HYBRID ENSEMBLE OF TRANSFER LEARNING AND CUSTOM CNN ARCHITECTURES FOR AUTOMATED CLASSIFICATION OF CERVICAL CANCER CELLS FROM PAP SMEAR IMAGES

Thesis Submitted
in Partial Fulfilment of the Requirements for the
Degree of

MASTER OF SCIENCE
in
BIOTECHNOLOGY

by
TANU PALIWAL
(24/MSCBIO/13)

Under the Supervision of

PROF. YASHA HASIJA
Professor and Head of Department
Department of Biotechnology
Delhi Technological University

Department of Biotechnology

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-110042, India

May, 2026

ACKNOWLEDGEMENT

I would like to express my sincere gratitude to my supervisor, Prof. Yasha Hasija, Department of Biotechnology, Delhi Technological University, for her invaluable guidance, constant encouragement, and insightful suggestions throughout the course of this thesis work. Her support and mentorship greatly contributed to the successful completion of this study.

I am also thankful to the faculty members and staff of the Department of Biotechnology, Delhi Technological University, for providing the necessary facilities, academic environment, and support required for this research.

I extend my heartfelt gratitude to PhD scholars, Ms. Neha, Mr. Shubham Shrivastav and Ms. Akansha Bisht for their unwavering support and guidance. Their expertise, patience, and mentorship was truly valuable.

I would like to thank my friends and family for their support and belief in me during every stage of this work. Their encouragement has always been my greatest strength.

Finally, I express my gratitude to everyone who directly or indirectly contributed to the completion of this thesis.

TANU PALIWAL

24/MSCBIO/13

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I, Tanu Paliwal, bearing Roll No. 24/MSCBIO/13 hereby certify that the work which is being presented in the thesis entitled "HYBRID ENSEMBLE OF TRANSFER LEARNING AND CUSTOM CNN ARCHITECTURES FOR AUTOMATED CLASSIFICATION OF CERVICAL CANCER CELLS FROM PAP SMEAR IMAGES" in partial fulfilment of the requirements for the award of the Degree of Master of Science, submitted in the Department of Biotechnology, Delhi Technological University is an authentic record of my own work carried out during the period from January 2026 to May 2026 under the supervision of Prof. Yasha Hasija.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other Institute.

Tanu Paliwal

24/MSCBIO/13

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR

Certified that **TANU PALIWAL** (24/MSCBIO/13) has carried out their search work presented in this thesis entitled **"HYBRID ENSEMBLE OF TRANSFER LEARNING AND CUSTOM CNN ARCHITECTURES FOR AUTOMATED CLASSIFICATION OF CERVICAL CANCER CELLS FROM PAP SMEAR IMAGES"** for the award of **Master of Science** from Department of Biotechnology, Delhi Technological University, Delhi, under my supervision. The thesis embodies results of original work, and studies are carried out by the student herself, and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

Date:

Prof. Yasha Hasija
Professor and Head of Department
Department of Biotechnology
Delhi Technological University
Delhi - 110042

"HYBRID ENSEMBLE OF TRANSFER LEARNING AND CUSTOM CNN ARCHITECTURES FOR AUTOMATED CLASSIFICATION OF CERVICAL CANCER CELLS FROM PAP SMEAR IMAGES"

Tanu Paliwal
24/MSCBIO/13

ABSTRACT

Background: The occurrence rate of cervical cancer has been increasing in recent years and the mortality rates because of it have increased rapidly and can be commonly seen in women across the world. This is greatly due to the fact that cervical cancer has no visible symptoms, or it is usually asymptomatic in its early stages. It is detected in the later stages when it has become malignant and is spreading and causing harm to other body systems. If it gets detected in early stages, treatment is possible and effective as compared to later stages. Screening for cervical cancer is usually done using conventional Pap Smear test in which cervical cells are examined for abnormalities using microscopic examinations. This process is cumbersome, time consuming and requires expert pathologists otherwise there will be a lot of variability and errors in the final results. To address these challenges, the present work aims to provide an automated and robust system using DTL and ensemble techniques for accurate classification of cervical cancer.

Methodology: This study uses three pre-trained CNNs - ResNet152, EfficientNetV2-S, ConvNeXt-Base and a Custom CNN with residual blocks and CBAM attention mechanism to classify cervical cancer cells of SIPaKMeD dataset into 5-classes. A novel ensemble of these four base learners was proposed by averaging the predictions from each of the four models and then finally predicting one class from the dataset for a given input image.

Result: For the four base learners the accuracy was 96.71%, 95.56%, 94.57% and 90.62% for ResNet152, EfficientNetV2-S, ConvNeXt-Base and Custom CNN respectively. The ensemble model in the current research gave a classification accuracy of 97.53% on the 5-class SIPaKMeD dataset, higher than those of individual base learners.

Conclusion: The performance of proposed ensemble reflects its superiority over base learners and to some the previous studies as well. A method for classifying cervical cancer cells from Pap smear images is provided by the current research project. Deployment of such systems in real world clinical settings can help medical professionals for better treatment plans and also improve patient experience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
HPV	Human Papilloma Virus
CAD	Computer Aided Diagnosis
TL	Transfer Learning
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
GAN	Generative Adversarial Network
WSI	Whole Slide Image
PCA	Principal Component Analysis
SVM	Support Vector Machine
D-CNN	Deep Convolutional Neural Network
CLAHE	Contrast Limited Adaptive Histogram Equalization
ViT	Vision Transformer
CBAM	Convolutional Block Attention Module
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve

CHAPTER 1

INTRODUCTION

Cervical cancer develops in the lower portion of the female reproductive tract arises from rapid and uncontrolled division of cervical cells, transforming into malignant tumour with time, spreading, and affecting other body systems [1]. It comes fourth amongst the most commonly seen cancer cases in the females across the world [2]. In early stages it is slow growing and usually asymptomatic in nature, which makes its detection difficult. If detected precisely in early stages, treatment is effective and cancer can be prevented from becoming malignant.

In most of the cases, HPV (Human Papilloma Virus) infection is the major cause of the disease in women. Currently, HPV vaccines are there in the market which can be used as one of the preventive methods [3] [4].

The Papanicolaou Smear test is a detection method for cervical cancer in which abnormalities in cervical cells are detected through microscopic examinations [5]. One of the limitations of this is that the procedure is complex, exhaustive and time consuming, requiring the expertise of a pathologist [6]. As the procedure is manual and human intervention is there, chances of erroneous results increase. Thus, accurate screening of cervical cancer requires advanced, automated, and faster methods which are also reliable to be used in the real world.

Due to the limitations of manual screening methods, research is being done to develop CAD systems which can analyse the cervical cell images precisely and classify them accurately into precancerous or cancerous lesions. These systems use DL algorithms because of their capability of precise classification of cervical cells efficiently [7]. Deep learning techniques are equipped with good feature extraction qualities from images so they can easily detect subtle changes in the morphology of cell cytoplasm and nucleus which is sometimes missed using manual methods [8] [9]. One of the requirements for using these deep learning algorithms is large datasets for training them. Medical datasets are usually limited and small sized due to privacy and other concerns.

To overcome this problem of limited medical dataset, Transfer Learning (TL) has become a widely adopted approach. In TL, a deep learning model is trained on larger datasets like ImageNet data, and this step is called pre-training. The pretrained model learns the general features of an image from the larger dataset, and in the next step of fine-tuning, same model is fine-tuned on a domain specific data to learn domain specific features. Fine-tuning is done by unfreezing few initial layers of the model to learn features of the given domain/problem while freezing the weights of remaining layers. Thus, TL helps in knowledge transfer and improves model performance even with less amount of data [10][11].

Different models have different strengths, and they make predictions based on those strengths, capturing different features from the same dataset. To make use of this, Ensemble techniques can be used. Ensemble is an approach which integrates predictions generated by multiple individual models (base learner) in order to produce one single model with better predictions as compared to any single base learner and thus making the classification more accurate and robust [12]. To construct more reliable systems for cervical cancer screening and diagnostic applications, ensemble models are being used these days, as they reduce the errors made by constituent learners [13] and thus resulting in better overall classification performance.

The primary contributions of the research conducted are summarized as follows:

1. To achieve competitive performance, a system for robust categorization of cervical cancer cells employing DL and TL approaches was designed.
2. ResNet152, EfficientNetV2-S, ConvNeXt Base, and a Custom CNN were used to form an ensemble that combined the advantages of several deep learning models.
3. To the best of our knowledge, ConvNeXt is not used very often for cervical cell classification, but here we used it as it has features of vision transformer and achieved good performance for the classification task.
4. The Custom CNN's architecture is completely novel with residual connections and CBAM attention for attending to relevant features.
5. The model proposed in the study outperformed many of the previous studies for SIPaKMeD 5-class classification problem.

CHAPTER 2

LITERATURE REVIEW

This chapter gives reviews and analysis of the previous research related to cervical cancer detection utilizing ML, DL and other computer assisted algorithms.

2.1 Data Augmentation Techniques

Data limitation remains a problem in medical domain; to address this many studies have used data augmentation techniques by synthetically generating additional synthetic images to enlarge the dataset. This can be done using traditional augmentation techniques or using generative models like GANs.

For increasing the size of cervical cytology dataset, [14] proposed a novel RES_DCGAN (residual deep convolutional GAN) and then for classification used ResNet50V2, Xception, and DenseNet121 with self-attention mechanism. RES_DCGAN helped with data augmentation and self-attention helped the network to focus on the important discriminative image features. This method improved performance of classification on SIPaKMeD dataset and helps in increasing the generalizability of the model.

[15] used GAN-based data augmentation methods to artificially generate new samples and expand the dataset size. They then incorporated EfficientNet for classification, and obtained an accuracy of 99.1% using GAN compared to 98.8% using conventional augmentation techniques.

2.2 Transfer Learning and Pre-trained CNN Architectures

TL helps address the issue of limited size of medical datasets. Pre-trained models can be used to learn generalized features of an image and that knowledge can later be transferred to do domain specific classification using fine-tuning of pre-trained models.

Kaur et. al. [11] compared many pre-trained DTL models using pap smear datasets (Herlev and SIPaKMeD) for classification of cervical cancer cells to see which TL model is giving good performance over others and can be used for proper screening. The results revealed that ResNet50 performed well for Herlev and VGG16 and DenseNet121 were good for SIPaKMeD.

[16] conducted a study using various pre-trained CNNs for classification of cervical cells on SIPaKMeD and Herlev datasets. They demonstrated that for classification of WSI cervical cells, segmentation is not necessary. This idea resulted from the

experiments that were conducted with segmentation, and their performance was not good compared to without segmentation.

The authors of [17] proposed a multi-deep transfer learning approach with smoothing cross entropy for detection of cervical cancer cells properly. The approach follows four stages – pre-processing of the images, extraction of features using various TL models and their fusion, PCA for feature reduction and finally classification using cross entropy loss. The model gave a competitive performance of 97% on SIPaKMeD dataset.

Hashem et. al. [18] stressed on the data imbalance challenges faced during classification of medical image data. They conducted their study on SIPaKMeD dataset's two distributions using DL models (ConvNets and Vision Transformers). The five-class distribution is balanced and achieved an accuracy of 100% by fine tuning of DenseNet-121 and EfficientNetv2 B0 while for two-class classification accuracy was 98.09% depicting a decrease in performance of the model with imbalanced dataset.

2.3 Custom CNN Architectures for Cervical Cell Classification

Several researchers have used ML and DL models trained from scratch or with some constraints. A multi-structural CNN, CerviXpert was given by [19] which was trained from scratch with limited convolutional layers and dense layers. When compared to SOTA CNNs, the system's ability to balance computational economy and accuracy made it suitable for deployment in resource-constrained applications.

DeepCELL, a deep convolutional neural network given by Fang et. al. [20] automatically classifies cervical cell images by utilizing learned feature representations from many kernels of varying size. Performance evaluation shows an accuracy level of 95.628% for the SIPaKMeD dataset and a decent performance on Herlev dataset as well.

A lightweight model named CCanNet was introduced by [21] with only 1.27M parameters and performance like other transfer learning and transformer-based models. The model was an integration of blocks containing residual connections, squeeze connections and skip connections. For enhancing the trust and reliability of the model, explainability was introduced by use of GradCam and thus making model predictions more transparent.

[22] proposes A2SDNet121 model which is a DenseNet121 with additional features. SE module is there to focus on nucleus and size of kernel is different and Atrous Dense Block is added to extract global and local features. The performance on SIPaKMeD dataset was good with accuracy >99% for multiclass classification.

These works demonstrated that custom CNNs with specific features trained from scratch achieves performance comparable to pre-trained CNNs.

2.4 Hybrid and Feature Fusion Based Methods

Hybrid learning approaches use both the DL and conventional ML approaches to build models with great outputs.

Chauhan et. al. [23] applied a Hybrid Learning Network on WSI of cervical cancer. They used ResNet-152 and VGG-16 for training the data that was expanded and resized to 224 x 224, 512 x 512 and 1024 x 1024. The concatenated features from two models were reduced by PCA (Principal Component Analysis). Then ML models (SVM and Random Forest) with majority voting were used for final classification.

A hybrid deep feature concatenated network (HDFCN) was given by [24] which implements two step data augmentation and combining the features coming after fine tuning of DL models. This was done on WSI of SIPaKMeD dataset for which the accuracy of classification was 97.45% for 5-class and 99.29% for 2-class.

2.5 Ensemble Learning Techniques

Ensemble learning is an emerging technique as it combines the strengths of different models and helps produce models that are more robust and have better performance and generalizability than base learners.

Hanzala et. al. [10] used four D-CNNs and four TL models and then compared their performances. Based on the performance evaluation they proposed a novel ensemble AZL using AlexNet, ZfNet and LeNet. The ensemble gave an accuracy of 99.92% for cervical cancer classification, better than individual D-CNN or TL models.

[12] proposed an ensemble model that achieves better performance than individual base learners (CNN, AlexNet, SqueezeNet) for categorization of squamous cells of the cervix into five classes.

[13] proposed an ensemble model with three base classifiers – Inception v3, Xception and DenseNet-169. A fuzzy rank-based ensemble of the base learners was used to increase the performance of the ensemble model on the SIPaKMeD and the Mendeley LBC datasets. The model was also tested on Zenodo 5K dataset to justify its performance.

A fuzzy distance-based ensemble was given by [25], with performance of 96.96% which was higher than those of the base learners (Inception V3, MobileNet V2 and

Inception ResNet V2) where some additional layers learning specific features of the data were incorporated. The final prediction is based on the minimum distance (loss) for each classification.

DeepCervixNet, an ensemble of ResNet101 and DenseNet169 incorporating SE (sequence and excitation) blocks was introduced by [26]. It gave a competitive performance of 99.89% for classification of cervical cells using Herlev dataset.

[27] uses U-Net for segmentation of the image and then classify it based on the prediction of an ensemble. The model was tested on 3 different datasets and for SIPaKMeD's binary classification, ensemble achieved an accuracy of 99%.

[28] used pre-trained CNNs for extraction of the features and PCA for reducing the number of dimensions and then classifying using an ensemble of different ML algorithms. The accuracy level of 97.03% (2-class) and 96.67% (5-class) was obtained using the model.

2.6 Emerging DL Trends in Detection of Cervical Cancer

Current research is also exploring advanced trends of deep learning like use of reinforcement learning and vision transformers. Vision transformers are based on attention mechanism which help focus on the important regions of an image for better classification performance.

RLCancerNet (Reinforcement Learning Cancer Network) is a novel framework introduced by [29] which combines CNN with reinforcement learning algorithms. EfficientNetV2 was the base model combined with supporter blocks for better performance.

[30] conducted a comprehensive comparison of 40 CNNs and 20 ViT models on SIPaKMeD with some augmentation and ensemble approaches. The results highlighted that Vit-based models outperformed CNN-based models and exceeds the results of previous studies.

2.7 Research Gaps

Previous studies have reported strong performance for classification of cervical cancer, though limitations are there. Most of the approaches don't combine strengths of different CNN families, relying only on one or two of them. Lightweight models are being used in some approaches but they compromise on accuracy of the model. Some of the ensemble approaches demonstrate great performance but using closely related

architectures, thus reducing the diversity in learning of models. Moreover, attention-based architectures are less explored in this field of cervical cancer classification.

Therefore, the current research proposes a hybrid ensemble framework combining heterogenous TL models and a custom CNN with residual blocks and CBAM attention mechanisms to complement each other and classify cervical cells accurately.

A summary of few of the above works is presented in Table 1 for better understanding of the previous literature.

Table 1: A comprehensive analysis of previous studies for automated cervical cancer classification of cervical cancer

Reference	**Approach**	**Key Models / Techniques**	**Dataset(s)**	**Performance**
[10]	Hybrid ensemble of D-CNN and transfer learning models	AZL ensemble (AlexNet + ZFNet + LeNet)	Cervical cancer dataset	99.92% accuracy
[11]	Comparison of pre-trained DTL models	16 DTL models (ResNet50 best on Herlev; VGG16 & DenseNet121 best on SIPaKMeD)	Herlev, SIPaKMeD	Best-performing models identified per dataset
[13]	Fuzzy rank-based ensemble of CNNs	Inception v3 + Xception + DenseNet-169 with fuzzy ranking	SIPaKMeD, Mendeley LBC, Zenodo 5K	98.55% and 95.43% accuracy of ensemble for 2-class and 5-class respectively
[29]	Reinforcement-learning-enhanced CNN framework (RLCancerNet)	EfficientNetV2 base + supporter blocks + RL	Herlev, SIPaKMeD	99.7% accuracy of proposed model
[19]	Multi-structural CNN trained from scratch	Custom lightweight CNN (limited conv + dense layers)	SIPaKMeD	98.6% accuracy (5-class) and 98.04% accuracy (3-class)
[25]	Fuzzy distance-based ensemble	Inception V3 + MobileNet V2 + Inception ResNet	SIPaKMeD	96.96% accuracy of ensemble

		V2 + task-specific layers; min-distance voting		
[30]	Large-scale comparison of CNNs vs. ViTs (with augmentation & ensembles)	40 CNN-based + 20 ViT-based models	SIPaKMeD	ViT models outperformed CNNs and prior SOTA
[23]	Hybrid learning network with feature concatenation + ML voting	ResNet-152 + VGG-16; PCA reduction; SVM + RF majority voting	SIPaKMeD, Mendeley LBC	Accuracy of 99.29% (2-class) and 98.47% (5-class) for SIPaKMeD dataset; 100% for LBC dataset
[18]	Deep learning models addressing data imbalance	ConvNets & ViTs on balanced vs. imbalanced data	SIPaKMeD (5-class & 2-class distributions)	100% (5-class); 98.09% (2-class)
[24]	Hybrid deep feature concatenated network (HDFCN)	Feature concatenation of fine-tuned DL models + two-step data augmentation	SIPaKMeD WSI (5-class & 2-class), LBC	97.45% (5-class), 99.29% (2-class)

CHAPTER 3

METHODOLOGY

The given chapter highlights the dataset used for the experiment, preprocessing done on the images, base learners used and the ensemble that was proposed for the experiment.

3.1 Dataset Description

The publicly available dataset SIPaKMeD was used for the research. It is made up of 4049 individual cell images that were extracted after cropping them from 966 pap smear slides of cervical tissue. It is divided into five classes: Superficial-Intermediate (SI), Parabasal (Pa), Koilocytotic (K), Metaplastic (Me), and Dyskeratotic (D). They can also be classified into three classes with Superficial-Intermediate and Parabasal cells being classified as normal or non-cancerous cells, Metaplastic as benign and Koilocytotic & Dyskeratotic as abnormal cells [16] [31]. The distribution of images per class is presented in Figure 1 using bar graph. Few of the raw sample images before transformation for each of the five classes are shown in Figure 2. The dataset was taken from https://www.cs.uoi.gr/~marina/sipakmed.html.

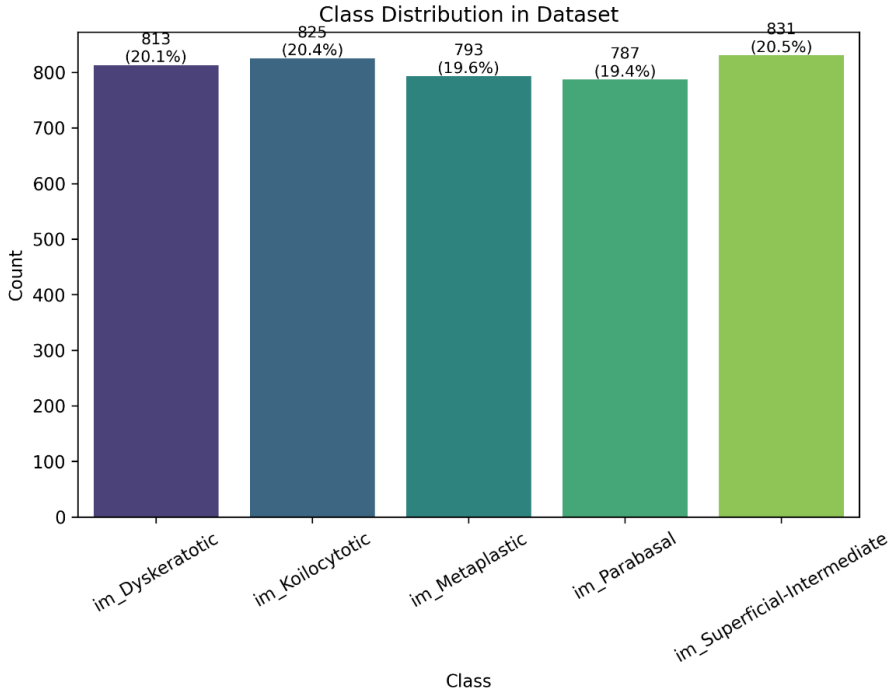

Fig. 1: Distribution of images in each class of the dataset

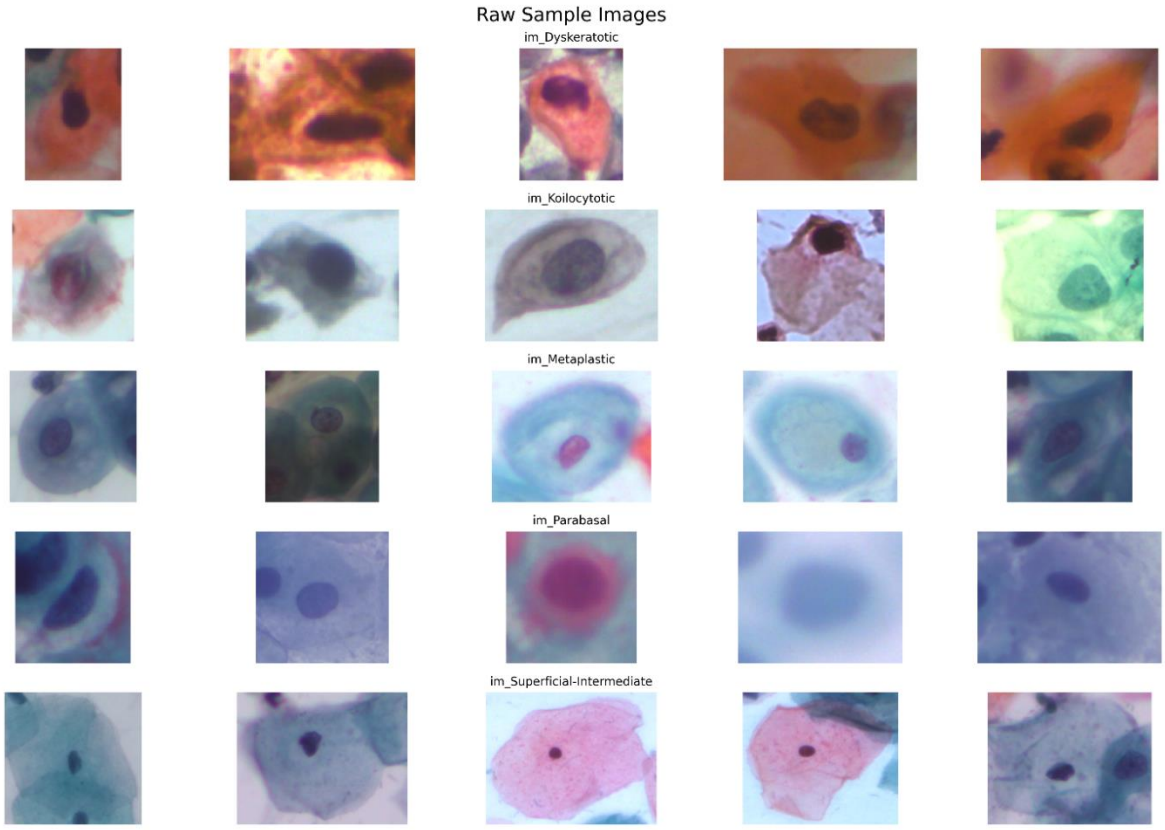

Fig. 2: Raw images for each of the dataset's five classifications

3.2 Data Preprocessing and Data Augmentation

The data was divided into subgroups for testing (15%), validation (15%), and training (70%) for balanced class representation across subgroups. This was done to provide sufficient samples for training the model, for hyperparameter tuning and performance evaluation. After splitting, 2834 training, 607 validation and 608 test images were obtained. The categorical labels were encoded to numeric ones using label encoding for compatibility with models.

The preprocessing steps applied before data augmentation included edge preserving smoothing using OpenCV which reduces the noise and helps preserve the boundaries of cells in the images which is critical for classification. Channel wise CLAHE enhancement was applied to LAB colour space to enhance local contrast for better visibility of nuclei and cytoplasm for better morphological features.

To increase the dataset, enhance model generalization, and lessen overfitting, augmentations were only used on the training set. This was done using Albumentations library. The transformations applied were:

- Affine (rotation ±15°, scaling 0.8-1.2×, translation ±15%, shearing ±2°, p=1.0) - geometric transformation technique used to increase dataset diversity, reduce overfitting and increase robustness of the model. With 100% probability each sample is transformed, and model sees a variation of that sample.
- Horizontal and Vertical Flips (p=1.0) - flips the image along its vertical and horizontal axis respectively with 100% probability [32].
- CLAHE [33] with clip limit 1-10, tile size 3x21 with 50% probability – helps enhance local features of the image while reducing noise.

After augmentation, each image was adjusted to a fixed size of 224 x 224, normalized to [-1,1] using mean (0.5, 0.5, 0.5) and standard deviation (0.5, 0.5, 0.5) and were converted to PyTorch tensors.

The size of validation and test sets were not increased using augmentations and only resizing, normalization and conversion to tensors were applied for consistent evaluation across all the images.

This preprocessing pipeline was designed to ensure that cytological features are enhanced while preserving the morphology of cells, for models to understand and classify them properly giving robust performance.

3.3 Base Learners

3.3.1 ResNet152

ResNet152 is a variant of ResNet family of deep learning CNNs with 152 layers. The problem with very deep networks is that performance decreases with increase in the number of layers but ResNet152 tackles it well by using identity mapping from one layer to the next by incorporating residual connections. This helps to improve the decreasing accuracy with deeper architecture and also helps with vanishing gradient problem during backpropagation for weight updation. The first 7 x 7 convolution and pooling layers of the design are followed by numerous bottleneck residual blocks of 1 x 1, 3 x 3, and 1 x 1 convolutions, average pooling, and the final classification layer [34]. The architecture of ResNet152 trained on ImageNet dataset is represented in Figure 3.

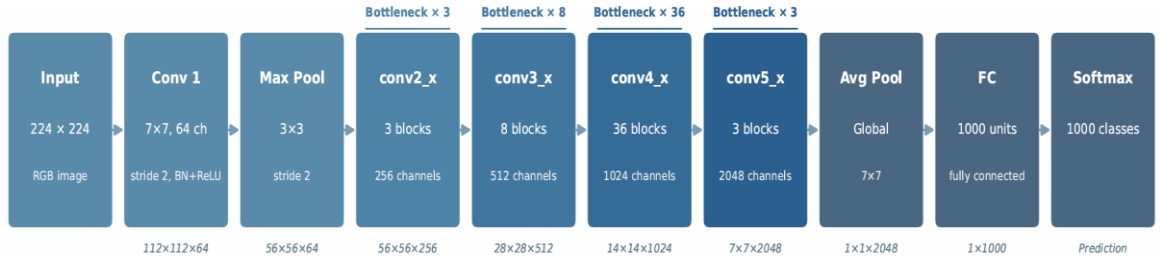

Fig. 3: Architecture of ResNet152 (Classes are 1000 as it is trained on ImageNet dataset, after fine-tuning to current dataset 5 classes are there)

3.3.2 EfficientNetV2-S

EfficientNetV2-S is a member of EfficientNetV2 family, which demonstrates faster training with efficient parameters. It utilizes a combination of both MBConv and fused-MBConv in the initial layers of the model. For better channel-wise feature recalibration, MBConv employs depth-wise separable convolutions with SE (squeeze and excitation) methods. The fused-MBConv uses standard convolutions instead of depth-wise to enhance training speed. The architecture starts with 3 x 3 convolutions followed by fused-MBConv and MBConv blocks and terminating with a 1 x 1 convolution, a pooling layer and a final classification head [35]. Figure 4 shows the architecture of EfficientNetV2-S model, trained on ImageNet dataset with 1000 classes.

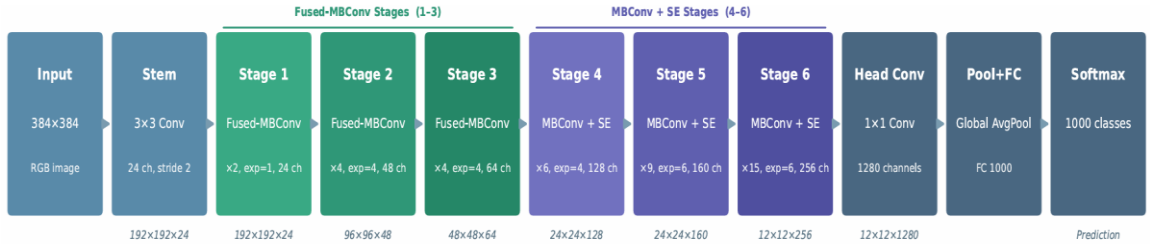

Fig. 4: Architecture of EfficientNetV2-S (Classes are 1000 as it is trained on ImageNet dataset, after fine-tuning to current dataset 5 classes are there)

3.3.3 ConvNeXt-Base

ConvNeXt has a CNN base with some features from modern vision transformers to improve performance. A standard CNN, ResNet was redesigned with ViT features but without attention mechanism to give ConvNeXt model. It uses ConvNeXt blocks which are composed of large (7 x 7) kernel size depthwise convolutions, batch normalization replaced with layer normalization, and ReLU activation replaced with GELU to improve feature extraction. It also employs inverted bottleneck design for enhanced efficiency. The network begins with a 4 x 4 patchify non overlapping

convolution followed by many ConvNeXt blocks stacked one by one and ends with a global average pool layer and a final linear classification head [36]. The design of ConvNeXt-Base is illustrated in Figure 5.

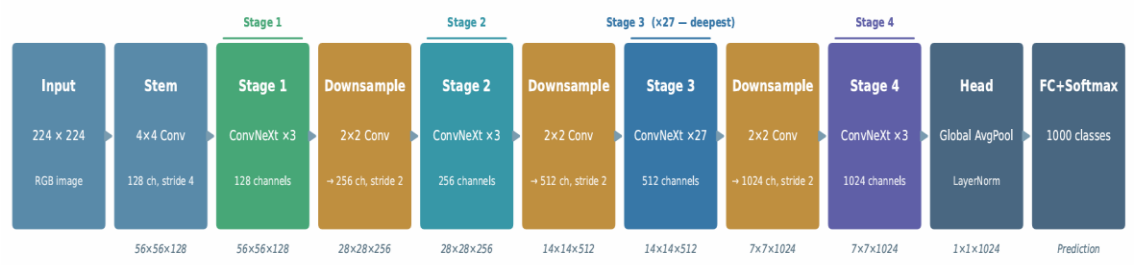

Fig. 5: Architecture of ConvNeXt-Base (Classes are 1000 as it is trained on ImageNet dataset, after fine-tuning to current dataset 5 classes are there)

3.3.4 Custom CNN

The custom CNN named ResidualCNN_CBAM was designed as a residual CNN with CBAM [37] to improve the discriminative feature learning process for cervical cytology. Architecture begins with a convolutional stem having 64 kernels of 7 x 7 size. Batch normalization, ReLU activation, and max pooling layers come after the stem. This stage helps in extracting basic features of an image like edges, textures etc.

After stem stage, there are three residual layers with progressively increasing feature channels. The number of channels increases from 64 to 128 in the first layer and it has three residual blocks. The second layer increases the channels from 128 to 256 and has three residual blocks. The final and third layer increases channels from 256 to 512 with two residual blocks. The first residual block of each layer has the function of down sampling and remaining blocks preserve the dimensions. This architecture helps in learning complex morphological features of cervical cells.

Each residual block is made of two 3 x 3 convolutions with batch normalization and activation by ReLU after first convolution and after residual addition. A shortcut connection is used to add original input to transformed output to maintain gradient flow. When spatial dimensions change, a 1 x 1 convolution with batch normalization is used to fix the mismatch of dimensions before and after addition.

CBAM is used in each residual block,z having both channel and spatial attention. The channel attention utilizes average pool and max pool which are passed through multi-layer perceptron and activated by sigmoid activation to give important feature channels. The spatial attention uses average and max pooled feature maps across channel dimension, concatenates them, applies 7 x 7 convolution and activate using sigmoid function to give important spatial regions. The attention provided by CBAM module helps focus on relevant areas of the cervical cell image for better classification.

Finally, after residual-attention stages is the adaptive average pool to reduce the number of feature maps to a 512-feature vector. This vector is passed through classifier having a linear layer with 256 neurons, ReLU activation, dropout with value 0.7 for regularization and reducing overfitting and a final linear classification head with five classes of SIPaKMeD dataset. Figure 6 represents the architecture of Custom CNN with residual blocks and CBAM.

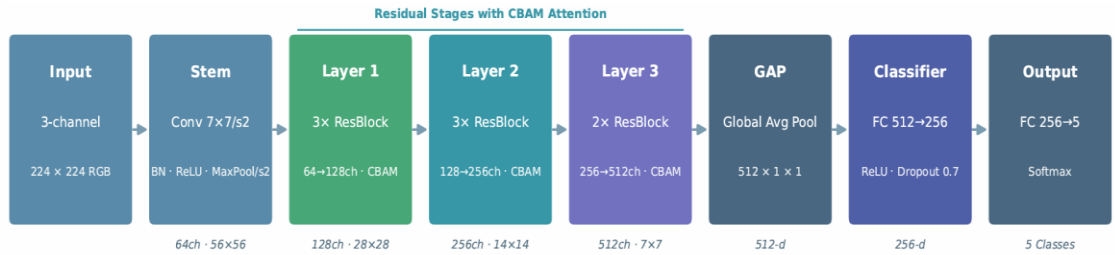

Fig. 6: Architecture of Custom CNN

3.4 Proposed Ensemble

To make the process of cervical cancer cell classification more robust, the predictions of individual base learners were combined into one to give an ensemble model. The ensemble consists of four trained models - ResNet152, EfficientNetV2-S, ConvNeXt, and Custom CNN. Each base learner has its own strengths depending on the architecture so they capture different features and helps in making the final decision while reducing any errors from any of the individual model.

In the proposed ensemble, element-wise averaging method was implemented. For a given input image, each model gives an output vector (logits). The logits obtained from each model were combined by averaging technique to give final prediction, thus ensuring equal contribution from individual models.

Mathematically, averaging technique for the ensemble can be represented as follows:

$$z_{final} = \frac{1}{4}\sum_{i=1}^{4} z_i$$

where z_{final} is the final output of the ensemble and z_i is the output from i^{th} base learner, with $i \in \{1, 2, 3, 4\}$.

The final prediction for an input then becomes the class label corresponding to the maximum value in the logit vector obtained from the ensemble model. The averaging technique was chosen for ensemble as it is simple to implement but highly effective in function. Thus, it helps improve performance of classification by diminishing

individual model biases. Figure 7 illustrates the design of proposed Ensemble Model and how the ensemble is made to do the classification task.

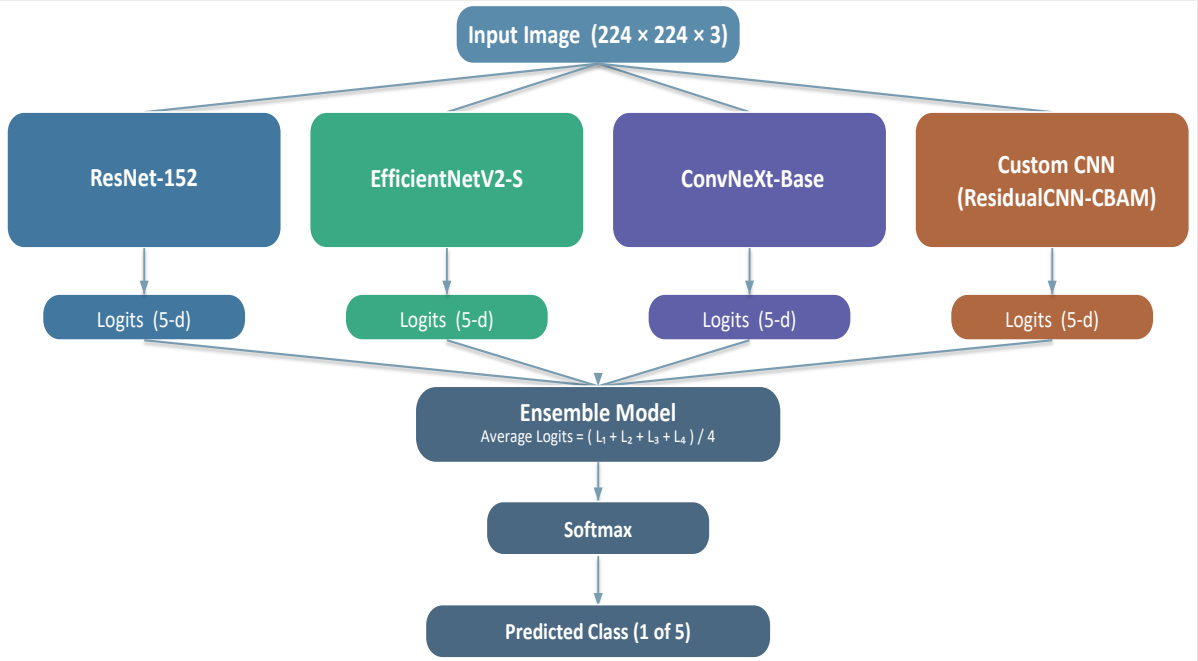

Fig. 7: Architecture of Proposed Ensemble Model

3.5 Evaluation Metrics

To assess how well the ensemble and individual base learners classified cervical cells in the SIPaKMeD dataset, some quantitative metrics were used. To calculate these metrics, TP (True Positives), TN (True Negatives), FP (False Positives), and FN (False Negatives) need to be defined. TP means a sample that is of positive class and classified correctly as positive. TN denotes the classification of a negative class sample as positive. FP denotes a negative class sample that has been categorized as positive. FN indicates a sample of positive class being classified as negative. In the present research, following evaluation metrics were used.

Accuracy

Accuracy is the main metric to judge the correctness of a model. It is described as the ratio of all true predictions to all of the model's predictions. This is how it is computed:

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

Apart from accuracy, precision, recall, and F1-score were also computed to analyse model performance better. They were computed using macro-averaging approach as it ensures that all the classes are equally treated without any bias as each class has its own significance in medical domain.

Precision

Precision is a metric which evaluates the actual positives from the total predicted positives by the model. It is computed by given formula:

$$Precision = \frac{TP}{TP + FP}$$

Recall

Recall also called sensitivity gives a ratio of true positive classification to the actual number of samples belonging to positive class. This metric is important for medical domain where minority/disease classes should be looked upon carefully. It is computed using following formula:

$$Recall = \frac{TP}{TP + FN}$$

F1-score

The harmonic mean of recall and precision is the F1 score. It is calculated as follows and provides a fair assessment of the model's categorization performance:

$$F1 - score = 2 * \frac{Precision\ .\ Recall}{Precision + Recall}$$

In addition to above stated metrics, confusion matrix was also used for performance analysis. It provides a report of misclassifications done by the model for different cervical cancer cell classes and tells which classes may look similar for model to distinguish between them and get confused. Apart from confusion matrix, loss and accuracy plots for train and validation datasets were also plotted for comparison.

To further evaluate the model, ROC and AUC curve were also computed for the ensemble model for each class using one-vs-rest approach. It helps determine how well a model can differentiate between different classes at different classification thresholds. The value of AUC closer to 1 indicates better classification.

3.6 Implementation Details

PyTorch framework was used to carry out all the experiments, along with few other libraries including OpenCV for image processing and Albumentations for data augmentation. The training and evaluation of models was done using Python based environment. For visualisation, libraries like Matplotlib and Seaborn were used for generating visually appealing plots and graphs for better understanding and clarity of the results.

The batch size of 32 was used for processing the dataset for batch wise training. ResNet152 was trained for 30 epochs. ConvNeXt was trained for 40 epochs and both EfficientNetV2-S and Custom CNN were trained for 50 epochs. The best model weights were saved for each model based on the validation performance.

Different optimizers were used for the process of optimization. For all the three pre-trained models, Adam optimizer was used for better learning rates and faster convergence. The AdamW optimizer was selected for Custom CNN because it separates the gradient update process from weight decay, improving regularization, stable training, and generalization.

Custom CNN was first trained with a learning rate of 0.001, and when the validation set's performance plateaued, ReduceLROnPlateau, a learning rate scheduler, was used to lower the learning rate. A learning rate of $1e^{-5}$ was used for ResNet152 and for EfficientNetV2-S and ConvNeXt lr of $1e^{-4}$ was used.

Since the cross-entropy loss function is appropriate for multi-class classification problems, it was utilized in all models. Label smoothing (0.1) was used to reduce the overconfidence of the model and for improving generalization. Class weights were also employed in the loss function calculation to account for any imbalance of minor class in the dataset.

The training of all the models was done using the same train, validation and test split to ensure uniformity and fair comparison. The experiments were run on a system enabled with GPU to make training and evaluation faster and easier, so that computationally intensive tasks can be handled easily.

CHAPTER 4

RESULTS AND DISCUSSION

In order to verify the robustness of the suggested approach, this section shows the results of the research carried out during the study and compares them with the current approaches found in the literature.

The evaluation metrics for ResNet152, EfficientNetV2-S, ConvNeXt-Base, Custom CNN and Proposed Ensemble are given in Table 2.

Table 2: Four base learners (ResNet152, EfficientNetV2-S, ConvNeXt-Base, Custom CNN) and the proposed ensemble's accuracy, precision, recall, and F1-score

Model	**Accuracy (%)**	**Precision**	**Recall**	**F1-Score**
ResNet152	96.71	0.9674	0.9673	0.9672
EfficientNetV2-S	95.56	0.9562	0.9560	0.9555
ConvNeXt-Base	94.57	0.9455	0.9460	0.9453
Custom CNN	90.62	0.9102	0.9071	0.9059
Proposed Ensemble	**97.53**	**0.9758**	**0.9757**	**0.9751**

The results of the experiment point to the fact that proposed ensemble achieved the best performance compared to base learners across all the evaluation metrics. It means combining different models with their specific strengths and architectures help improve the ensemble's performance.

The individual base learners demonstrated robust performances on their own. Among them, the strongest was ResNet152 followed by EfficientNetV2-S. The ResNet's architecture of deep residual network helped it achieve competitive performance. ConvNeXt's architecture inspired from vision transformers, helped it achieve decent performance. Custom CNN, although simpler compared to pre-trained architectures performed well by incorporating attention mechanism for specific task.

The ensemble was able to achieve such great performance due to the heterogeneity of the individual learners used. Each model captures different aspects of the cervical cell image. Thus, after averaging the predictions of different models, the bias and variance are reduced in the ensemble model showcasing better performance.

Accuracy during training and validation and loss plots also helps a lot in the performance analysis of the model. They record how a model is performing on the train dataset and then how its performance changes on validation dataset which is usually unseen by the model.

Train accuracy initially is low as the model is looking at the dataset for the first time, but as training progresses, it should usually increase with the epochs. The model's predictions on the unobserved validation dataset are represented as validation accuracy. It shows how well a model is generalizing to the data that was not seen by the model. It also helps analyse whether a model is overfitting or not i.e. is it learning the patterns or memorizing the data. If the training accuracy increases and the accuracy of the validation set is significantly less than the train accuracy, it shows overfitting.

The loss is a metric which measures how far is the prediction from true value. Initially, training loss is high because the model is looking at the data for the first time, gradually with more number of epochs loss decreases as the model starts learning the relevant patterns in the samples. Validation loss is the error made by the model on unseen validation dataset. This helps in assessing model's generalizability and performance on real world data. Overfitting of the model happens when training loss is reducing but loss of the validation set is still increasing. So, both training and validation losses need to be tracked properly for a good model performance for its applicability in the real-world.

The accuracy and loss charts for ResNet152 during training and validation are displayed in Figure 8. For EfficientNetV2-S, similar plots are shown in Figure 9 while for ConvNeXt-Base and Custom CNN, accuracy and loss plots are illustrated in Figure 10 and Figure 11 respectively.

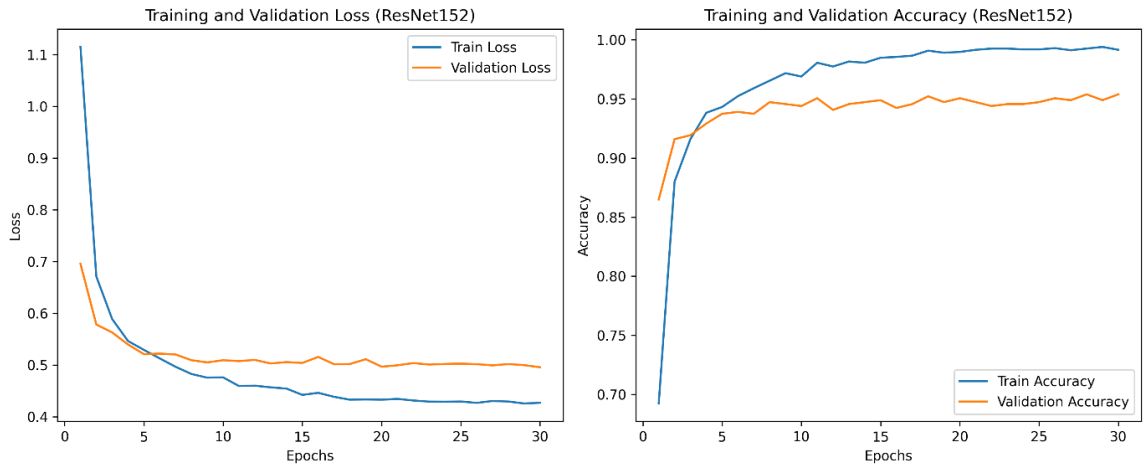

Fig. 8: ResNet152 training and validation loss and accuracy plot

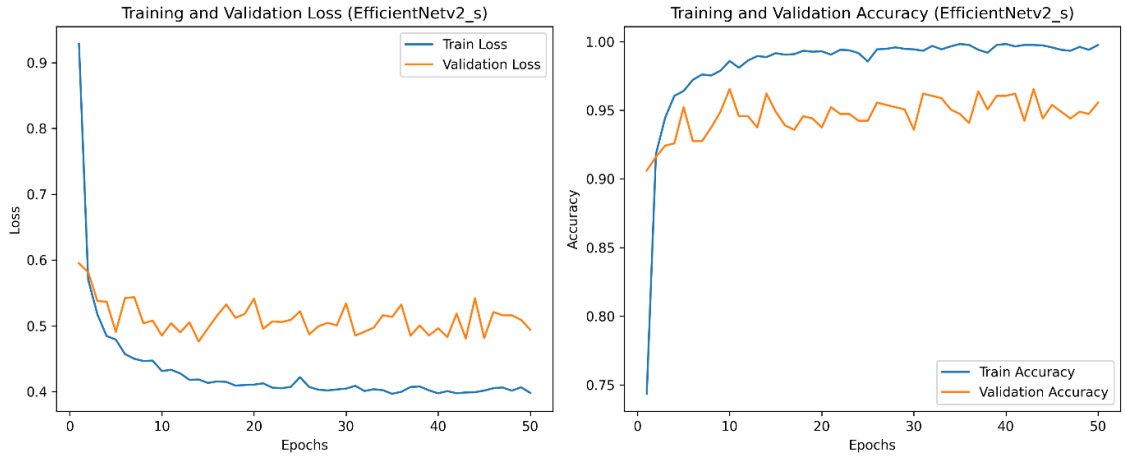

Fig. 9: EfficientNetV2-S training and validation loss and accuracy plot

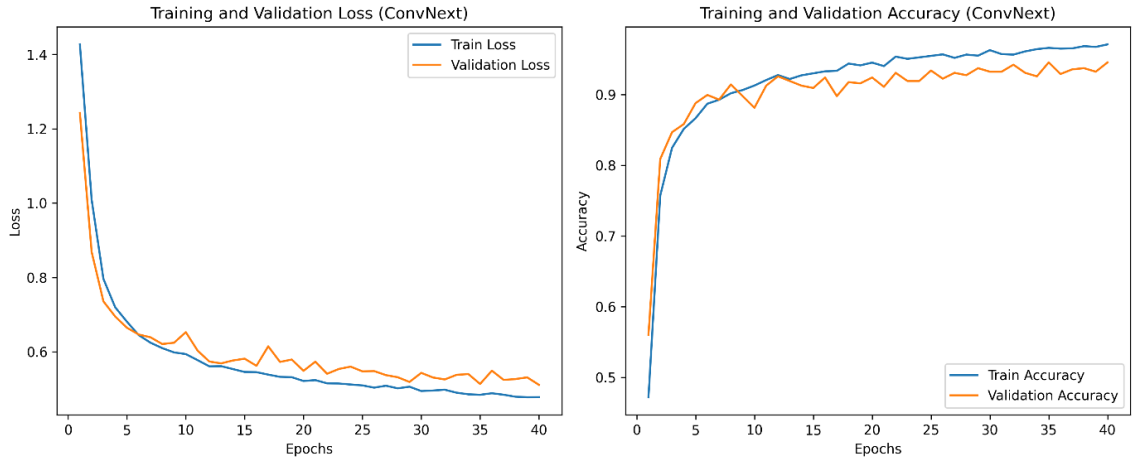

Fig. 10: ConvNeXt-Base training and validation loss and accuracy plot

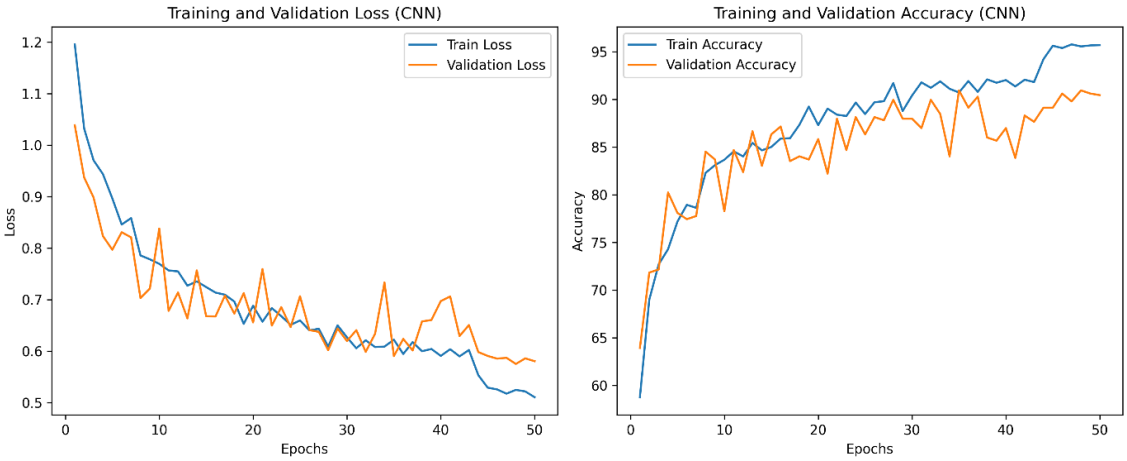

Fig. 11: Custom CNN's training and validation loss and accuracy plot

Confusion matrix helps in assessing the misclassifications done by the model. It helps understand which classes have similar morphological features and the model is not able to distinguish between them properly and is confused about which class to assign to a sample. Confusion matrix for ResNet152 is presented in Figure 12. Figure 13 shows confusion matrix for EfficientNetV2-S, similarly confusion matrix for ConvNeXt-Base, Custom CNN and ensemble are also shown in Figure 14, Figure 15 and Figure 16 respectively.

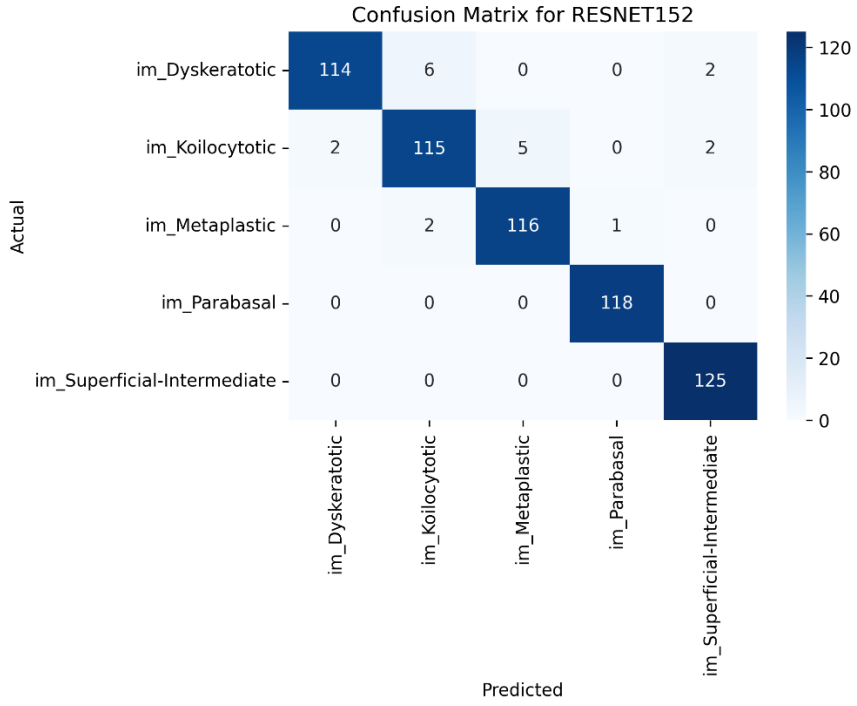

Fig. 12: Confusion Matrix for ResNet152

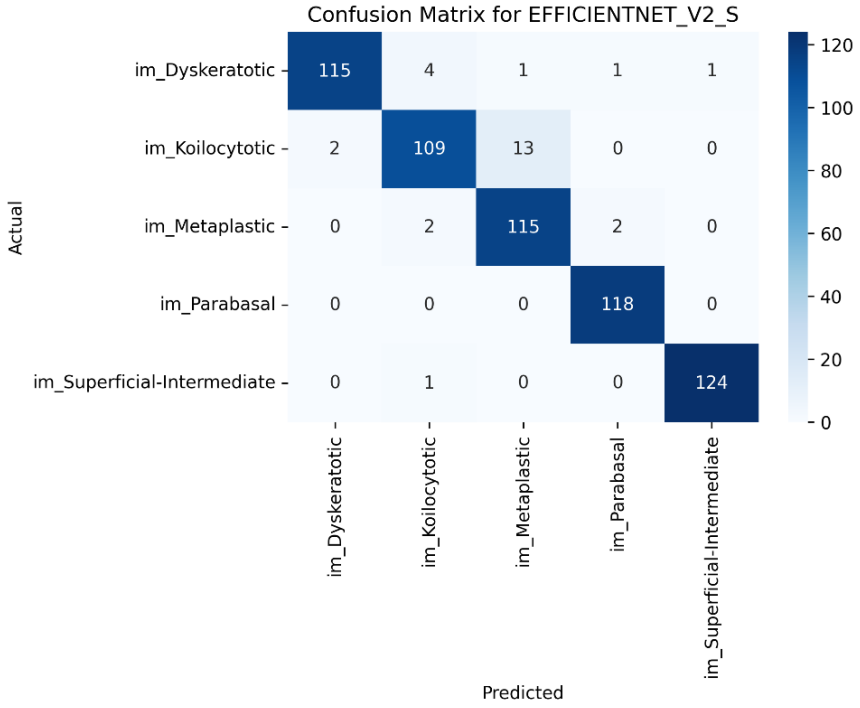

Fig. 13: EfficientNetV2-S Confusion Matrix

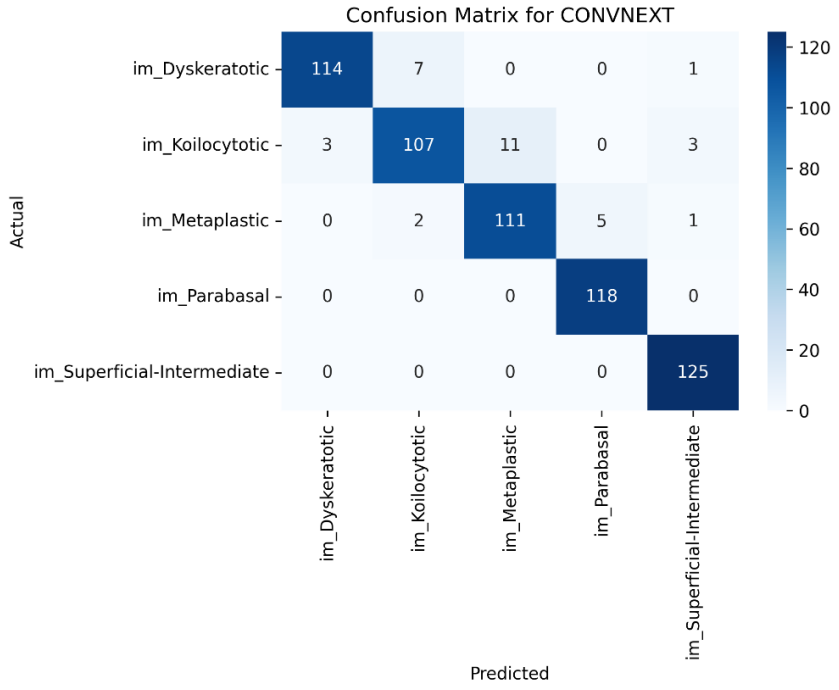

Fig. 14: Confusion Matrix for ConvNeXt-Base

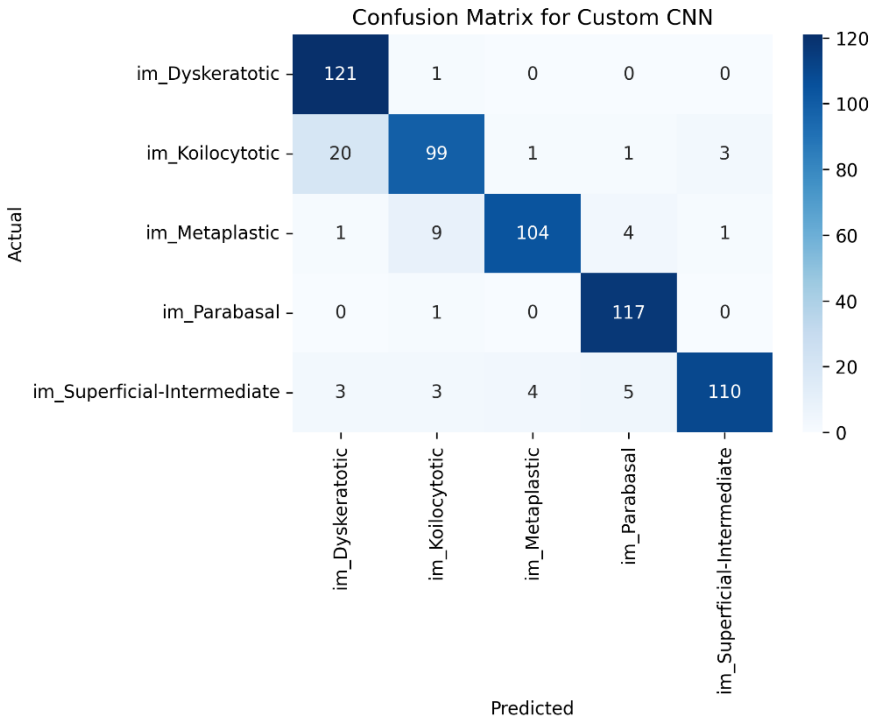

Fig. 15: Custom CNN's Confusion Matrix

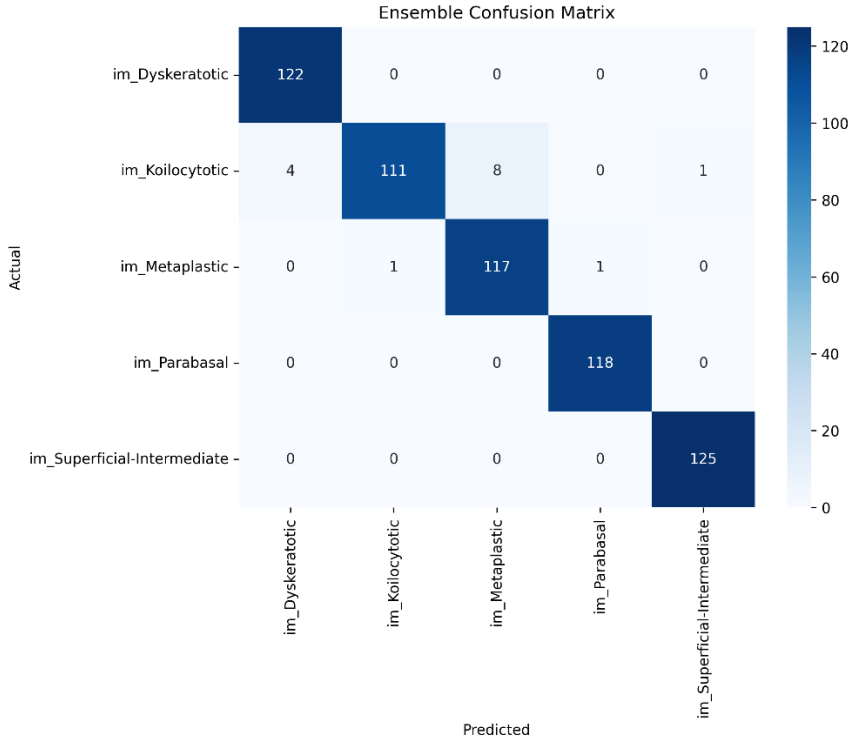

Fig. 16: Ensemble Model's Confusion Matrix

From the confusion matrix of different models and the final ensemble model, it is visible that class 'Koilocytotic' is misclassified most of the times by every model to some other class, usually Metaplastic or Dyskeratotic. It means model finds it difficult to learn distinguishing features of Koilocytotic class properly.

ROC-AUC curve for ensemble is given in Figure 17. It displays the True Positive Rate versus False Positive Rate per class (one vs rest). AUC of 1 means perfect classification. From the figure and confusion matrix of ensemble model, it is clear that classes 'Superficial-Intermediate', 'Parabasal' and 'Dyskeratotic' are perfectly classified while classes 'Metaplastic' and 'Koilocytotic' have few misclassifications, decreasing their AUC.

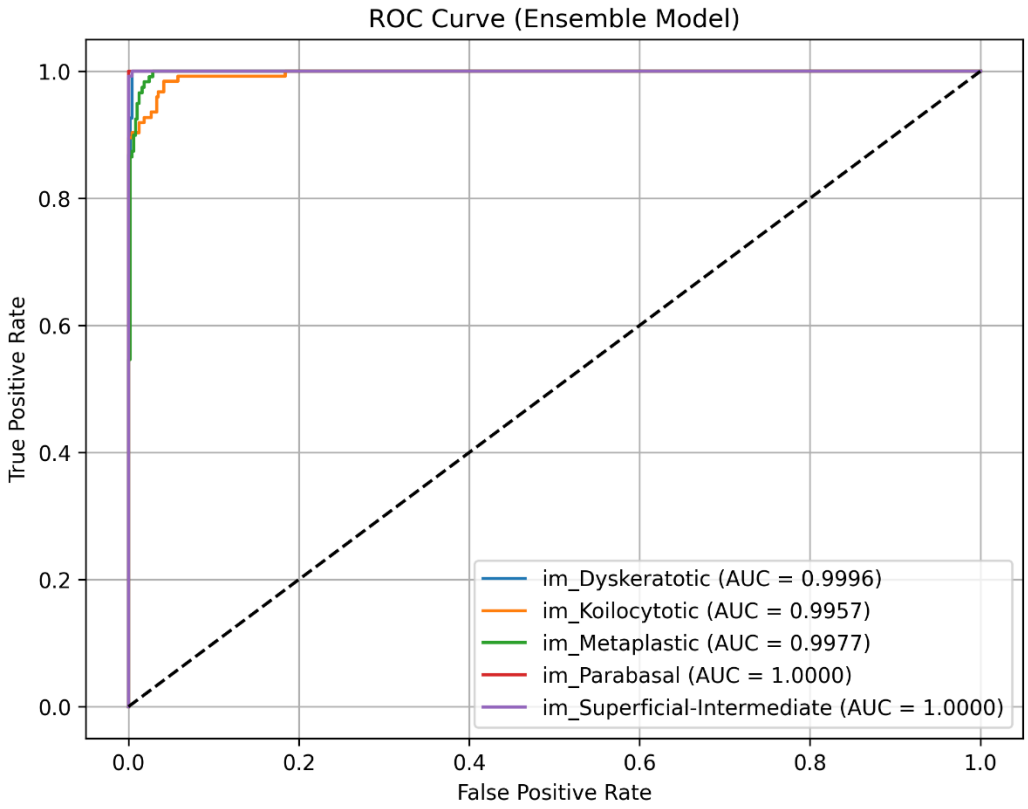

Fig. 17: ROC-AUC Curve for the Proposed Ensemble

A comparison of the proposed ensemble with other models from the literature for the classification of cervical cancer cells is shown in Table 3. Gangrade et al. [12] proposed a novel ensemble with CNN, AlexNet and SqueezeNet that produced a 94% classification accuracy for cervical cancer. Manna et al. [13] created an ensemble with Xception, DenseNet-169, and Inception V3, utilizing fuzzy rank-based fusion of the separate learners. The accuracy was 95.43% on SIPaKMeD's 5-class classification task. Pramanik et al. [25] provided a fuzzy distance-based ensemble of Inception V3,

MobileNet V2, and Inception ResNet V2 and some additional layers in each model for learning specific features of the given dataset. This proposed model achieved a competitive accuracy of 96.96% compared to base learners. Sahoo et al. [38] gave a method combining three pre-trained CNNs into an ensemble with some advanced data augmentation techniques. Kudva et al. [39] demonstrated a method using hybrid transfer learning technique. In this model, they trained a CNN from scratch using initial weights as those identified from pre-trained CNNs filters' which were relevant for the task of classification that need to be performed. The proposed ensemble gave an accuracy of 97.18% on the SIPaKMeD dataset.

Table 3: A comparison of the proposed ensemble model's performance with a few previous research studies

Method	**Dataset**	**Performance**	**Reference**
Ensemble of CNN, AlexNet and SqueezeNet	SIPaKMeD	94% accuracy	[12]
Fuzzy rank-based fusion of Inception V3, Xception, and DenseNet-169 ensemble	SIPaKMeD	95.43% accuracy (5-class)	[13]
Inception V3, MobileNet V2, and Inception ResNet V2 fuzzy distance-based ensemble	SIPaKMeD	96.96% accuracy	[25]
Ensemble with advanced data augmentation	SIPaKMeD	97.18% accuracy	[38]
Novel Hybrid Transfer Learning Technique	Cervical cell images	91.46% accuracy	[39]
CASPNet	SIPaKMeD	97.07%	[40]
Proposed Model	**SIPaKMeD**	**97.53%**	**-**

Discussion

In order to improve performance and automate the process, the current study presents a method for classifying cervical cell images as either cancerous or non-cancerous utilizing contemporary deep learning frameworks. The idea of combining different CNN architectures into an ensemble gave good results as can be seen through ensemble performance. The performance of both base learners and ensemble was good, thus adding to the reliability of the approach. For its use in real world settings, more deep research with good quality real world data is required for deployment.

Though the study conducted shows strong performance, certain limitations remain which need to be looked upon in the upcoming studies. The size of the dataset used was limited which may not be a true representation of diversity in the real-world data, thus impairing the model's capacity to effectively generalize to actual clinical

scenarios. Therefore, more varied data is needed to train DL models. The architectures used for base learners were quite heavy for real world deployment. Thus, focus should be on deployment of lightweight models without compromising the model's performance. The ensemble technique used was equal weight averaging which is simple and may not represent strengths of individual models properly, so advanced techniques should be explored for improvements.

CHAPTER 5
CONCLUSION AND FUTURE SCOPE

The current study used the SIPaKMeD dataset to present an ensemble model based on deep learning techniques for the categorization of cervical cancer cells. The approach combined four heterogenous base learners – ResNet152, EfficientNetV2-S, ConvNeXt and Custom CNN with attention mechanism (CBAM) to form an ensemble enhancing the performance of overall model.

The ensemble was created using the averaging technique, which simply averages the predictions from various distinct models to get a single final output. The results reveal that ensemble outperformed individual models across the measured evaluation metrics with an accuracy of 97.53%. Thus, it shows that combination of different models helps extract complementary features from the image, enhancing the classification performance.

Despite strong current approach, future works can have more improvements. Dataset limitation can be overcome to a certain extent by use of synthetic image generation [41]. For better image classification transformer-based architectures should be incorporated. A proper separate real world cervical cancer dataset should be used for validation. For deployment in real world, such models should be built which are lightweight without compromising on the model's efficiency and accuracy [42]. The models used should not be a black box, so interpretability should be added to model predictions using explainable AI techniques for making the system more robust for deployment in real world clinical settings [43].

REFERENCES

[1] M. Kalbhor, S. Shinde, P. Wajire, and H. Jude, "CerviCell-detector: An object detection approach for identifying the cancerous cells in pap smear images of cervical cancer," *Heliyon*, vol. 9, no. 11, p. e22324, Nov. 2023, doi: 10.1016/j.heliyon.2023.e22324.

[2] International Agency for Research on Cancer, "Cancer Today (GLOBOCAN 2022)," International Agency for Research on Cancer (IARC). Accessed: Mar. 09, 2026. [Online]. Available: https://gco.iarc.fr/today

[3] D. Saslow *et al.*, "American Cancer Society, American Society for Colposcopy and Cervical Pathology, and American Society for Clinical Pathology Screening Guidelines for the Prevention and Early Detection of Cervical Cancer," *Am. J. Clin. Pathol.*, vol. 137, no. 4, pp. 516–542, Apr. 2012, doi: 10.1309/AJCPTGD94EVRSJCG.

[4] G. Bogani *et al.*, "HPV-related lesions after hysterectomy for high-grade cervical intraepithelial neoplasia and early-stage cervical cancer: A focus on the potential role of vaccination," *Tumori Journal*, vol. 110, no. 2, pp. 139–145, Apr. 2024, doi: 10.1177/03008916231208344.

[5] B. S. Deo, M. Pal, P. K. Panigarhi, and A. Pradhan, "CerviFormer: A Pap-smear based cervical cancer classification method using cross attention and latent transformer," Mar. 2023.

[6] N. Merlina, A. Prasetio, I. Zuniarti, N. A. Mayangky, D. N. Sulistyowati, and F. Aziz, "Deep CNN Models for Detecting Cervical Cancer in Pap Smear Images," *TEM Journal*, pp. 1073–1083, May 2025, doi: 10.18421/TEM142-09.

[7] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017, doi: 10.1016/j.media.2017.07.005.

[8] R. Lozano, "Comparison of computer-assisted and manual screening of cervical cytology," *Gynecol. Oncol.*, vol. 104, no. 1, pp. 134–138, Jan. 2007, doi: 10.1016/j.ygyno.2006.07.025.

[9] M. A. Raza *et al.*, "Advanced Feature Extraction for Cervical Cancer Image Classification: Integrating Neural Feature Extraction and AutoInt Models," *Sensors*, vol. 25, no. 9, p. 2826, Apr. 2025, doi: 10.3390/s25092826.

[10] A. Hanzala, T. Akter, and Md. S. Rahman, "A hybrid approach for cervical cancer detection: Combining D-CNN, transfer learning, and ensemble models," *Array*, vol. 27, p. 100434, Sep. 2025, doi: 10.1016/j.array.2025.100434.

[11] H. Kaur, R. Sharma, and J. Kaur, "Comparison of deep transfer learning models for classification of cervical cancer from pap smear images," *Sci. Rep.*, vol. 15, no. 1, p. 3945, Jan. 2025, doi: 10.1038/s41598-024-74531-0.

[12] J. Gangrade, R. Kuthiala, S. Gangrade, Y. P. Singh, M. R, and S. Solanki, "A deep ensemble learning approach for squamous cell classification in cervical cancer," *Sci. Rep.*, vol. 15, no. 1, p. 7266, Mar. 2025, doi: 10.1038/s41598-025-91786-3.

[13] A. Manna, R. Kundu, D. Kaplun, A. Sinitca, and R. Sarkar, "A fuzzy rank-based ensemble of CNN models for classification of cervical cytology," *Sci. Rep.*, vol. 11, no. 1, p. 14538, Jul. 2021, doi: 10.1038/s41598-021-93783-8.

[14] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Classification of cervical cells from the Pap smear image using the RES_DCGAN data augmentation and ResNet50V2 with self-attention architecture," *Neural Comput. Appl.*, vol. 36, no. 34, pp. 21801–21815, Dec. 2024, doi: 10.1007/s00521-024-10404-x.

[15] S. Nayar, D. Priya, and V. P. J, "Deep Learning Approach for Automated Data Augmentation and Multi-class Classification of Pap Smear Images," *Procedia Comput. Sci.*, vol. 235, pp. 2205–2214, 2024, doi: 10.1016/j.procs.2024.04.209.

[16] R. Austin and R. Parvathi, "CNN based method for classifying cervical cancer cells in pap smear images," *Sci. Rep.*, vol. 15, no. 1, p. 23936, Jul. 2025, doi: 10.1038/s41598-025-10009-x.

[17] R. Ahmed, N. Dahmani, G. Dahy, A. Darwish, and A. Ella Hassanien, "Early Detection and Categorization of Cervical Cancer Cells Using Smoothing Cross Entropy-Based Multi-Deep Transfer Learning," *IEEE Access*, vol. 12, pp. 157838–157853, 2024, doi: 10.1109/ACCESS.2024.3485888.

[18] F. Hashem *et al.*, "Exploring Data Imbalance Challenges in Cervical Cancer Detection Using Advanced Deep Learning Models," in *2024 International Conference on Smart Systems and Power Management (IC2SPM)*, IEEE, Nov. 2024, pp. 95–101. doi: 10.1109/IC2SPM62723.2024.10841356.

[19] R. S. Akash, R. Islam, S. S. I. Badhon, and K. T. Hossain, "CerviXpert: A multi-structural convolutional neural network for predicting cervix type and cervical cell abnormalities," *Digit. Health*, vol. 10, Jan. 2024, doi: 10.1177/20552076241295440.

[20] M. Fang, X. Lei, B. Liao, and F.-X. Wu, "A Deep Neural Network for Cervical Cell Classification Based on Cytology Images," *IEEE Access*, vol. 10, pp. 130968–130980, 2022, doi: 10.1109/ACCESS.2022.3230280.

[21] Md. H. K. Mehedi, M. Khandaker, S. Ara, Md. A. Alam, M. F. Mridha, and Z. Aung, "A lightweight deep learning method to identify different types of cervical cancer," *Sci. Rep.*, vol. 14, no. 1, p. 29446, Nov. 2024, doi: 10.1038/s41598-024-79840-y.

[22] Y. Zhang, C. Ning, and W. Yang, "An automatic cervical cell classification model based on improved DenseNet121," *Sci. Rep.*, vol. 15, no. 1, p. 3240, Jan. 2025, doi: 10.1038/s41598-025-87953-1.

[23] N. K. Chauhan *et al.*, "A hybrid learning network with progressive resizing and PCA for diagnosis of cervical cancer on WSI slides," *Sci. Rep.*, vol. 15, no. 1, p. 12801, Apr. 2025, doi: 10.1038/s41598-025-97719-4.

[24] N. K. Chauhan, K. Singh, A. Kumar, and S. B. Kolambakar, "HDFCN: A Robust Hybrid Deep Network Based on Feature Concatenation for Cervical Cancer

Diagnosis on WSI Pap Smear Slides," *Biomed Res. Int.*, vol. 2023, no. 1, Jan. 2023, doi: 10.1155/2023/4214817.

[25] R. Pramanik, M. Biswas, S. Sen, L. A. de Souza Júnior, J. P. Papa, and R. Sarkar, "A fuzzy distance-based ensemble of deep models for cervical cancer detection," *Comput. Methods Programs Biomed.*, vol. 219, p. 106776, Jun. 2022, doi: 10.1016/j.cmpb.2022.106776.

[26] D. B. Talpur, A. R. Rang, A. Khowaja, and A. Shah, "DeepCervixNet: An Advanced Deep Learning Approach for Cervical Cancer Classification in Pap Smear Images", doi: 10.21015/vtcs.v12i1.1812.

[27] B. Z. Wubineh, A. Rusiecki, and K. Halawa, "Deep learning-based automatic segmentation and classification for cervical cancer detection using an improved U-Net and ensemble methods," *Sci. Rep.*, vol. 16, no. 1, p. 5184, Jan. 2026, doi: 10.1038/s41598-026-35299-7.

[28] R. Singh, H. Kaur, and J. Malhotra, "Cervical cancer detection through Pap smear images using hybrid deep feature extraction and ensemble machine learning," *Multiscale and Multidisciplinary Modeling, Experiments and Design*, vol. 8, no. 5, p. 233, May 2025, doi: 10.1007/s41939-025-00834-y.

[29] S. Muksimova, S. Umirzakova, S. Kang, and Y. I. Cho, "CerviLearnNet: Advancing cervical cancer diagnosis with reinforcement learning-enhanced convolutional networks," *Heliyon*, vol. 10, no. 9, p. e29913, May 2024, doi: 10.1016/j.heliyon.2024.e29913.

[30] I. Pacal and S. Kılıcarslan, "Deep learning-based approaches for robust classification of cervical cancer," *Neural Comput. Appl.*, vol. 35, no. 25, pp. 18813–18828, Sep. 2023, doi: 10.1007/s00521-023-08757-w.

[31] N. M. Fahad, S. Azam, S. Montaha, and Md. S. H. Mukta, "Enhancing cervical cancer diagnosis with graph convolution network: AI-powered segmentation, feature analysis, and classification for early detection," *Multimed. Tools Appl.*, vol. 83, no. 30, pp. 75343–75367, Feb. 2024, doi: 10.1007/s11042-024-18608-y.

[32] V. S. Jadhav, R. Yakkundimath, G. Saunshi, and G. Konnurmath, "Deep Learning Based Classification of Cervical Cancer Stages Using Transfer Learning Models," *Indian Journal of Medical and Paediatric Oncology*, vol. 46, no. 05, pp. 480–488, Oct. 2025, doi: 10.1055/s-0045-1809907.

[33] G. Verma and A. Barthwal, "Attention-enhanced deep learning for cervical cytology: combining convolutional networks with multi-head attention and fuzzy logic," *Pol. J. Radiol.*, vol. 90, pp. 414–430, Aug. 2025, doi: 10.5114/pjr/207475.

[34] S. K. Mathivanan, D. Francis, S. Srinivasan, V. Khatavkar, K. P, and M. A. Shah, "Enhancing cervical cancer detection and robust classification through a fusion of deep learning models," *Sci. Rep.*, vol. 14, no. 1, p. 10812, May 2024, doi: 10.1038/s41598-024-61063-w.

[35] M. Tan and Q. V. Le, "EfficientNetV2: Smaller Models and Faster Training," Jun. 2021.

[36] Z. Liu, H. Mao, C.-Y. Wu, C. Feichtenhofer, T. Darrell, and S. Xie, "A ConvNet for the 2020s," Mar. 2022.

[37] Z. Cömert, F. Efil, and M. Türkoğlu, "Convolutional Block Attention Module and Parallel Branch Architectures for Cervical Cell Classification," *Int. J. Imaging Syst. Technol.*, vol. 35, no. 2, Mar. 2025, doi: 10.1002/ima.70048.

[38] P. Sahoo, S. Saha, S. Mondal, M. Seera, S. K. Sharma, and M. Kumar, "Enhancing Computer-Aided Cervical Cancer Detection Using a Novel Fuzzy Rank-Based Fusion," *IEEE Access*, vol. 11, pp. 145281–145294, 2023, doi: 10.1109/ACCESS.2023.3346764.

[39] V. Kudva, K. Prasad, and S. Guruvare, "Hybrid Transfer Learning for Classification of Uterine Cervix Images for Cervical Cancer Screening," *J. Digit. Imaging*, vol. 33, no. 3, pp. 619–631, Jun. 2020, doi: 10.1007/s10278-019-00269-1.

[40] J. Mondal, R. Chatterjee, M. Kumar Gourisaria, M. Sahni, and E. León-Castro, "Cervical cancer classification using a novel hybrid approach," *Front. Oncol.*, vol. 15, Dec. 2025, doi: 10.3389/fonc.2025.1703772.

[41] G. Shandilya *et al.*, "Enhancing advanced cervical cell categorization with cluster-based intelligent systems by a novel integrated CNN approach with skip mechanisms and GAN-based augmentation," *Sci. Rep.*, vol. 14, no. 1, p. 29040, Nov. 2024, doi: 10.1038/s41598-024-80260-1.

[42] O. Attallah, "Light-XAI: a CADx for explainable cervical cancer detection via attention-based lightweight convolutional neural networks and layer-wise feature fusion," *BioData Min.*, vol. 19, no. 1, p. 26, Apr. 2026, doi: 10.1186/s13040-026-00540-6.

[43] M. I. H. Siddiqui *et al.*, "Accelerated and accurate cervical cancer diagnosis using a novel stacking ensemble method with explainable AI," *Inform. Med. Unlocked*, vol. 56, p. 101657, 2025, doi: 10.1016/j.imu.2025.101657.

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis 'Hybrid Ensemble of Transfer Learning and Custom CNN Architectures for Automated Classification of Cervical Cancer Cells from Pap Smear Images' Total Pages 42 Name of the Scholar: Tanu Paliwal (24/MSCBIO/13).

Supervisor

Prof. Yasha Hasija

Professor and Head of Department

Department of Biotechnology

Delhi Technological University

Department of Biotechnology

This is to report that the above thesis was scanned for similarity detection. Process and outcome are given below:

Software used: Turnitin, Similarity Index: 12%, Total Word Count: 9150

Date:

Candidate's Signature **Signature of Supervisor**

Tanu_Thesis_WC

ORIGINALITY REPORT

12%	8%	10%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	Anu V Kottath. "SWING EFFNET: A HYBRID FRAMEWORK FOR MONKEYPOX DISEASES DETECTION USING SWING TRANSFORMERS AND EFFICIENTNET-B0", International Journal of Applied Mathematics, 2025 Publication	1%
2	link.springer.com Internet Source	<1%
3	www.mdpi.com Internet Source	<1%
4	K. V. Sambasivarao, Anasuya Sesha Roopa Devi Bhima. "Artificial Intelligence, Computational Intelligence, and Inclusive Technologies - Proceedings of International Conference on Artificial Intelligence, Computational Intelligence, and Inclusive Technologies (ICRAIC2IT – 2025)", CRC Press, 2026 Publication	<1%
5	www.frontiersin.org Internet Source	<1%
6	Submitted to University of Malaya Student Paper	<1%
7	arxiv.org Internet Source	<1%
8	eprints.utar.edu.my Internet Source	<1%
9	assets-eu.researchsquare.com Internet Source	<1%
10	www.nature.com Internet Source	<1%
11	mdpi-res.com Internet Source	<1%
12	"Advances in Emerging Technologies and Computing Innovations", Springer Science and Business Media LLC, 2025 Publication	<1%
13	dspace.daffodilvarsity.edu.bd:8080 Internet Source	<1%
14	proceedings.neurips.cc Internet Source	<1%
15	"Advances in Data Science and Management", Springer Science and Business Media LLC, 2026 Publication	<1%
16	"Advances in Smart Computing and Applications", Springer Science and Business Media LLC, 2026 Publication	<1%
17	"Proceedings of Data Analytics and Management", Springer Science and Business Media LLC, 2025 Publication	<1%

18	Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Intelligent Computing and Communication Techniques - Volume 2", CRC Press, 2025 Publication	<1 %
19	www.preprints.org Internet Source	<1 %
20	Submitted to University of Hull Student Paper	<1 %
21	Submitted to University of Warwick Student Paper	<1 %
22	Submitted to Vardhaman College of Engineering, Hyderabad Student Paper	<1 %
23	harvest.usask.ca Internet Source	<1 %
24	www.researchgate.net Internet Source	<1 %
25	Submitted to Georgia Institute of Technology Main Campus Student Paper	<1 %
26	Mehdi Ghayoumi. "Generative Adversarial Networks in Practice", CRC Press, 2023 Publication	<1 %
27	essay.utwente.nl Internet Source	<1 %
28	www.ijirset.com Internet Source	<1 %
29	Submitted to Higher Education Commission Pakistan Student Paper	<1 %
30	Nitin Kumar Chauhan, Krishna Singh, Amit Kumar, Ashutosh Mishra et al. "A hybrid learning network with progressive resizing and PCA for diagnosis of cervical cancer on WSI slides", Scientific Reports, 2025 Publication	<1 %
31	dspace.dtu.ac.in:8080 Internet Source	<1 %
32	Arvind Dagur, Sohit Agarwal, Dhirendra Kumar Shukla, Shabir Ali, Sandhya Sharma. "Artificial Intelligence and Sustainable Innovation - Volume 3", CRC Press, 2026 Publication	<1 %
33	Omneya Attallah. "CerCan·Net: Cervical Cancer Classification Model via Multi-layer Feature Ensembles of Lightweight CNNs and Transfer Learning", Expert Systems with Applications, 2023 Publication	<1 %
34	Omneya Attallah. "CerCan·Net: Cervical cancer classification model via multi-layer feature ensembles of lightweight CNNs and transfer learning", Expert Systems with Applications, 2023 Publication	<1 %
35	S. S. Saranya, C. Santhanakrishnan, K. Pradeep Mohan Kumar, P. Anil Kumar, Bibhuti Bhusan Dash, Saroja Kumar Rout, Kanchan	<1 %

	Bala. "Advancements in fusion-based deep representation learning for enhanced cervical precancerous lesion classification using biomedical image analysis", Scientific Reports, 2025 Publication	
36	Sidrah Mumtaz, Mudassar Raza, Ofonime Okon, Saeed Rehman, Adham Ragab, Hafiz Rauf. "A Hybrid Framework for Detection and Analysis of Leaf Blight Using Guava Leaves Imaging", Agriculture, 2023 Publication	<1 %
37	Yixuan Xin, Qinghua Li, Dianyuan Li. "LGPNet: A dual-branch parallel network for cervical cell image classification", Biomedical Signal Processing and Control, 2026 Publication	<1 %
38	researchexchange.ocrahope.org Internet Source	<1 %
39	www.researchbank.ac.nz Internet Source	<1 %
40	"Proceedings of International Conference on Artificial Intelligence and Networks", Springer Science and Business Media LLC, 2026 Publication	<1 %
41	A. Naresh Kumar, S. Sageengran, Parikshit Narendra Mahalle, Loganathan Guganathan. "Cervical cancer image analysis: Detection and segmentation using self-guided quantum GANs and musical chairs optimization", Computers in Biology and Medicine, 2026 Publication	<1 %
42	K.P. Swain, Soumya Ranjan Nayak, Ayusee Swain, Gopinath Palai. "Automated classification of cervical cancer cells using deep learning models: A comparative analysis of CNN architectures", Elsevier BV, 2026 Publication	<1 %
43	Md Mamunur Rahaman, Chen Li, Xiangchen Wu, Yudong Yao, Zhijie Hu, Tao Jiang, Xiaoyan Li, Shouliang Qi. "A Survey for Cervical Cytopathology Image Analysis Using Deep Learning", IEEE Access, 2020 Publication	<1 %
44	P. Karthika, M. Premkumar. "An in-depth exploration of CNN-based deep learning models in cervical carcinoma analysis", Advances in Biomarker Sciences and Technology, 2026 Publication	<1 %
45	Pranab Sahoo, Jay Gala, Sriparna Saha, Sujit Chowdhury, Samrat Mondal. "PFSNet: A prototypical federated siamese network for few-shot cervical cancer detection", Biomedical Signal Processing and Control, 2026 Publication	<1 %
46	ebin.pub Internet Source	<1 %
47	eprints.qut.edu.au Internet Source	<1 %
	minds.wisconsin.edu	

CURRICULUM VITAE

Tanu Paliwal

paliwaltanu02@gmail.com | 7505506708 | Ghaziabad, UP | LinkedIn

EDUCATION

Master of Science in Biotechnology, *Delhi Technological University* — *2024 – 2026 | Delhi*
CGPA – 9.09/10.00

Minor in Data Science and Machine Learning, *IIT Mandi* — *2024 – 2025 | Mandi, HP*
CGPA – 8.60/10.00

Bachelor of Science in Zoology, *University of Delhi* — *2021 – 2024 | Delhi*
CGPA – 9.365/10.00

PROJECTS

Automated Waste Classification | Link *PyTorch, Scikit-learn, NumPy, Matplotlib*

- Fine-tuned pre-trained CNNs on a multi-class waste image dataset, achieving 97% classification accuracy - outperforming baseline models after targeted preprocessing.
- Engineered full preprocessing pipeline (augmentation, normalisation, class-balancing) to reduce overfitting and improve real-world generalisability.
- Delivered an end-to-end web interface that returns waste category prediction with tailored recycling guidance, making the model immediately usable by non-technical users.

SKILLS

Programming & Data: Python, Pandas, NumPy, Matplotlib, Seaborn
ML / Deep Learning: PyTorch, Scikit-learn, CNN fine-tuning
Wet Lab: Mammalian Cell Culture, GLP Compliance, DNA Extraction, HPLC
Bioinformatics: Primer Design, Homology Modelling, MSA

EXPERIENCE

Research Intern, *Bosch Global Software Technologies, Bengaluru* — *01/2026 – 06/2026*

- Architected a multi-modal biological age estimation framework fusing wearables, lab biomarkers, and imaging data - reducing dependency on single-modality inputs and increasing robustness.
- Engineered deviation-scoring components using reference-based quantile regression on healthy cohorts, enabling per-biomarker attribution of age acceleration with interpretable outputs for clinical review.
- Modelled system-level aging trajectories across cardiovascular, metabolic, and musculoskeletal systems.

Animal Cell Culture Training, *Dabur Research Foundation* — *06/2025 – 08/2025*

- Independently executed sub-culturing, cryopreservation, and cell-line revival protocols under GLP conditions, developing a working foundation in mammalian cell biology.
- Gained hands-on proficiency with sterile technique, biosafety cabinet operation, and quality control checkpoints relevant to pharmaceutical R&D environments.

ACHIEVEMENTS

National Competitive Exams
GATE Biotechnology 2024 (Qualified) | IIT JAM 2024 (AIR 113)

Academic Awards

- Proficiency Prize – highest marks across all years of B.Sc.; highest marks in 1st and 3rd year of B.Sc.
- Instituted Prize – highest marks in 1st and 2nd year of B.Sc.