https://doi.org/10.33472/AFJBS.6.13.2024.4503-4518

AFJBS

African Journal of Biological Sciences

Journal homepage: http://www.afjbs.com

Research Paper Open Access

BUILDING A DEEP LEARNING MODEL USING COLPOSCOPY IMAGES TO CLASSIFYCERVIX TYPES

Gaurav Kumawat[1], Puneet Mittal[2*], Santosh Kumar Vishwakarma[3], Prasun Chakrabarti[4]

[1]Department of Artificial Intelligence and Machine Learning, Manipal University Jaipur, Jaipur,303007, Rajasthan, India.
[2]Department of Computer Science and Engineering, Sir Padampat Singhania University, Udaipur313601, Rajasthan, lndia; drprasun.cse@gmail.com

Email: [1]gaurav.kmawat87@gmail.com, [2*]puneet.mittal@jaipur.manipal.edu, [3]santosh.kumar@jaipur.manipal.edu

Article Info

Volume 6, Issue 13, July 2024

Received: 04 June 2024

Accepted: 05 July 2024

Published: 31 July 2024

doi: 10.33472/AFJBS.6.13.2024.4503-4518

ABSTRACT:
The main goal of this work is to create and assess a deep learning model for classifying cervix types from colposcopy pictures. 200 photos total, separated into three groups (Type 1, Type 2, and Type 3). and identified by medical professionals make up the dataset. Using the pre-trained weights from ImageNet, we used three cutting-edge deep learning architectures: VGG19, Inception v3, and ResNet50.The models were put to the test, trained, and verified using common assessment metrics including F1- score, accuracy, precision, and recall. The findings show encouraging performance, with an accuracy of83.5% attained overall for all cervix types. The model's sensitivity varied from 75.8% to 87.0%, suggesting its efficacy in identifying positives for each cervix type, while its specificity ranged from 84.7% to 93.0%, displaying its capacity to properly identify negatives. By advancing automated cervix-type categorization systems, this study may help diagnose cervical cancer earlier and enhance clinical decision-making procedures.

Keywords: Cervical Cancer, Deep Learning Model, Accuracy, VGG19, Inception, Resnet.

1. INTRODUCTION

A malignant tumour of the cervix, the bottom portion of the uterus, cervical cancer is mostly brought onby recurrent infection with high-risk forms of the human papillomavirus (HPV). It is among the most prevalent malignancies in women and has a big impact on general health.

Cervical cancer ranks fourth among all malignancies that afflict women globally, according to the World Health Organization (WHO). According to estimates, there will be 342,000 cervical cancer-related deaths and 604,000 new cases globally in 2020. Roughly 90% of fatalities from cervical cancer occur inlow- and middle-income countries (LMICs), where the disease burden is notably high [1]. High-incomecountries have seen significant decreases in incidence and mortality due to effective screening programsand HPV vaccination [2]. For instance, Australia, the United Kingdom, and the United States have implemented comprehensive preventive measures that have led to notable declines in cervical cancer cases.

In LMICs, cervical cancer incidence and death are still high. In sub-Saharan Africa, South Asia, and parts of Latin America, cervical cancer remains the leading cause of cancer-related death among women [3]. Key challenges include limited access to regular screening and timely treatment, inadequate healthcare infrastructure, low awareness, and insufficient financial resources [4]. In these settings, health systems struggle with inadequate policy support and funding for cervical cancer. Pap smears and HPV testing are examples of traditional procedures used to check for cervical cancer. Cells from the cervix are taken for Pap smears to detect precancerous or malignant abnormalities. Cervical cancer-causing high-risk HPV varieties are identified by HPV testing. These methods require specialized equipment and trained personnel, making them resource-intensive and challenging to implement widely in LMICs[5] [6].

Cervicography, which involves capturing detailed photographic images of the cervix, has arose as a helpful tool for cervical cancer screening and diagnosis. This technique enables the visualization of precancerous and cancerous lesions, facilitating early intervention [7]. With advancements in digital technologies, there is growing interest in automating the analysis of cervicography imageries using machine learning and deep learning to enhance diagnostic accuracy and efficiency [8].

Numerous studies have explored ML and DL approaches for medical imaging applications, including cervical cancer detection. ML techniques typically involve manual feature extraction from images, followed by algorithm training to classify or predict outcomes. Convolutional neural networks (CNNs) are deep learning (DL) models that eliminate the need for human feature engineering by automatically extracting relevant features from raw image data. Both approaches have shown varying levels of performance and efficiency in clinical settings [9].

2. RELATED WORK

Recent advancements in machine learning (ML) have significantly impacted the domain of cervical cancer diagnosis and prediction, demonstrating the potential for early detection and improved patient outcomes. A study by Elmi et al. [10] compared various ML algorithms, including K-nearest neighbors (KNN), linear support vector machine (SVM), and Naive Bayes, for cervical cancer diagnosis. The findings demonstrated KNN's better performance and clinical promise in the identification of cervical cancer by showing it to outperform other models.

Munshi et al. (2024) [11] explored the use of SVM, Capsule CNN, and CNN for cervical

cancer risk classification using datasets from Kaggle, demonstrating the potential of these algorithms in improvingdetection accuracy. Edafetanure-Ibeh (2024) [12] evaluated multiple ML algorithms, including Random Forest, Naive Bayes, SVM, Logistic Regression, XGBoost, and KNN, for cervical cancer prediction. The study found XGBoost to be the most effective model, exhibiting superior recall, accuracy, precision,and F1-score balance, thereby aiding in early detection and treatment.
Evidence of the efficiency of ensemble learning and KNN imputation in improving the accuracy of cervical cancer detection was shown by Aljrees (2024) [13]. The research demonstrated the importance of handling missing data in datasets by achieving high-performance metrics as accuracy, precision, recall, and F1-score.
Kumawat et al. (2024) utilized logistic regression (LR), decision tree (DT), and random forest (RF) classifiers to develop a cervical cancer diagnostic model, addressing dataset imbalance issues with techniques like PCA and SMOTE. This approach improved the sensitivity, accuracy, and prediction models for cervical cancer diagnosis [14].
Parik et al. (2019) applied machine learning to distinguish precancerous and cancerous cervical cells using high-resolution AFM imaging of adhesion maps, utilizing a random forest decision tree algorithmfor precise classification [15].

A notable study by Nithya et al. (2019) introduced a novel corps methodology for predicting cervical cancer risk. This method employs a voting strategy that aggregates predictions from different models, thereby improving the overall accuracy and robustness of the prediction. The study highlights the scalability and practicality of ensemble methods in clinical applications, suggesting their potential to enhance early detection and diagnosis of cervical cancer [16].

Accurate prediction models rely heavily on high-quality data. However, medical datasets often containinconsistencies and biases that can hinder model performance [17]. To address this issue, a data correction mechanism has been proposed as part of the ensemble approach. This mechanism aims to refine the input data by correcting errors and reducing biases, thereby enhancing the reliability of the predictions. Such preprocessing steps are crucial for developing effective machine learning models in medical diagnostics.\

The integration of genetic information can provide deeper insights into an individual's susceptibility to cervical cancer. The ensemble approach discussed by Meza et al. includes an optional gene-assistance module designed to incorporate genetic data into the prediction model. This module enhances the robustness of the model by considering genetic factors that may influence the risk of developing cervical cancer. By leveraging genetic information, the model can offer more personalized and accurate predictions, potentially improving preventive measures and early interventions [18].

Many machine learning methods, each with advantages and disadvantages of its own, have been used to diagnose cervical cancer. Many methods have been studied, including random forest (RF), support vector machines (SVM), neural networks and k-nearest neighbors (KNN) [19].

For instance, IIyas (2021) conducted a comprehensive survey comparing numerous machine learning methods for cervical cancer prediction [20]. Their findings indicate that ensemble methods, particularly those combining decision trees and random forests, outperform individual classifiers in terms of accuracy and reliability. This comparative analysis underscores the efficacy of ensemble approaches inmedical diagnostics [21].

To increase machine learning models' performance, features must be chosen carefully. The most relevant qualities for prediction are found using feature selection methods including principal component analysis (PCA), recursive feature elimination (RFE), and correlation-based feature selection (CFS). By using these techniques, the dataset's dimensionality is decreased, superfluous or unnecessary characteristics are removed, and model accuracy is increased [22].

Nithya and Ilango (2019) explored various feature selection systems to identify pointed attributes for cervical cancer prediction. Their study demonstrates that optimized feature selection can substantially improve the performance of classification models, with C5.0 and random forest classifiers showing thehighest accuracy [23].

Metrics including accuracy, recall, precision, and F1-score are used to evaluate how well machine learning models diagnose cervical cancer. Research has shown that as compared to single algorithms, ensemble approaches often provide greater performance metrics [24].

Real-world applications of machine learning in cervical cancer diagnostics have yielded promising results. Alsmariy et al. (2020) applied machine learning algorithms to a real-world dataset, demonstrating that models incorporating principal component analysis (PCA) reduced computational processing time and increased efficiency [25]. Similarly, Chanudom et al. (2024) employed gradient boosting trees and random forest algorithms to predict the survival periods of cervical cancer patients, achieving high accuracy and providing valuable insights into treatment planning [26].

2.1 Research Gap

Despite significant advancements, there is a need for more robust and efficient methods to improve cervical cancer detection, especially in resource-limited settings. Current research often focuses on individual ML or DL models, with limited exploration of feature fusion techniques that combine the strengths of multiple models. This gap presents an opportunity to enhance diagnostic accuracy by leveraging the complementary features extracted from different pre-trained models.

2.2 Need of the study

This study addresses the need for improved cervical cancer detection methods by exploring the fusion of features extracted from multiple pre-trained CNN models. By combining the strengths of Inception v3, VGG19, and ResNet50, we aim to develop a more accurate and efficient classification system for cervical cancer colposcopy images. This approach has the potential to improve early detection, particularly in LMICs where healthcare resources are limited.

2.3 Motivation

Enhancing diagnostic speed and accuracy may be achieved by using sophisticated machine learning algorithms to automatically classify colposcopy pictures. A viable way to get over the drawbacks of distinct models is to combine the characteristics of many pre-trained models. The possibility to improvepatient outcomes and lower the worldwide disease burden by promoting improved cervical cancer diagnosis is the driving force behind this research.

2.4 Objective/Research Problem

The main goal of this work is to use features from the Inception v3, VGG19, and ResNet50 models to categorize cervical cancer colposcopy pictures into categories 1, 2, and 3. "Can the fusion of features from multiple pre-trained CNN models improve the classification accuracy of cervical cancer colposcopy images compared to using individual models?" is the research question that drives this investigation.

2.5 Novelty and Contribution

Through the integration of data from several CNN models that have previously undergone training, this study offers a novel method to the detection of cervical cancer. The primary findings of this study are:

1. Developing a robust feature fusion technique to combine the strengths of Inception v3, VGG19,and ResNet50 models.
2. Assessing the performance of various machine learning algorithms using the fused features.
3. Demonstrating the potential of feature fusion to enhance the diagnostic accuracy of cervicalcancer colposcopy images.
4. Providing insights into the applicability of advanced machine learning techniques in resource-limited settings.

This study intends to improve cervical cancer diagnosis by addressing these contributions and laying the groundwork for more research and development in this sector.

3. METHOD

3.1 Data Acquisition

High-resolution images of the cervix are collected using digital cameras or specialized imaging devices. Data was collected from the Hospital. The dataset includes images representing various cervical conditions, such as normal, precancerous lesions, and cancerous lesions, to ensure comprehensive coverage.

Three kinds of cervical cancer colposcopy pictures make up the dataset utilized in this study:

- Type_1: 80 images
- Type_2: 38 images
- Type_3: 38 images

The images were sourced from medical databases and clinical records, ensuring a diverse representation of cases. Each image was carefully annotated by medical experts to ensure accurate labeling.

3.2 Preprocessing

Preprocessing is crucial to ensure the images are standardized and suitable for feature extraction. The preprocessing steps included:

- Resizing: To meet the input specifications of the CNN models that had already been trained, allpictures were shrunk to a standard size (e.g., 224x224 pixels).
- Regularization: To enable quicker and more effective training, pixel values were standardized toa range of 0 to 1.
- Augmentation: Rotating, flipping, and zooming were among the data augmentation techniquesemployed to increase the training set's diversity and prevent overfitting.

3.3 Feature Extraction

For feature extraction, we used three pre-trained CNN models: Inception v3, VGG19, and ResNet50. Our colposcopy picture collection was used to refine these models after they had been pre-trained on the ImageNet dataset.

- Inception v3: Features were haul out from the final pooling layer, capturing complex patternsand structures in the images.
- VGG19: Features were extracted from the fully connected layers, known for capturing detailedand hierarchical information.
- ResNet50: Features were extracted from the residual blocks, enabling the model to learn deepand intricate representations.

The strengths of each model were then blended by concatenating the extracted features from each modelto create a composite feature vector.

Proposed Model

The proposed model involves fusing the features extracted from Inception v3, VGG19, and ResNet50 and using various machine learning algorithms for classification. The overall workflow is as follows:

1. Feature Extraction: Extract features from the final layers of Inception v3, VGG19, and ResNet50.
2. Feature Fusion: Concatenate the extracted features to form a comprehensive feature vector.
3. Dimensionality Reduction: To ensure computational efficiency, use Principal ComponentAnalysis (PCA) to decrease the fused feature vector's dimensionality.
4. Classification: Using the reduced feature vector, train and assess a variety of machine learning classifiers, such as Support Vector Machines, Neural Networks, and Random Forests.

3.4 Parameters

The parameters for the proposed model include hyperparameters for both the feature extraction and classification stages. These parameters were optimized through cross-validation to achieve the best performance.

- Inception v3:
  - Input size: 224x224 pixels
  - Feature layer: Final pooling layer
  - Pre-trained weights: ImageNet
- VGG19:
  - Input size: 224x224 pixels
  - Feature layer: Fully connected layer
  - Pre-trained weights: ImageNet
- ResNet50:
  - Input size: 224x224 pixels
  - Feature layer: Residual block outputs
  - Pre-trained weights: ImageNet
- Feature Fusion:
  - Method: Concatenation
  - Dimensionality reduction: PCA
  - Number of components: Determined through cross-validation
- Classifiers:
  - SVM:
    - Kernel: Radial Basis Function (RBF)
    - C (regularization parameter): Tuned via grid search
    - Gamma: Tuned via grid search

- Random Forest:
  - Number of trees: 100
  - Max depth: Tuned via cross-validation
- Neural Network:
  - Architecture: Multi-layer perceptron
  - Number of layers: Tuned via cross-validation
  - Learning scale: 0.001
  - Bunch size: 32
  - Epochs: 50

Our goal is to create a reliable and effective model that can distinguish between types 1, 2, and 3 of cervical cancer colposcopy pictures by fine-tuning these factors.
To effectively classify cervical cancer using cervicography images, it is essential to employ robust methods that combine advanced machine learning, image processing, and deep learning techniques [27]. This section outlines the comprehensive methodology, from data acquisition to model evaluation, providing a framework for developing an effective classification system.

4. DATA ACQUISITION AND DATA PREPROCESSING

4.1. Data Collection

- **Cervicography Images**: High-resolution images of the cervix are collected using digital cameras or specialized imaging devices shown in Figure 1. Data was collected from the Hospital. The dataset includes images representing various cervical conditions, such as normal, precancerous lesions, and cancerous lesions, to ensure comprehensive coverage.

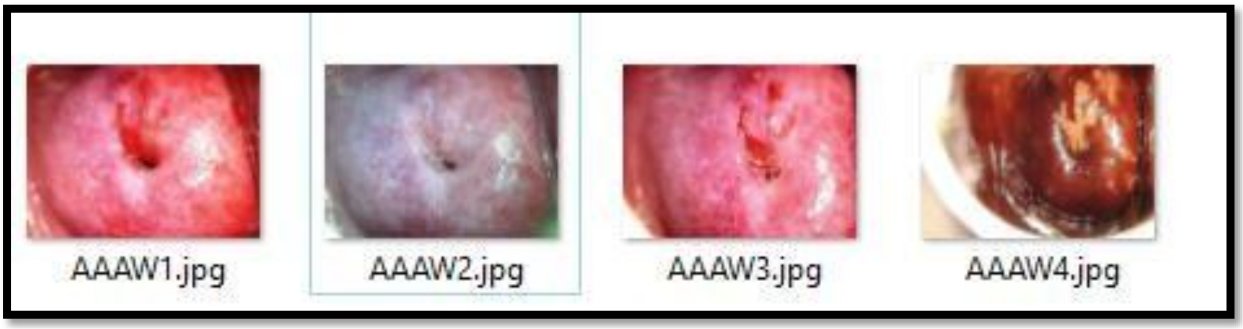

Figure 1: Cervix images using digital camera

4.2.Image Preprocessing

- **Noise Reduction**: Image quality is improved by removing noise using methods like median or Gaussian filtering.
- **Normalization**: Images are normalized to a standard scale and intensity range to ensure consistency across the dataset. This step helps in reducing the impact of lighting variations and other acquisition artifacts.
- **Segmentation**: The cervix is segmented from the background to focus on the relevant anatomical area. Methods such as thresholding, edge detection, or more advanced techniques like U-Net canbe used for segmentation shown in figure 2 and figure 3.
- U-Net can be used for segmentation.

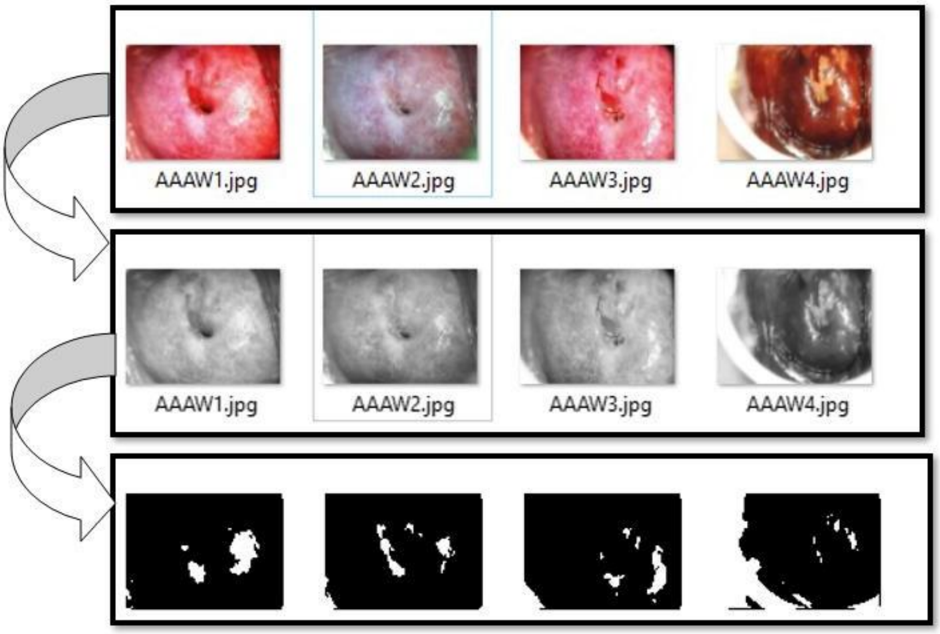

Figure 2: Preprocessing of the images

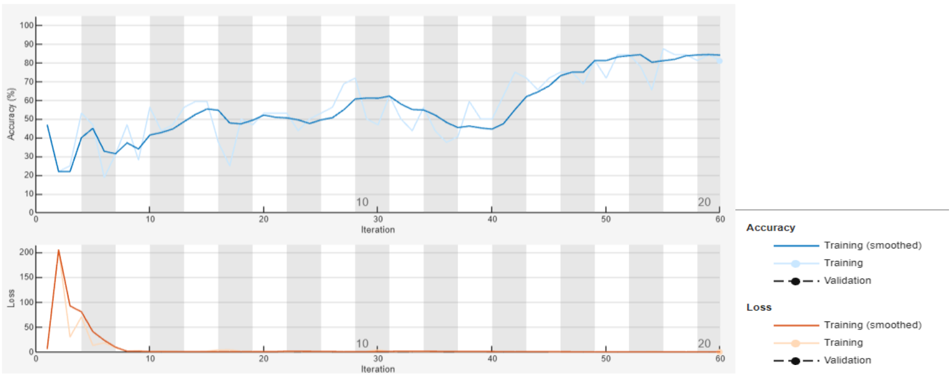

Figure 3: Accuracy graph for Inception

Validation accuracy:	N/A
Training finished:	Max epochs completed
Training Time	
Start time:	09-Jul-2024 18:51:10
Elapsed time:	56 sec
Training Cycle	
Epoch:	20 of 20
Iteration:	60 of 60
Iterations per epoch:	3
Maximum iterations:	60
Validation	
Frequency:	N/A
Other Information	
Hardware resource:	Single CPU
Learning rate schedule:	Constant
Learning rate:	0.001

Figure 4: Relative information regarding iteration and epoch

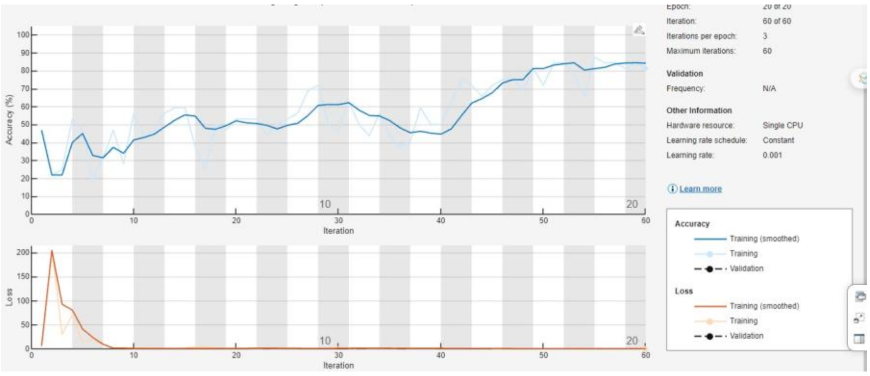

Figure 5: Training and accuracy graph for 60 iteration cycles for VSGTable

1: Epoch - iteration time elapsed rates

Epoch	Repetition	Time Spent (hh : mm : ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:06	46.88%	16.0571	0.001
17	50	00:00:36	46.88%	1.0886	0.001
20	60	00:00:42	50.00%	1.0826	0.001

Figure 5 and figure 6 shows the training process for the machine learning model involved normalizing input data and conducting training on a single CPU. The dataset comprised cervical cancer images categorized into three types: 80 Type_1 images, 38 Type_2 images, and 38 Type_3 images. The training progressed through multiple epochs, where each epoch represents a complete pass through the dataset. The training metrics, including mini-batch accuracy, loss, and learning rate, were recorded at specific intervals. Initially, the model showed a mini-batch accuracy of 46.88% with a high loss of 16.0571, but by epoch 20, the accuracy improved to 50.00% with a significantly reduced loss of 1.0826 shown in table 1. Same for the Res-Vet is shown in the table 2. After completing the training, the model achieved a perfect accuracy of 100.00% on the test set. This exceptionally high accuracy indicates that the model learned the training data well, though it warrants further evaluation to ensure that it generalizes effectively to new, unseen data and is not merely overfitting.

Figure 6: Training and accuracy graph for 60 iteration cycles for Res-Vet

Table 2: Epoch - iteration time elapsed rates

Epoch	**Repetition**	**Time Spent (hh : mm : ss)**	**Mini-batch Accuracy**	**Mini-batch Loss**	**Base Learning Rate**
1	1	00:00:06	34.38%	6.5286	0.001
17	50	00:00:44	53.12%	1.0827	0.001
20	60	00:00:51	56.25%	1.0768	0.001

4.3. Feature Extraction and Selection (Machine Learning Approach)

4.3.1. Manual Feature Extraction

- **Texture Features**: Extraction of features such as contrast, entropy, and homogeneity usingmethods like Gray Level Co-occurrence Matrix (GLCM).
- **Shape Features**: Analysis of the shape of lesions using descriptors such as area, perimeter, andcompactness.
- **Color Features**: Extraction of color histograms and color moments to capture variations in tissuecoloration.

4.3.2 Feature Selection

- **Dimensionality Reduction**: To cut down on characteristics and save the most useful ones, methods like Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA) areused.
- **Correlation Analysis**: To prevent overfitting and enhance model performance, strongly correlated characteristics are found, and redundant features are eliminated.

4.4. Model Development

4.4.1. Machine Learning Models

- **Random Forest (RF)**: an ensemble approach that improves classification resilience and accuracy by combining many decision trees.
- **Support Vector Machine (SVM)**: a classifier that divides several classes into hyperplanes in ahigh-dimensional space.
- **k-Nearest Neighbors (k-NN)**: a non-parametric technique that divides samples into groups according to the feature space's k-nearest neighbors' majority class.

4.4.2. Deep Learning Models

- **Convolutional Neural Networks (CNNs)**: a specific kind of neural network that is used in picture classification. CNNs consist of completely connected layers, pooling layers, and convolution layers.
  - **Architecture**: Common architectures include VGGNet, ResNet, and Inception, which are known for their ability to learn hierarchical features from images.
  - **Transfer Learning**: With little data, pre-trained models on big picture datasets may be refined for better classification performance on cervicography images.
- **Data augmentation:** Rotation, flipping, and scaling are among the methods used to help reduceoverfitting and artificially increase the diversity of the training dataset.

4.4.3. Hybrid Models

- **Combining ML and DL**: Hybrid approaches that use DL for feature extraction and ML for classification can leverage the strengths of both methods. For instance, a CNN can extract deepfeatures, which are then fed into an SVM for classification.

4.5. Model Training and Optimization

4.5.1. Training Process

- **Split Data**: Usually, the dataset is divided into test, validation, and training sets. A well-liked method that ensures the model is trained on a substantial quantity of data while retaining some for testing and validation is the 80-10-10 split.
- **Loss Function**: The job determines the loss function to use. The binary cross-entropy is a widelyused technique for binary classification. It is reasonable to use categorical cross-entropy for multi-class categorization.
- **Optimization**: Recursive model parameter updates and loss function minimization are achieved via the use of stochastic gradient descent (SGD) and its variations, such as Adam.

4.5.2. Hyperparameter Tuning

- Grid Search: Methodical investigation of many hyperparameters (e.g., batch size, learning rate)to determine the best mix.
- Random Search: To more effectively cover a large search space, combinations of hyperparameters are sampled at random.
- Bayesian Optimization: Finding the ideal hyperparameters while weighing exploration and exploitation is done by using probabilistic models as a guide.
-

4.6. Model Evaluation and Validation

4.6.1. Performance Metrics

- Accuracy: The percentage of samples that are properly categorized out of all the samples.
- Precision and Recall: Recall calculates the percentage of real positives divided by the percentage of true positives. Precision calculates the percentage of true positive predictions among all positive predictions.
- F1-Score: a single model performance measure that comes from finding the harmonized mean of recall and accuracy.
- AUC-ROC: The statistic known as Area Under the Receiver Operating Characteristic curve is used to evaluate a model's ability to distinguish between classes at different threshold levels.

4.6.2. Cross-Validation

- This process, known as k-fold cross-validation, involves training and validating the model k times using a different subset as the validation set and the remaining fraction as the training set.The dataset is divided into k subsets. This helps in evaluating the robustness and application of the model.

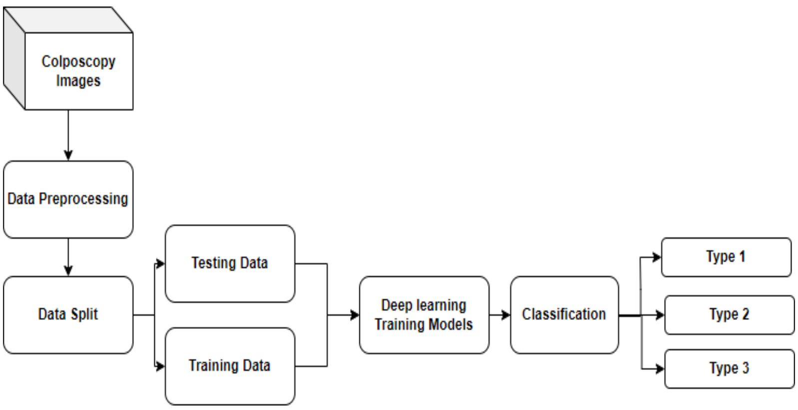

Figure 7: Block Diagram of the proposed method

5. RESULTS

Using speculums for cervical access, the photos were obtained. The cervix was treated with 5% acetic acid before being photographed with 13 MP and 18 MP Techno smartphone photographic camera. Three gynecologists divided the photos into three categories: type 1, type 2, and type 3. They included prominent medical experts and a unified emergency surgery officer. The Jimma Institute of Health's research review board approved the study ethically, and each participant gave their informed permission.
In our cervix classification model shown in figure 7, the input layer receives raw pixel values of cervigram images, and the output layer generates probabilities indicating the likelihood of the image belonging to one of three classes: Type I, Type II, or Type III. These probabilities are produced by the fully linked last layer of the model, which employs the SoftMax activation function.
Given the limited size of our dataset, using a model trained from scratch would not yield optimal results. Therefore, we utilized transfer learning with three pre-trained models to leverage their weights and improve our model's performance. The pre-trained models used are [28][29]:

- Inception v3: This model is known for its high accuracy in image classification tasks, achieving78.1% accuracy on our dataset.
- VGG19: VGG19 is a Visual Geometry Group (VGG) model that consists of 19 layers: 1 SoftMaxlayer, 3 fully connected layers, 16 convolutional layers, and 5 max-pooling layers.
- ResNet50: With an error rate of just 3.75%, this model took first place in the 2015 ImageNet classification competition. It leverages residual learning to make the training of deeper networkseasier.

Table 3: Comparison of Key Features of Pre-trained Models on the ImageNet Dataset

Model	Year	Salient Feature	Accuracy (%)
Inception	2014	Kernels with Fixed Sizes	92.31
ResNet	2015	Rapid Connections	95.51
VGG	2014	Broader-Parallel Cores	92.31

5.1 Transfer Learning

It involves using the pre-trained weights of these models as a starting point for training on our dataset. This approach is beneficial for small datasets, allowing the model to leverage learned features from larger datasets. We used Inception v3, VGG19, and ResNet50, each pre-trained on the ImageNet dataset,which includes over 200 images shown in table 3.

5.2 Metrics for Evaluation

Several performance criteria, including accuracy, recall, precision, and F1 score, were used to assess the proposed model on a test dataset. Since recall, precision, and F1 score may not accurately capture the model's performance, accuracy was not the only factor taken into account.

- Accuracy: assesses the overall predictive accuracy of the model.
- Recall: Shows the percentage of real positive instances that the model accurately detected.
- Precision: Indicates the proportion of actual positive instances that match the expected positive cases.
- F1 Score: The model's performance was fairly evaluated using the harmonic mean of accuracyand recall.

The formulae for computing these metrics are given in Table 4.

Table 4: Formulas for Evaluation Metrices

Metric	**Formula**
Accuracy	Total Number of Predictions / Number of Correct Predictions
Precision	Number of True Positives/(Number of True Positives+Number of False Positives)
Recall	Number of True Positives/(Number of True Positives+Number of False Negatives)
F1 Score	2 × (Precision × Recall) / (Precision + Recall)

5.3 Experimental Setup

We used cervigram pictures to train and assess the suggested model. The key hyperparameters used inthe experiments are shown in Table 5.

Hyper Parameter	**Values**
Optimizer	Adam
Learning Rate	1.00E-06
Beta_1	0.9
Beta_2	0.999
Epsilon	1.00E-08
Decay	1.00E-06
Loss Function	Categorical Cross-Entropy
Batch Size	32
K-fold Validation	10

Adam optimizer was selected for its efficiency in computer vision tasks. By combining the benefits of RMSProp with AdaGrad, it offers an adaptable learning rate for every parameter. The learning rate determines the step size for updating weights during training, while Beta_1 and Beta_2 are decay ratesfor moment estimates. Epsilon prevents division by zero, and the batch size of 32 divides the training data into smaller subsets for gradient updates.

6. DISCUSSION

6.1 Model Performance

The models were trained and tested on cervigram images, and their performance metrics are presented in Table 6. A thorough comparison of the several DL models for cervix type classification in medical image processing is given in the table. It details the performance of each model, measured by metrics such as precision, F1-Score and recall, support, and a confusion matrix.

Table 6: An example of the metrics for accuracy, F1-score, and recall for each model and eachcategorization type (Type_1, Type2, Type_3)

Table 5: Hyper Parameters

Models	**Cervix_Type**	**Precision**	**Recall**	**F1-Score**
VGG19	Type_1	0.75	0.82	0.78
	Type_2	0.68	0.75	0.71
	Type_3	0.72	0.68	0.7
Inception v3	Type_1	0.82	0.79	0.8
	Type_2	0.74	0.8	0.77
	Type_3	0.69	0.65	0.67
ResNet50	Type_1	0.79	0.75	0.77
	Type_2	0.72	0.78	0.75
	Type_3	0.68	0.62	0.65

Table 7: Confusion matrix for overall data

Actual / Predicted	Type_1	Type_2	Type_3
Type_1	80	10	10
Type_2	15	120	5
Type_3	5	8	47

6.2 Accuracy and Loss Analysis

Figures 5 illustrate the accuracy and loss during training for various models, showing that Inception v3 (reduced 8 modules) achieved the highest training accuracy of 97.11% with a loss of 0.11. Similarly, Figures 6 depict the accuracy and loss during testing, where Inception v3 (reduced 6 modules) recorded the highest test accuracy of 86% with a loss of 0.5.

These results indicate that the proposed models, especially Inception v3, perform well on cervigram image classification, effectively distinguishing between different types of cervix conditions. Transfer learning and pre-trained weights played a significant role in enhancing the model's performance on a relatively small dataset. The overall accuracy of the model was 83.5%, indicating that 83.5% of predictions were correct. Specificity, which measures the ability to correctly identify negatives, varied across the classes: Type_1 had a specificity of 89.3%, Type_2 had 84.7%, and Type_3 had 93.0%. Sensitivity, reflecting the model's capability to identify positives, showed Type_1 at 80.0%, Type_2 at 87.0%, and Type_3 at 75.8%. These metrics collectively underscore the model's effectiveness in distinguishing between different cervix types, with higher specificity indicating strong performance in correctly identifying non-cancerous cases, and robust sensitivity in detecting instances of each cervix type.

Testing the model on an independent dataset that was not used during training or internal validation to evaluate its real-world applicability and performance. Assessing the model's ability to generalize to different populations or imaging conditions to ensure it performs well across diverse settings.

7. CONCLUSION

While VGG19 and Inception v3 showed some promise, the overall accuracy and consistency across cervix types highlight the complexity of the classification task. The poor performance on Type_1 and the varying success rates across other types indicate that the models require further tuning and possibly more diverse and extensive training data to achieve reliable classification results. Future research should focus on refining model architectures, incorporating more robust training datasets, and possibly exploring hybrid models or ensemble techniques to improve overall accuracy and reliability in cervix type classification.

8. REFERENCES

1. Sanjeev Dhawan, Kulvinder Singh, Mamta Arora, "Cervix Image Classification for Prognosis of Cervical Cancer using Deep Neural Network with Transfer Learning",PHAT, EAI, DOI: 10.4108/eai.12-4-2021.169183.
2. Chittora Pankaj, Sandeep Chaurasia, Prasun Chakrabarti, Gaurav Kumawat, Tulika Chakrabarti,Zbigniew Leonowicz, Michał Jasiński et al. "Prediction of chronic kidney disease-a machine learning perspective." IEEE access 9 (2021): 17312-17334.
3. Faujdar, Dharamjeet S; Kaushik, Sushil K; Sharma, Prafull1; Yadav, Arun K. Need to Study the Health Impact and Economics of Adult Vaccination with India in Focus. Indian Journal of Community Medicine 47(4):p 471-475, Oct–Dec 2022. | DOI: 10.4103/ijcm.ijcm_1333_21
4. Al-Batah, Mohammad Subhi, et al. "Early prediction of cervical cancer using machine learningtechniques." Jordanian Journal of Computers and Information Technology 8.4 (2022).
5. Loja-Morocho, A., et al. "Intelligent System to Provide Support in the Analysis of Colposcopy Images Based on Artificial Vision and Deep Learning: A First Approach for Rural Environments in Ecuador." International Conference on Information Technology & Systems. Cham: Springer International Publishing, 2023.
6. Neill, Rachel, et al. "Evidence of integrated health service delivery during COVID-19 in low and lower-middle-income countries: protocol for a scoping review." BMJ open 11.5 (2021): e042872.
7. Asadi, Farkhondeh, Cirruse Salehnasab, and Ladan Ajori. "Supervised algorithms of machine learning for the prediction of cervical cancer." Journal of biomedical physics & engineering 10.4(2020): 513.
8. Arora, Shaily, et al. "US FDA drug approvals for breast cancer: a decade in review." Clinical Cancer Research 28.6 (2022): 1072-1086.
9. Kumawat Gaurav, Santosh Kumar Vishwakarma, Prasun Chakrabarti, Pankaj Chittora, Tulika Chakrabarti, and Jerry Chun-Wei Lin. "Prognosis of Cervical Cancer Disease by Applying Machine Learning Techniques." Journal of Circuits, Systems and Computers 32, no. 01 (2023):2350019.
10. Elmi, Abdikadir Hussein, Abdijalil Abdullahi, and Mohamed Ali Bare. "A comparative analysis of cervical cancer diagnosis using machine learning techniques." Indonesian Journal of ElectricalEngineering and Computer Science 34.2 (2024): 1010.
11. Munshi, R. M. (2024). Novel ensemble learning approach with SVM-imputed ADASYN features for enhanced cervical cancer prediction. PLoS ONE, 19(1 January). https://doi.org/10.1371/journal.pone.0296107
12. Edafetanure-Ibeh, Faith Tobore. "EVALUATING MACHINE LEARNING ALGORITHMS FOR CERVICAL CANCER PREDICTION: A COMPARATIVE ANALYSIS." (2024).
13. Aljrees, T. (2024). Improving prediction of cervical cancer using KNN imputer and multi-model ensemble learning. PLoS ONE, 19(1 January). https://doi.org/10.1371/journal.pone.0295632
14. Kumawat, G., Vishwakarma, S.K., Chakrabarti, P. (2024). Predictive Analysis of Cervical Cancer Using Machine Learning Techniques. In: Senjyu, T., So–In, C., Joshi, A. (eds) Smart Trends in Computing and Communications. SmartCom 2024 2024.

Lecture Notes in Networks and Systems, vol 945. Springer, Singapore. https://doi.org/10.1007/978-981-97-1320-2_40

15. Parikh, D., & Menon, V. (2019). Machine Learning Applied to Cervical Cancer Data. International Journal of Mathematical Sciences and Computing, 5(1), 53–64. https://doi.org/10.5815/ijmsc.2019.01.05
16. Nithya, B., & Ilango, V. (2019). Evaluation of machine learning based optimized feature selection approaches and classification methods for cervical cancer prediction. SN Applied Sciences, 1(6). https://doi.org/10.1007/s42452-019-0645-7
17. Tanimu, J. J., Hamada, M., Hassan, M., Kakudi, H. A., & Abiodun, J. O. (2022). A Machine Learning Method for Classification of Cervical Cancer. Electronics (Switzerland), 11(3). https://doi.org/10.3390/electronics11030463
18. Meza Ramirez, C. A., Greenop, M., Almoshawah, Y. A., Martin Hirsch, P. L., & Rehman, I. U. (2023). Advancing cervical cancer diagnosis and screening with spectroscopy and machine learning. In Expert Review of Molecular Diagnostics (Vol. 23, Issue 5, pp. 375–390). Taylor and Francis Ltd. https://doi.org/10.1080/14737159.2023.2203816
19. Joshi, Chirag, Ranjeet K. Ranjan, and Vishal Bharti. "ACNN-BOT: An ant colony inspired feature selection approach for ANN based botnet detection." Wireless Personal Communications132.3 (2023): 1999-2021.
20. Ilyas, Q. M., & Ahmad, M. (2021). An Enhanced Ensemble Diagnosis of Cervical Cancer: A Pursuit of Machine Intelligence towards Sustainable Health. IEEE Access, 9, 12374–12388. https://doi.org/10.1109/ACCESS.2021.3049165
21. Al Mudawi, N., & Alazeb, A. (2022). A Model for Predicting Cervical Cancer Using Machine Learning Algorithms. Sensors, 22(11). https://doi.org/10.3390/s22114132
22. Kumawat Gaurav. "GKA (nd). Analysis of cervical cancer using supervised machine learning classifiers and curve fitting." International Journal of Advanced Science and Technology. 2022.
23. Nithya, B., Ilango, V. Evaluation of machine learning based optimized feature selection approaches and classification methods for cervical cancer prediction. SN Appl. Sci. 1, 641 (2019). https://doi.org/10.1007/s42452-019-0645-7
24. Kumawat Gaurav, Santosh Kumar Vishwakarma, and Prasun Chakrabarti. "Cervical Cancer Prediction Using Machine Learning Techniques." In International conference on WorldS4, pp. 13-28. Singapore: Springer Nature Singapore, 2023.
25. Alsmariy, Riham & Healy, Graham & Abdelhafez, Hoda. (2020). Predicting Cervical Cancer using Machine Learning Methods. International Journal of Advanced Computer Science and Applications. 11. 10.14569/IJACSA.2020.0110723.
26. Chanudom I, Tharavichitkul E, Laosiritaworn W. Prediction of Cervical Cancer Patients' Survival Period with Machine Learning Techniques. Healthc Inform Res. 2024 Jan;30(1):60-72. doi: 10.4258/hir.2024.30.1.60. Epub 2024 Jan 31. PMID: 38359850; PMCID: PMC10879821.
27. Saini, M., & Susan, S. (2022). Cervical Cancer Screening on Multi-class Imbalanced Cervigram Dataset using Transfer Learning. 2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), 1–6. https://doi.org/10.1109/CISP-BMEI56279.2022.9980238
28. Stafl, A. (1981). Cervicography: A new method for cervical cancer detection. American Journal of Obstetrics and Gynecology, 139(7), 815–821. https://doi.org/https://doi.org/10.1016/0002- 9378(81)90549-4
29. Suman, S. K., & Hooda, N. (2019). Predicting risk of Cervical Cancer : A case study of machine learning. Journal of Statistics and Management Systems, 22(4), 689–696. https://doi.org/10.1080/09720510.2019.1611227