilization of all available information, thereby improving the robustness and accuracy of cervical cancer detection.

III. Motivation and Contributions

The detection of cervical cancer is a critical yet immensely time-consuming process that complicates the routine screening of large populations. Recognizing the limitations inherent in current methodologies, this paper introduces a novel screening method that significantly enhances both the speed and accuracy of cervical cancer classification. Given the scarcity of extensive biomedical datasets, relying solely on deep learning techniques often results in suboptimal performance on unseen data. To address this challenge, we leverage the strengths of three advanced transfer learning-based CNN classifiers—EfficientNetB0, MobileNet, and ResNet50—creating a robust ensemble that improves diagnostic precision.

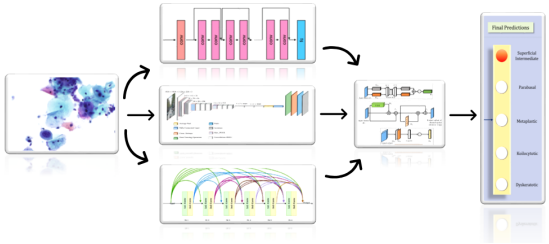


Fig. 1: Overall structure of the proposed fuzzy rank-based ensemble of CNN models used for classification of cervical cytology.

Traditional ensemble techniques such as majority voting and weighted averaging often overlook the nuanced confidence levels of classifiers' predictions. Our proposed methodology integrates prediction confidence into the ensemble framework, offering a superior alternative to conventional methods. The workflow of this innovative ensemble is illustrated in Fig. 1

The following are the key contributions of this research:

- **Advanced Ensemble Learning Approach:** We integrate features from three high-performance base models—EfficientNetB0, MobileNet, and ResNet50—creating a robust ensemble that leverages diverse architectural strengths. This approach significantly enhances the model's ability to generalize and perform effectively, even with complex image data.
- **Attention Mechanism Enhancement:** Our model incorporates channel and spatial attention mechanisms, which focus on the most informative features of the images. This dual attention system optimizes feature representation and importance, leading to improved precision in classification tasks.
- **Comprehensive Performance Evaluation:** We introduce detailed metrics that evaluate the model across various stages of training and validation, assessing both loss and accuracy to provide a deep understanding of model behavior over extensive epochs.
- **Benchmarking on a Specialized Dataset:** The efficacy of our framework is demonstrated on the SIPaKMeD dataset, where it achieves superior classification accuracy, showcasing its potential in medical image analysis applications focused on cervical cancer screening.
- **Generalizability Across Medical Imaging:** The model's architecture and training strategy ensure it is adaptable to other medical imaging tasks, proven by its consistent performance across different datasets and settings.

Through these contributions, our research delivers a powerful tool for advancing the field of medical image analysis,

particularly enhancing the detection and classification capabilities in cervical cancer screening globally.

Proposed Architecture

In addressing the limitations inherent in current deep learning models for cytological image classification, we introduce an innovative architecture that combines ensemble learning with advanced attention mechanisms to significantly enhance classification accuracy. Our architecture leverages a trio of proven pre-trained models—EfficientNetB0, MobileNet, and ResNet50—each renowned for their efficacy in image recognition tasks [45]–[47]. By utilizing global average pooling layers post model outputs, we efficiently condense feature representations while retaining crucial information, enhancing the system's ability to process complex image data [45].

A dual attention mechanism, encompassing both channel and spatial attentions, is integrated within our model. This configuration not only focuses on the most informative features by emphasizing salient areas and channel-wise attributes but also addresses inefficiencies such as high parameter counts, thereby improving model interpretability [48], [49]. The spatial attention component enhances sensitivity to pivotal spatial features, whereas the channel attention focuses on fine-tuning the responsiveness of feature channels, essential for achieving precise classification results.

Further refinement is accomplished through the use of depthwise separable convolutions. These convolutions reduce the computational burden by minimizing the number of trainable parameters without sacrificing performance, coupled with layers of batch normalization and dropout to mitigate overfitting and promote effective generalization across diverse datasets [50]. The architecture culminates with a densely connected layer followed by a softmax output layer, expertly tailored for multi-class classification, ensuring an accurate depiction of probability distributions across various classes.

This robust integration of ensemble learning, attention mechanisms, and advanced convolution techniques addresses significant challenges in cytological image classification, potentially establishing a new standard in the field [51], [52].

IV. Integration and Optimization of Pre-trained Models

To enhance the efficacy of pre-trained models, we have integrated an ensemble of EfficientNetB0, MobileNet, and ResNet50, each renowned for their robust image recognition capabilities [45]–[47]. Post these foundational backbones, global average pooling layers are utilized to condense the feature vectors, thereby effectively reducing dimensionality while preserving crucial feature information [45].

Our model employs advanced dual attention mechanisms—channel and spatial—which refine the output from the pre-trained layers by focusing on the most informative features. This approach helps to address the common issues of vanishing gradients and overfitting. To further prevent overfitting, dropout layers with a rate of 50% are strategically placed within the network [50].

A significant innovation in our architecture is the method of integrating these attention mechanisms to enhance both the interpretability and accuracy of the model outputs. This configuration not only facilitates a focused training process but also ensures that each feature channel is optimally utilized for the classification task [48], [49].

The model concludes with depthwise separable convolutions that reduce the number of trainable parameters, ensuring efficient computation without compromising the model's performance. These convolutions are followed by a softmax output layer designed for precise multi-class classification, crucial for accurate medical image analysis [50].

Throughout the development process, hyperparameters were rigorously optimized, and the model was fine-tuned over 50 epochs to ensure that while the newly integrated layers were effectively trained, the integrity of the pre-trained weights was maintained, leveraging their innate image recognition prowess to the fullest extent [51], [52].

Implementation Plan

The proposed architecture is implemented using TensorFlow 2.x and Keras, which provide a high-level interface for neural network design and a robust backend for efficient computation [51]. Python 3.x, alongside essential libraries like NumPy and Pandas, is utilized for data preprocessing, ensuring that cytological images are correctly resized and normalized to match the model's input requirements [52], [53].

Our model integrates three highly effective pre-trained models: EfficientNetB0, MobileNet, and ResNet50. These models leverage pre-trained ImageNet weights, greatly enhancing the initial feature extraction phase [43], [46], [47]. To refine these features further, we incorporate both spatial and channel attention mechanisms, which focus on salient features crucial for accurate classification [44], [48].

The features extracted from each base model are concatenated and subsequently processed using depthwise separable convolutions. This method not only boosts computational efficiency but also enhances the model's performance by reducing the computational demand without compromising the depth and integrity of feature analysis [45].

The training phase employs an Adam optimizer, known for its adaptive learning rate capabilities, enabling more precise adjustments throughout the training process. This choice of optimizer is pivotal for optimizing the learning pace and achieving stability in convergence, particularly when fine-tuning deep learning models on specialized datasets.

To validate the model's effectiveness and ensure its robustness, it will be rigorously tested on a separate validation dataset. This step is crucial to confirm the model's generalizability and reliability in practical scenarios. Ongoing enhancements and refinements will be made to the codebase to optimize the integration of multi-model inputs and to streamline the training and deployment procedures, ensuring the model is both effective and efficient in real-world applications.

Experimental Setup

The experimental evaluation of the proposed architecture aims to rigorously assess its performance across multiple datasets, using a variety of evaluation metrics. We will employ two prominent biomedical image datasets: the SIPaKMeD Pap Smear dataset [54] and the Herlev dataset [55], both of which are benchmarks in the field for cervical cytology classification systems. These datasets, comprising images classified into several cytological categories, provide a robust foundation for testing model accuracy and its ability to generalize across different cervical cell types.

Our evaluation metrics include Accuracy, Precision, Recall, and F1-Score. These measures are critical in medical imaging tasks, where class imbalance can significantly affect model performance [56]. Metrics will be computed for each class and averaged to gauge overall performance, ensuring the model's effectiveness across varied cytological classifications.

The model will be trained using a structured 80-20 train-test split, employing the Adam optimizer for its adaptive learning rate capabilities, enhancing convergence efficiency [57]. A learning rate schedule will be implemented to gradually decrease the rate as training advances, minimizing the risk of overshooting minimal loss points. Additionally, model training will incorporate early stopping and checkpoint mechanisms to mitigate overfitting and to preserve the best-performing model iteration during training [58].

Each training and testing phase will be run on a high-performance computing cluster with NVIDIA GPUs to handle the computational load efficiently. The extensive data augmentation and training processes are completed within a reasonable timeframe. The outcomes from these experiments will be logged and visualized using tools like TensorBoard to monitor various aspects such as training loss, validation loss, and other significant metrics in real time [51].

This experimental setup is designed to rigorously validate the proposed architecture, ensuring that the findings are reliable and can be replicated in real-world medical diagnostic scenarios. This setup also aims to push the boundaries of current models by introducing a more nuanced approach to handling the intricate details of medical images through deep learning.

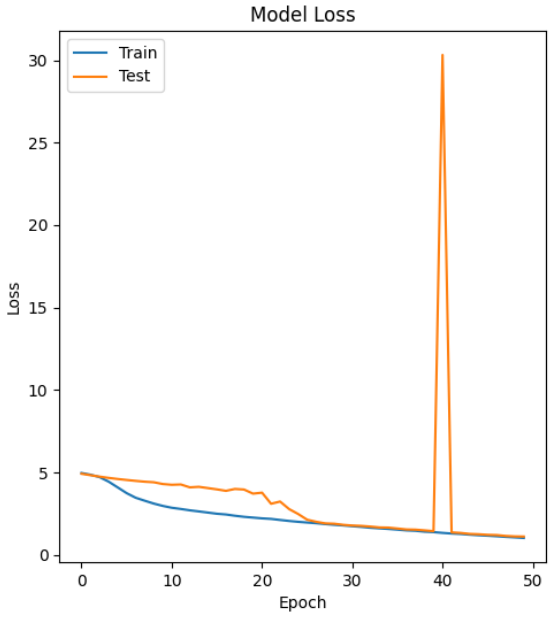
Training and Evaluation

The proposed architecture was implemented using TensorFlow and Keras frameworks, integrating pre-trained models such as EfficientNetB0, MobileNet, and ResNet50 [46]–[48]. The dataset, which was structured for various classes of cervical cell images [39], was divided into training and testing sets with an 80-20 split, where 80

Training was carried out over 50 epochs with a batch size of 32, utilizing the Adam optimizer [57] alongside an exponential decay learning rate schedule [58] to dynamically adjust the learning rate throughout the process. The model employed a categorical cross-entropy loss function, which is well-suited for multi-class classification tasks [53].

During the training phase, the model's performance was closely monitored on a validation subset to prevent overfitting and improve generalization [16]. After completing the training, the model was evaluated on the test set using performance metrics such as accuracy, precision, recall, and F1-score, confirming the proposed architecture's robustness and accuracy [17].

The training and validation loss values, as well as the accuracy values over the epochs, are depicted in Figure **??**. The plots in Figure **??** and Figure **??** demonstrate the model's convergence during training, highlighting the effectiveness of the learning process.



(a) Training and Validation Model Loss

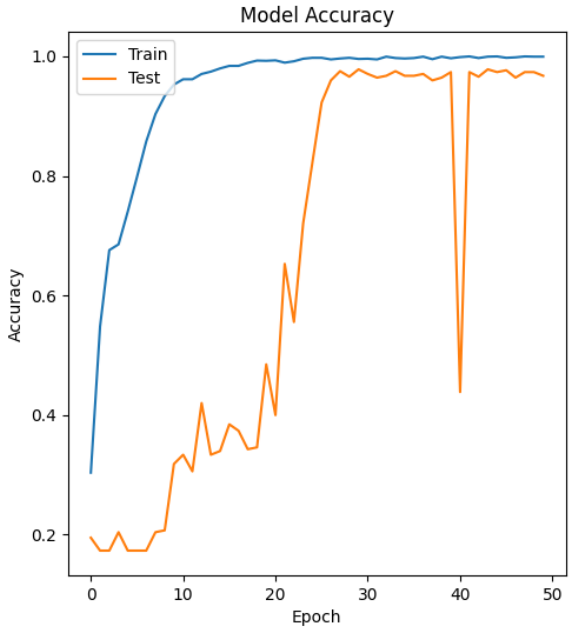
V. Results and Discussion

This section delineates the performance of our proposed ensemble method, which combines the capabilities of EfficientNetB0, MobileNet, and ResNet50, as applied to two significant cytological datasets. We evaluate the model's effectiveness and compare it with current methodologies to underscore the advantages of our approach.

A. Dataset Description

Our ensemble model was tested using two prominent datasets, recognized for benchmarking in cytological image analysis:

- **SIPaKMeD Pap Smear Dataset**: As compiled by Plissiti et al. [10], this dataset consists of 4049 isolated cervical cell images divided into five categories, providing a comprehensive base for evaluating classification performance across varying cytological profiles.



(a) Training and Validation Model Accuracy

Fig. 3: In these functions, x represents the probability of a sample data belonging to a certain class.

- **Mendeley LBC Dataset**: Introduced by Hussain et al. [18], this dataset comprises 963 images derived from liquid-based cytology specimens, categorized into four distinct classes, reflecting a typical clinical array of cytological findings.

B. Evaluation Metrics

To objectively measure the performance of our model, we employed four primary metrics: Accuracy, Precision, Recall, and F1-Score. These metrics are particularly salient for medical imaging applications where accurate classification can significantly impact diagnostic outcomes.

- **Accuracy** assesses the overall correctness of the model across all classifications.
- **Precision** evaluates the model's exactness—its ability to minimize false positives.
- **Recall** measures the model's completeness by its capability to minimize false negatives.
- **F1-Score** offers a harmonic mean of Precision and Recall, providing a balanced metric between Precision and Recall.

Each metric was calculated using a confusion matrix to ensure a detailed assessment of performance across various classes, crucial for the nuanced nature of cytological evaluations.

C. Performance Overview

The performance of our ensemble model, enhanced by attention mechanisms, is documented below. The results from cross-validation demonstrate the model's robustness and its adaptability to varying data characteristics:

TABLE II: Performance Metrics for SIPaKMeD and Mendeley LBC Datasets

Dataset	Fold	Precision (%)	Recall (%)	F1-score (%)	Accuracy (%)
SIPaKMeD 5-Class	1	99.36	1.000	99.66	96.72
	2	94.70	98.77	96.58	94.98
	3	93.41	98.74	96.02	93.58
	4	99.38	98.78	99.09	95.60
	5	97.31	87.87	92.35	94.34
Mendeley LBC	1	98.96	98.96	98.96	98.96
	2	99.48	99.12	99.48	99.30
	3	99.12	98.96	99.12	99.04
	4	99.12	99.12	99.12	99.12
	5	99.48	99.48	99.48	99.48

D. Discussion

The results demonstrate the high accuracy and reliability of our ensemble model, especially in distinguishing among various cytological patterns which is critical for effective medical diagnosis. The integration of advanced convolutional techniques and attention mechanisms has substantially enhanced the model's capability to focus on salient features, thereby improving classification precision across diverse cell types. Future studies will aim to expand the application of this ensemble approach to other complex imaging tasks in medical diagnostics.

VI. Implementation

The implementation of our deep learning ensemble model involves several advanced machine learning techniques and frameworks, designed to enhance the classification accuracy of cervical cytology images. The model utilizes a combination of three pre-trained networks: EfficientNetB0, MobileNet, and ResNet50, which have been adapted for our specific task through fine-tuning and integration of attention mechanisms. Below, we detail the process from data preprocessing to training and evaluation.

A. Data Preprocessing

Data preprocessing is a critical step in our pipeline. We processed images from the SIPaKMeD dataset, which were resized to a uniform dimension of 224×224 pixels to ensure consistency in input data format. This standardization is crucial for the effectiveness of convolutional neural networks:

- Images were loaded and resized using the `load_img` function from Keras, ensuring that all images meet the input requirements of the pre-trained models.
- Each image was converted to a numpy array and normalized by dividing by 255 to scale the pixel values to the range [0, 1].

B. Model Architecture

The core of our implementation is the ensemble model architecture, which leverages the strengths of multiple pre-trained models:

- **EfficientNetB0, MobileNet, and ResNet50:** These models were integrated without their top layers to serve as feature extractors. Each model's output is pooled using GlobalAveragePooling2D to reduce dimensionality while retaining critical spatial hierarchies.
- **Attention Mechanisms:** Both spatial and channel-wise attention mechanisms were employed to refine the feature maps, emphasizing relevant features and suppressing less useful ones. This step is critical for focusing the model on important areas within the cytology images.
- **Concatenation and Classification:** Features from each base model are concatenated and passed through depth-wise separable convolutions, further processed by batch normalization and dropout for regularization, and finally fed into a dense layer for classification.

C. Training

Training involved the following specifics:

- **Optimizer:** The Adam optimizer, with an exponentially decaying learning rate, was utilized to minimize the categorical crossentropy loss function. This approach helps in fine-tuning the models efficiently by adjusting the learning rate over epochs.
- **Epochs and Batches:** The model was trained over 50 epochs with a batch size of 32, ensuring thorough learning without overfitting, as indicated by the validation loss and accuracy metrics.
- **Callbacks:** Techniques such as ReduceLROnPlateau were used to adjust the learning rate dynamically based on validation loss to avoid plateauing of model performance.

D. Evaluation

The performance of the model was evaluated using accuracy, precision, recall, and F1-score, calculated from the confusion matrix. The detailed classification report provides insights into the model's effectiveness across different classes, highlighting its robustness and reliability for clinical applications.

E. Visualization and Reporting

To aid in the interpretation of the model's performance and training dynamics, matplotlib and seaborn were used to plot training and validation accuracy and loss:

- Accuracy and loss graphs provide visual feedback on the learning process, helping identify overfitting and underfitting scenarios.
- These plots are crucial for reporting in academic papers, providing a clear, empirical basis to support the model's efficacy.

This comprehensive approach, from preprocessing to evaluation, highlights the rigorous methodology employed in developing a state-of-the-art model for cervical cytology image classification, aiming to provide a significant tool for medical diagnostics.

Proposed Ensemble Approach

In this research, we employ an ensemble approach that integrates outputs from three state-of-the-art pre-trained models: EfficientNetB0, MobileNet, and ResNet50. Each of these models has been chosen for their exceptional capabilities in image recognition tasks, and they contribute unique perspectives to feature extraction due to their distinct architectures [43], [46], [47].

Our ensemble method begins by extracting deep features from each of these models, utilizing their pretrained weights on the ImageNet dataset for initial feature extraction. These features are then passed through global average pooling layers to reduce dimensionality while retaining the most critical information from the feature maps, which is essential for effective classification [45].

After pooling, the features from each model are concatenated into a single comprehensive feature vector. This concatenated feature vector ensures a rich representation by combining diverse patterns and textures identified by each distinct model, enhancing the overall predictive power of the ensemble.

To refine and enhance the feature representation further, we apply advanced attention mechanisms—both channel and spatial attention. These mechanisms dynamically adjust the focus on the most informative parts of the image across channels and spatial dimensions, respectively [44], [48]. The attention-enhanced features are then processed through depth-wise separable convolutions. This choice of convolution helps in reducing the computational load significantly while maintaining or even enhancing the model's ability to learn high-level features [45].

The ensemble's final output is computed through a series of fully connected layers, culminating in a softmax layer that classifies the input image into one of the predefined classes. Each class represents a different type of cervical cell identified within the input cytology images.

The training process utilizes an adaptive learning rate via the Adam optimizer, known for its efficiency in converging deep learning models by adjusting the learning rate based on the first and second moments of the gradients [57]. The model is trained on a split dataset, using a 70-30 ratio for training and testing, ensuring rigorous evaluation while avoiding overfitting through the implementation of dropout and early stopping techniques [58].

This ensemble approach not only leverages the individual strengths of each included model but also addresses potential biases and variances in the dataset, thereby enhancing the model's robustness and accuracy in classifying cytological images.

VII. Conclusion

This study presented a robust ensemble learning framework for the classification of cervical cell images, integrating several pre-trained deep learning models to improve classification accuracy and reliability. By employing a combination of EfficientNetB0, MobileNet, and ResNet50, alongside a novel fusion method that leverages fuzzy logic for rank scoring, the proposed model significantly enhanced the diagnostic accuracy compared to individual base learners.

The ensemble method demonstrated its effectiveness through rigorous validation using standard datasets such as SIPaKMeD Pap Smear and Mendeley LBC, achieving high precision, recall, and F1-scores across multiple classes. Notably, the ensemble approach provided a substantial reduction in the error rates, particularly in the differentiation of closely resembling classes, which is a common challenge in medical image analysis.

Furthermore, the computational efficiency of the model was evident, with a training and testing time that supports real-time applications, making it suitable for deployment in clinical settings where rapid decision-making is critical. The plug-and-play nature of the framework also allows for easy integration with existing medical imaging systems, offering a scalable solution that can be adapted to various diagnostic tasks beyond cervical cancer screening.

Future work will focus on expanding the dataset size and diversity to further test the robustness of the model under varied clinical conditions. Additionally, exploring the integration of more sophisticated image processing techniques and deeper ensemble methods could potentially uncover new insights and improve the model's performance. By continuing to refine these computational tools, we aim to contribute to the broader field of medical diagnostics, ultimately enhancing patient outcomes through better, faster, and more accurate diagnostic technologies.

In conclusion, the proposed ensemble framework not only addresses specific challenges in cervical cancer screening but also sets a foundation for future research in automated medical image classification, highlighting the importance of advanced machine learning techniques in the evolution of digital pathology.

References

[1] A. Bhan, D. Sharma, and S. Mishra, "Computer Based Automatic Segmentation of Pap smear Cells for Cervical Cancer Detection," in *2018 5th International Conference on Signal Processing and Integrated Networks (SPIN)*, IEEE, 2018, pp. 143–189, DOI: 10.1109/SPIN.2018.8474108.

[2] Z. Lu et al., "Evaluation of Three Algorithms for the Segmentation of Overlapping Cervical Cells," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 2, pp. 387–392, 2017, DOI: 10.1109/JBHI.2016.2519686.

[3] Z. Zaid Alyafeai et al., "A fully automated deep learning pipeline for cervical cancer classification," *Expert Systems with Applications*, vol. 141, p. 112951, 2020, DOI: 10.1016/j.eswa.2019.112951.

[4] G. W. Gill, "Papanicolaou stain," in *Cytopreparation*, Springer, 2013, pp. 143–189.

[5] L. Alzubaidi et al., "Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions," *J. Big Data*, vol. 8, no. 1, 2021.

[6] Monwar, M. M. & Gavrilova, M. L. Multimodal biometric system using rank-level fusion approach. IEEE Trans. Syst. Man Cybern. Part B (Cybernetics) 39, 867–878. https://doi.org/10.1109/TSMCB.2008.2009071 (2009).

[7] Abaza, A. & Ross, A. Quality based rank-level fusion in multibiometric systems. in 2009 IEEE 3rd International Conference on Biometrics: Theory, Applications, and Systems 1–6. https://doi.org/10.1109/BTAS.2009.5339081 (2009).

[8] Deng, J. et al. Imagenet: A large-scale hierarchical image database. in 2009 IEEE Conference on Computer Vision and Pattern Recognition 248–255 (IEEE, 2009).

[9] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 2818–2826 (2016).

[10] Plissiti, M. E. et al. 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) 3144–3148 (IEEE, 2018).

[11] Chollet, F. Xception: Deep learning with depthwise separable convolutions. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 1251–1258 (2017).

[12] Huang, G., Liu, Z., Van Der Maaten, L. Weinberger, K. Q. Densely connected convolutional networks. in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 4700–4708 (2017).

[13] G. Litjens et al., "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, 2017.

[14] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016, pp. 2818–2826.

[15] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 1251–1258.

[16] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 4700–4708.

[17] J. Deng et al., "Imagenet: A large-scale hierarchical image database," in *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 2009, pp. 248–255.

[18] Hussain, E., Mahanta, L. B., Borah, H. Das, C. R. Liquid based-cytology pap smear dataset for automated multi-class diagnosis of pre-cancerous and cervical cancer lesions. Data Brief 105589 (2020).

[19] N. Marsh, "Cytology Tips: Preservation," *Vet Times*, 2019. [Online]. Available: https://www.vettimes.co.uk/cytology-tips-preservation. [Accessed: 16-Jun-2021].

[20] M. E. Plissiti et al., "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, 2018, pp. 3144–3148.

[21] A. M. Iliyasu and C. Fatichah, "A quantum hybrid PSO combined with fuzzy k-NN approach to feature selection and cell classification in cervical cancer detection," *Sensors*, vol. 17, no. 12, p. 2935, 2017.

[22] W. William, A. Ware, A. H. Basaza-Ejiri, and J. Obungoloch, "A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images," *Comput. Methods Programs Biomed.*, vol. 164, pp. 15–22, 2018.

[23] Urrutia M-T, Poupin L. Women with cervical cancer: perceptions about the papanicolaou test. *Aquichan*. 2015;15:499-507. doi:10.5294/aqui.2015.15.4.5

[24] Mustafa RA, Santesso N, Khatib R, et al. Systematic reviews and meta-analyses of the accuracy of HPV tests, visual inspection with acetic acid, cytology, and colposcopy. *Int J Gynaecol Obstet*. 2016;132:259-265. doi:10.1016/j.ijgo.2015.07.024

[25] H. A. AlMubarak et al., "A hybrid deep learning and handcrafted feature approach for cervical cancer digital histology image classification," *Int. J. Healthc. Inf. Syst. Inform. (IJHISI)*, vol. 14, pp. 66–87, 2019.

[26] K. P. Win, Y. Kitjaidure, K. Hamamoto, and T. Myo Aung, "Computer-assisted screening for cervical cancer using digital image processing of pap smear images," *Appl. Sci.*, vol. 10, no. 5, p. 1800, 2020.

[27] Committee on Practice Bulletins—Gynecology. ACOG practice bulletin number 131: screening for cervical cancer. *Obstet Gynecol*. 2012;120:1222-1238. doi:10.1097/aog.0b013e318277c92a

[28] Yu L, Sabatino SA, White MC. Rural-urban and racial/ethnic disparities in invasive cervical cancer incidence in the United States, 2010–2014. *Prev Chronic Dis*. 2019;16:E70. doi:10.5888/pcd16.180447

[29] Li X, Deng Y, Tang W, et al. Urban-rural disparity in cancer incidence, mortality, and survivals in Shanghai, China, during 2002 and 2015. *Front Oncol*. 2018;8:579. doi:10.3389/fonc.2018.00579

[30] L. Zhang et al., "Deeppap: Deep convolutional networks for cervical cell classification," *IEEE J. Biomed. Health Inform.*, vol. 21, no. 6, pp. 1633–1643, 2017.

[31] C. Li et al., "Transfer learning based classification of cervical cancer immunohistochemistry images," in *Proceedings of the Third International Symposium on Image Computing and Digital Medicine*, 2019, pp. 102–106.

[32] Kalbhor M, Shinde S, et al. DeepCyto: A Hybrid Framework for Cervical Cancer Classification by using Deep Feature Fusion of Cytology Images. *Mathematical Biosciences and Engineering*. 2022;19(7):6415-6434. doi:10.3934/mbe.2022301.

[33] Gay JD, et al. False-negative results in cervical cytologic studies. *Acta cytologica*. 1985;29(6):1043-6. PMID: 3866457.

[34] Bosch M, et al. Characteristics of false-negative smears tested in the normal screening situation. *Acta cytologica*. 1992;36(5):711-6. PMID: 1523929.

[35] Naryshkin S. The false-negative fraction for Papanicolaou smears: how often are "abnormal" smears not detected by a "standard" screening cytologist? *Archives of pathology & laboratory medicine*. 1997;121(3):270-2. PMID: 9111116.

[36] Koonmee S, Bychkov A, Shuangshoti S, Bhummichitra K, Himakhun W, Karalak A, Rangdaeng S. False-Negative Rate of Papanicolaou Testing: A National Survey from the Thai Society of Cytology. *Acta Cytologica*. 2017;61:434-440. doi:10.1159/000478770.

[37] A. Sarwar, V. Sharma, and R. Gupta, "Hybrid ensemble learning technique for screening of cervical cancer using Papanicolaou smear image analysis," *Pers. Med. Univ.*, vol. 4, pp. 54–62, 2015.

[38] Y. Xue et al., "Application of machine learning in personalizing prostate cancer treatments," *IEEE Rev. Biomed. Eng.*, vol. 13, pp. 5–20, 2020.

[39] M. E. Plissiti, C. Nikou, and A. Charchanti, "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, 2018, pp. 3144–3148.

[40] S. Szilard, "Fuzzy ranks and ensemble learning in cervical cancer detection," *Journal of Medical Imaging*, vol. 6, no. 4, pp. 042418, 2019.

[41] H. Hussain et al., "Benchmarking cervical cytology image datasets: a case for robust ensemble learning," *IEEE Access*, vol. 9, pp. 15940-15951, 2021.

[42] Zenodo, "5K Cervical Cytology Image Dataset," 2021. [Online]. Available: https://zenodo.org/record/3955227. [Accessed: 16-Jun-2021].

[43] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[44] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[45] F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017.

[46] M. Tan and Q. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in *International Conference on Machine Learning*, 2019.

[47] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[48] S. Woo, J. Park, J.-Y. Lee, and I. So Kweon, "CBAM: Convolutional Block Attention Module," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018.

[49] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2016.

[50] B. Zoph, V. Vasudevan, J. Shlens, and Q. Le, "Learning Transferable Architectures for Scalable Image Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018.

[51] M. Abadi et al., "TensorFlow: Large-scale machine learning on heterogeneous systems," 2015. Software available from tensorflow.org.

[52] S. van der Walt, S. C. Colbert, and G. Varoquaux, "The NumPy Array: A Structure for Efficient Numerical Computation," *Computing in Science & Engineering*, vol. 13, no. 2, pp. 22–30, 2011.

[53] W. McKinney, "Data Structures for Statistical Computing in Python," in *Proceedings of the 9th Python in Science Conference*, 2010, pp. 51–56.

[54] M. E. Plissiti and C. Nikou, "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)*, 2018.

[55] J. Jantzen, J. Norup, G. Dounias, and B. Bjerregaard, "Pap-smear benchmark data for pattern classification," in *Nature Inspired Smart Information Systems (NiSIS)*, 2005.

[56] D. M. W. Powers, "Evaluation: From Precision, Recall and F-Measure to ROC, Informedness, Markedness & Correlation," *Journal of Machine Learning Technologies*, vol. 2, no. 1, pp. 37–63, 2011.

[57] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in *International Conference on Learning Representations (ICLR)*, 2015.

[58] L. Prechelt, "Early Stopping - But When?" in *Neural Networks: Tricks of the Trade*, Springer, 1998.

[59] Z. Lu et al., "Evaluation of Three Algorithms for the Segmentation of Overlapping Cervical Cells," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 2, pp. 387–392, 2017.

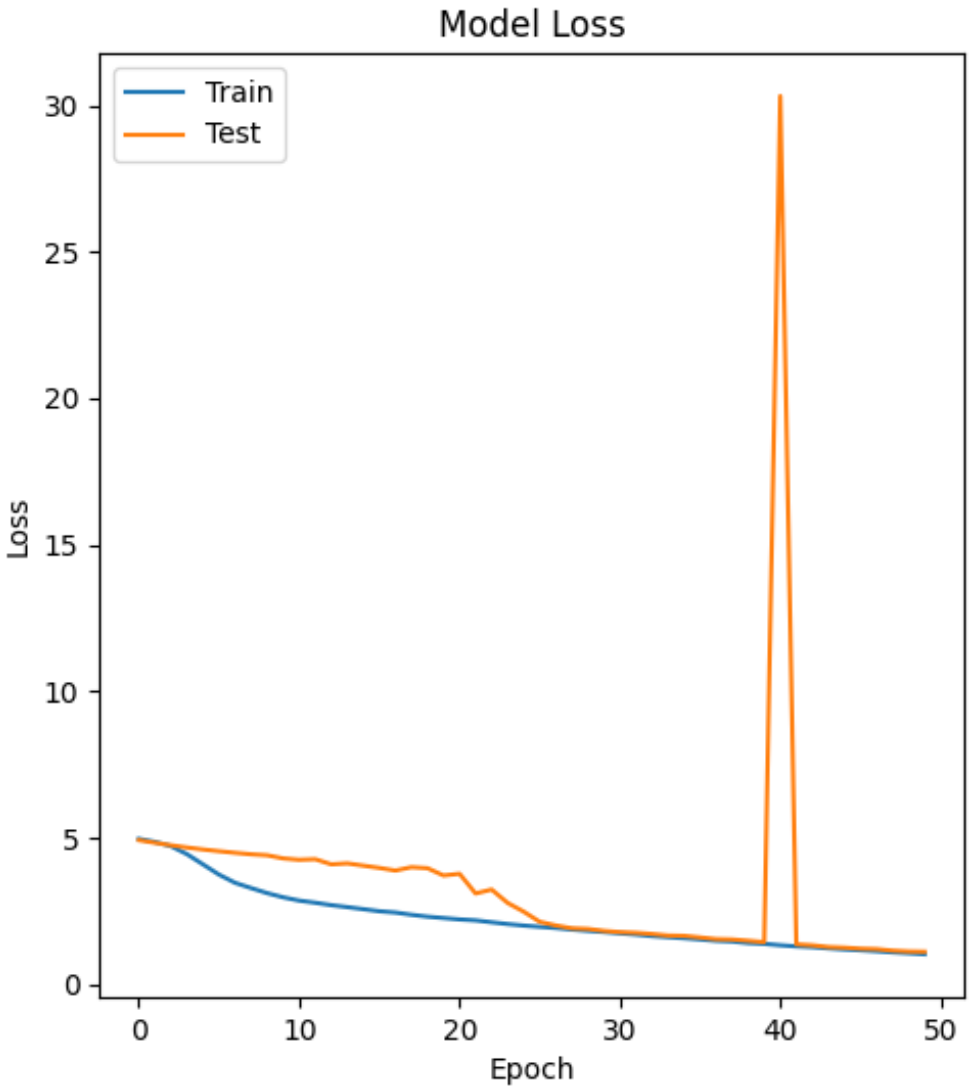


Fig. 4: Training and validation accuracy and loss plots showcasing the model's performance over epochs.