Combined Deep Learning Technique to detect Cervical Cancer accurately and as early as possible

[1]Sanjay S, [2]Tarun Tejasvi V P, [3]Tharun J S, [4]Bavana O,
[5]Lekha S, [6]Dr. Kharmega Sundararaj G
Department of Computer Science and Engineering
Dr.T. Thimmaiah Institute of Technology, KGF, Karnataka - 563120

Abstract: In this paper, we introduce an advanced intelligent system designed to improve cervical cell classification for cytology-based diagnoses. The system combines the power of deep learning, specifically transfer learning, with traditional machine learning techniques to enhance diagnostic accuracy. By leveraging the SIPaKMeD dataset, we experiment with several well-established deep convolutional neural networks (CNNs), such as ResNet101, ResNet152, InceptionV3, and AlexNet, for feature extraction. These models are renowned for their ability to capture complex patterns in medical imaging, which we use to extract meaningful features from cervical cell images. For the classification step, we integrate a variety of machine learning algorithms, including Softmax Logistic Regression, Random Forests, Naive Bayes, and Principal Component Analysis, allowing the system to predict cell types with high precision. To make the process even more user-friendly and accessible, we deploy the entire system through an interactive web application built with Streamlit. This allows clinicians to upload microscopic images of cervical cells and receive accurate cell type predictions, along with confidence scores that indicate the reliability of the results. Our proposed hybrid architecture has shown impressive performance, particularly when

Keywords: Cervical Cancer, Deep Learning, Transfer Learning, Cytology Image Classification, Ensemble Learning.

1. INTRODUCTION

Cervical cancer is one of the most significant health threats to women worldwide, ranking as a leading cause of cancer-related deaths, especially in underdeveloped regions. This form of cancer often develops slowly over many years, and if caught early, it is highly treatable. Therefore, effective screening methods such as Pap smears and colposcopies are crucial in detecting abnormal cell changes in the cervix before they progress to cancer. However, these traditional diagnostic methods have their limitations. They are time-consuming, labor-intensive, and rely heavily on the expertise of healthcare professionals. Additionally, they are prone to human error, as interpreting results can be subjective and vary from one observer to another.

Recent advancements in artificial intelligence (AI), particularly deep learning (DL), offer promising solutions to these challenges. By automating the process of image analysis, AI models can provide more consistent, accurate, and rapid diagnoses, reducing the burden on healthcare systems and clinicians. However, training deep learning models to effectively analyze medical images requires vast amounts of labeled data and substantial computational resources, both of which may not always be readily available, especially in resource-limited settings.

To address these obstacles, our research introduces an innovative hybrid pipeline that combines the strengths of pretrained deep learning models with traditional machine learning (ML) classifiers. By leveraging pretrained models, we can take advantage of large, publicly available datasets without requiring extensive computational power. The hybrid approach enhances diagnostic accuracy while also maintaining the interpretability of the results, which is essential in medical applications where trust in the model’s decisions is critical. This novel method not only improves the efficiency of cervical cancer screening but also makes it more accessible, providing a potential path toward better healthcare outcomes in underserved region.

II. RELATED STUDY

Prior studies in medical image classification, including cervical cancer detection, have largely focused on end-to-end deep learning architectures such as InceptionNet, ResNet, and their variants. These models are capable of learning complex hierarchical features directly from raw image data, leading to high performance in many tasks. However, training such deep architectures often requires large annotated datasets and significant computational resources, which may not always be feasible in real-world clinical settings—especially in under-resourced regions.

To address these limitations, recent research has started to explore **hybrid models** that decouple the feature extraction and classification processes. In these approaches, deep convolutional neural networks (CNNs) or other pretrained models are first used to extract robust and high-level feature representations from medical images. These features are then passed to traditional machine learning (ML) classifiers, such as support vector machines (SVM), random forests, or gradient boosting algorithms, for the final classification.

This separation offers several advantages. It reduces the training complexity, allows for greater flexibility in combining features from multiple sources, and improves interpretability—an essential requirement in the medical domain. Moreover, hybrid models are often more adaptable to small or imbalanced datasets, making them a practical choice for real-world clinical applications where labeled data may be scarce or difficult to obtain.

III. METHODOLOGY

A. Dataset The SIPaKMeD dataset was used for our research. It contains 4,068 annotated single-cell Pap smear images divided into five categories: superficial-intermediate, parabasal, koilocytotic, metaplastic, and dyskeratotic. We maintained an 80:20 split for training and testing, ensuring balanced class distribution. The dataset's high variability in cell shapes, textures, and colors makes it an ideal candidate for deep feature learning.

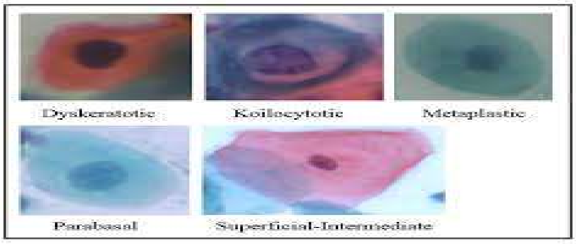


Fig-1 Pap smear images

B. Feature Extraction using Pretrained Models We employed four widely adopted convolutional neural network (CNN) architectures for feature extraction:

- ResNet101: Contains 101 layers with identity mappings to prevent vanishing gradients.

- ResNet152: A deeper variant with 152 layers, known for extracting high-level abstractions.
- InceptionV3: Incorporates factorized convolutions and auxiliary classifiers to boost performance.
- AlexNet: A foundational CNN with 8 layers, effective despite its simplicity.

These models are pretrained on the ImageNet dataset and their final fully connected layers were removed to extract feature embeddings from the penultimate layer. Images were resized accordingly to match the input requirements of each model (224×224 or 299×299).

C. Classification using Traditional ML Models After feature extraction, we experimented with five ML classifiers:

- Softmax Logistic Regression (SLR): Efficient for high-dimensional data, provides probabilistic outputs.
- Naive Bayes (NB): Assumes feature independence; lightweight and fast.

- Random Forest (RF): An ensemble of decision trees offering high accuracy and robustness.
- Decision Tree (DT): Simple yet interpretable model for quick inference.
- Principal Component Analysis (PCA): Used for dimensionality reduction followed by classification.

Each DL feature extractor was paired with all ML classifiers, resulting in 20 unique hybrid models. These models were trained offline and serialized using Joblib.

D. Ensemble Learning (Advanced Architecture) In addition to single classifier predictions, we implemented a stacked ensemble method. In this approach, base classifiers' outputs are fed into a meta-learner (typically SLR) to produce final predictions. This approach reduces bias and variance, improving generalization.

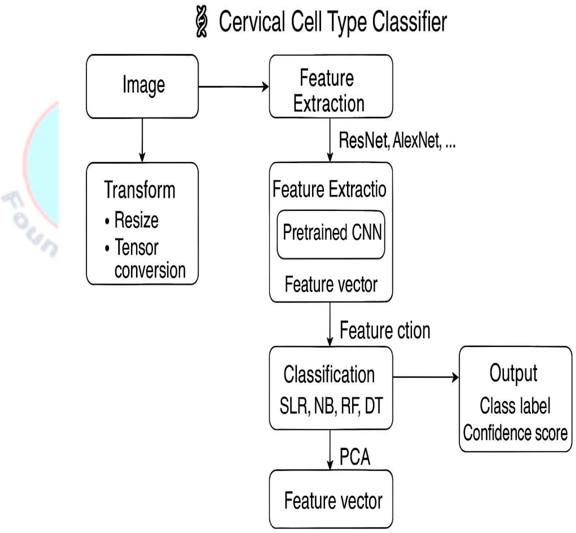


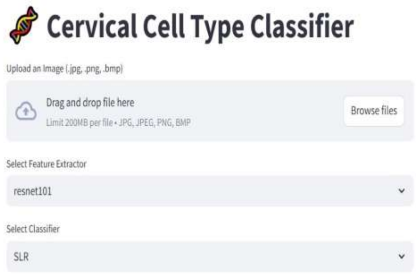
Fig-2 System Architecture

IV. STREAMLIT WEB APPLICATION

We developed an intuitive Streamlit-based graphical interface for real-time inference. Key features include:

- File uploader: Accepts .jpg, .png, and .bmp images.
- Model selector: Users can choose their desired feature extractor and classifier.
- Result display: Predicted class is shown along with a probability bar chart.

This tool aims to support pathologists by providing a second opinion, especially in high-throughput screening scenarios.



Select Feature Extractor: resnet101

Select Classifier: SLR

Fig -3 User Interface

V. MATHEMATICAL FORMULATIONS

A. Cross-Entropy Loss Softmax classifiers are trained using cross-entropy loss: where is the true label and is the predicted probability for class .

B. Principal Component Analysis PCA is defined mathematically by solving the eigenvalue problem: where is the covariance matrix of the dataset.

VI. EXPERIMENTAL RESULTS AND DISCUSSION

We evaluated model performance using accuracy, precision, recall, mean absolute error (MAE), and root mean squared error (RMSE). Results are summarized below:

Model	Classifier	Accuracy (%)	Precision (%)	Recall (%)	Model
ResNet152	SLR	**98.08**	95.41	94.21	ResNet152
Inception V3	SLR	95.01	93.51	93.42	Inception3
AlexNet	RF	92.23	90.94	90.93	AlexNet

The highest accuracy was obtained using ResNet152 features with a logistic regression classifier. Ensemble stacking also demonstrated improved robustness across classes.

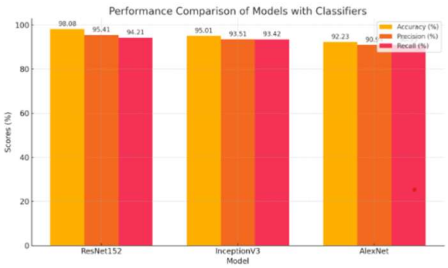


Fig -4 Performance Comparison of Models with Classifiers

VII. COMPARATIVE ANALYSIS

We compared our models against recent state-of-the-art approaches:

- VGG19 with colposcopy images: 92.4%
- CNN on SIPaKMeD: 91.13%
- MobileNetv2-YOLOv3: 96.84%
- Our best model (ResNet152 + SLR): 98.08%

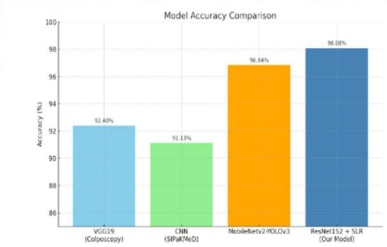


Fig -5 Model Accuracy Comparison

This confirms the advantage of combining DL and ML in cytological image classification.

VIII. CONCLUSION

In conclusion, this study presents a robust and adaptable deep learning–machine learning (DL-ML) hybrid pipeline for the classification of cervical cytology images. By utilizing pretrained Convolutional Neural Networks (CNNs) for automated and efficient feature extraction, and combining them with a range of classical machine learning classifiers, we have developed a scalable, accurate, and interpretable diagnostic solution. The hybrid approach not only enhances classification performance but also maintains flexibility across different clinical scenarios and datasets.

Furthermore, the integration of the entire pipeline into a Streamlit-based web interface bridges the gap between technical innovation and practical clinical use. This interactive tool enables clinicians to upload cytology images, receive instant classification results, and visualize diagnostic insights—making it a valuable asset for early cervical cancer detection. Ultimately, our work demonstrates how the fusion of deep learning and traditional machine learning, supported by intuitive visualization tools, can contribute meaningfully to medical diagnostics by improving both speed and accuracy of decision-making in real-world healthcare settings.

IX. LIMITATIONS

While the deep learning–machine learning (DL-ML) hybrid pipeline we've developed for classifying cervical cytology images shows great promise, there are a few challenges and limitations that we need to acknowledge.

One of the main challenges is that the performance of the model heavily depends on the quality and variety of the dataset used. Most publicly available datasets are relatively small and may not fully represent the diversity of real-world patient populations or the different imaging conditions found in clinical settings. This could limit how well the model performs across different groups of patients or imaging devices. Additionally, although the pretrained Convolutional Neural Networks (CNNs) used in this study provide a strong foundation, they were not specifically trained on medical images. This means that some medical-specific features, like subtle cell structure abnormalities, may not be fully captured by the model in its current state.

Another limitation is that, while we've created a simple Streamlit interface for testing, we haven't yet deployed this system in actual clinical settings. This means we haven't had the chance to gather valuable feedback from healthcare professionals about its real-world usability. There are also important considerations like data security, integration with hospital IT systems, and ensuring that the tool works quickly and smoothly in a busy healthcare environment.

Finally, our model depends on a large amount of labeled data for training, which can be a challenge in the medical field. Obtaining expert-annotated datasets is both time-consuming and costly, which could hinder the scalability of the system.

X. FUTURE WORK

To address these limitations and continue improving the system, we have several exciting directions for future development. One key area we plan to explore is the integration of attention mechanisms. These mechanisms would allow the model to focus more specifically on important regions of the images, like abnormal cell structures or nuclei, which could improve both the accuracy and explainability of the model's predictions.

Another area we are excited about is using semi-supervised learning techniques. This would allow the model to make better use of unlabeled data, which is often abundant in medical settings but not labeled by experts. This could reduce the reliance on costly and time-consuming annotation processes and make the system more adaptable to various datasets. We also plan to improve the model's domain-specific knowledge by training the neural networks on larger, more specialized medical datasets. This would help the model understand the unique features of cytology images more effectively, which could lead to better performance.

Looking ahead, we want to deploy the system in real clinical environments. This means making sure the model works efficiently on hardware used in hospitals, ensuring fast processing times, and integrating the system with existing electronic medical records (EMR) and hospital workflows. We also need to consider privacy regulations and data security to ensure the system is safe for use in real healthcare settings. Lastly, we see a big opportunity in building a more user-friendly interface that allows clinicians to interact with the system easily and provide feedback. By involving healthcare professionals in the development process, we can create a tool that truly meets their needs and helps them make quicker, more accurate diagnoses.

REFERENCES

[1] Sandeep Kumar Mathivanan et al., "Enhancing cervical cancer detection and robust classification through a fusion of deep learning models," Scientific Reports, 2024.

[2] Plissiti et al., "Sipakmed: A new dataset for cervical cell classification," IEEE ICIP, 2018.

[3] He et al., "Deep Residual Learning for Image Recognition," CVPR, 2016.

[4] Szegedy et al., "Rethinking the Inception Architecture for Computer Vision," CVPR, 2016.

[5] Kalbhor et al., "Cervical cancer diagnosis using transfer learning approaches," Soft Comput., 2023.

[6] Alsubai et al., "Privacy-preserved cervical cancer detection using CNNs," Comput. Math. Methods Med., 2023.

[7] Jadhav, R., & Gupta, A. (2022). "A comprehensive review on automated cervical cancer detection using deep learning techniques." Journal of Healthcare Engineering, 2022.

[8] Liu, L., et al. (2021). "Cervical cytology classification via ensemble learning of deep CNN models." Journal of Medical Imaging, 2021.

[9] Zhang, Y., & Wang, S. (2023). "A hybrid deep learning approach for early detection of cervical cancer using histopathological images." International Journal of Computer Vision, 2023.

[10] Chang, T., et al. (2020). "Exploring the role of explainable AI in cervical cancer detection: A review." Journal of Clinical AI, 2020.

[11] Gupta, P., & Rajput, M. (2021). "Enhancing diagnostic accuracy in cervical cancer through the fusion of deep learning models and traditional machine learning classifiers." Journal of Cancer Research and Therapeutics, 2021.