

Deep Learning and Ensemble Approach for Cervical Cancer Detection using Medical Image Analysis

P. C. N. Lalithya^{1*}, Damodharan Susi², Adimulam Janaki², Challa Srilekha², Gaddam Maheswari²

¹Assistant Professor, ²UG Student, ^{1,2}Department of Electronics and Communication Engineering

^{1,2}Geethanjali Institute of Science and Technology, Nellore-Bombay Highway, S.P.S.R, Andhra Pradesh 524137, India

*Correspondence: P. C. N. Lalithya

ABSTRACT

Cervical cancer remains a significant global health challenge and is one of the leading causes of cancer-related mortality among women, particularly in developing regions where access to early screening and diagnostic facilities is limited. Early detection plays a crucial role in improving survival rates, with over 90% success in early-stage diagnosis. However, conventional histopathological examination is time-consuming, labour-intensive, and subject to inter-observer variability, leading to inconsistent results. To address these challenges, this study proposes a Deep Ensemble Learning (DEL) approach named Cervical Cancer Detection Network (CVD-Net). Initially, histopathology images undergo preprocessing techniques including resizing, normalization, contrast enhancement, and noise reduction to improve image quality. Traditional machine learning models such as Logistic Regression (LR), Gaussian Naive Bayes (GNB), and Quadratic Discriminant Analysis (QDA) are used as baseline classifiers. The proposed method leverages Convolutional Neural Network (CNN)-based feature extraction combined with a Random Forest Classifier (RFC) to automatically learn high-level spatial and textural features, minimizing reliance on handcrafted features. The model classifies images into four categories: Squamous Cell Carcinoma (SCC), Negative for Intraepithelial Malignancy (NILM), Low-grade Squamous Intraepithelial Lesion (LSIL), and High-grade Squamous Intraepithelial Lesion (HSIL). Experimental results demonstrate superior performance of the proposed approach, achieving an accuracy of 97.5%, precision of 97.9%, recall of 96.6%, and F1-score of 97.1%. These findings highlight the effectiveness of the proposed system for accurate and early detection of cervical cancer, potentially aiding clinical decision-making and reducing diagnostic errors.

Keywords: Cervical Cancer, CVD-Net, Convolutional Neural Network, Random Forest Classifier, Deep Ensemble Learning.

1. INTRODUCTION

According to the World Health Organization (WHO), the most common causes of untimely death among women across the world are malignancies of the female reproductive tract, including breast, cervical, and ovarian cancer. Cervical cancer, also known as cancer of the cervix (the lowermost section of the uterus), is a malignant tumour that develops when tissue cells covering the cervix begin to expand and multiply uncontrolled without following the appropriate mechanism for cell division. This research proposes a new approach for predicting cervical cancer given the preliminary

resulting of screening from individual medical records. The proposed prediction model would help in the intelligent detection of cervical cancer using deep learning and swarm optimization approaches. Fig. 1 illustrates the burden of cervical cancer in India using Disability-Adjusted Life Years (DALYs) age-standardized rates (ASR) per 100,000 population, clearly showing strong regional variation. The North-East region records the highest burden with a DALYs ASR of 290.1 per 100,000, indicating a significantly higher impact of cervical cancer compared to other regions. This is followed by the Central region

with 276.9 and the North region with 276.7, both reflecting a high disease burden and substantial loss of healthy life years due to cervical cancer.

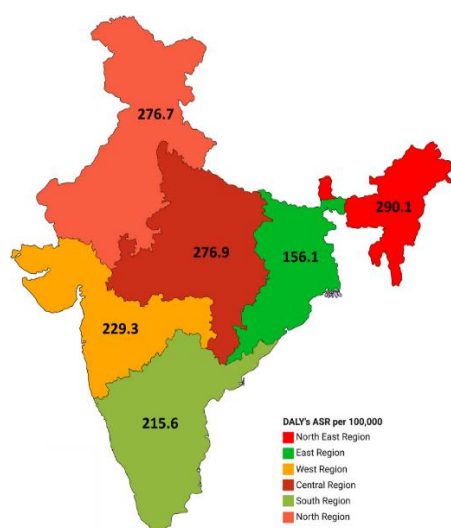


Fig. 1: Burden of cervical cancer in India.

2. LITERATURE SURVEY

Alzakari et al. [1] focused on a major global health concern, emphasizing the high mortality rate of cervical cancer in less-developed countries. It introduced an early classification approach that combined a hybrid feature selection algorithm (BWWPAPSO) with an optimized MLP classifier (WWPAPSO+MLP). The method was tested using a publicly available dataset that underwent preprocessing with SMOTE-Tomek to manage missing values and class imbalance. The results showed notable improvements in specificity, sensitivity, and accuracy. Statistical tests, including the Wilcoxon signed-rank test and ANOVA, further confirmed the effectiveness of the proposed approach. Meenakshi Sundaram A et al. [2] emphasized cervical cancer as the third most common cancer and highlighted the need for early detection, noting its strong link to HPV and the limitations of Pap smear tests due to human error. While machine learning models had been used for prediction, their performance declined with large datasets. To improve accuracy, the researchers introduced a hybrid deep learning model combining three CNNs with an averaging method. The ensemble

approach delivered better diagnostic performance, achieving an accuracy of 0.96 compared to 0.95, 0.91, and 0.95 for RF, SVM, and DT.

Renan Chaves de Lima et al. [3] showed that deep learning models offered promising solutions for cervical cancer screening in underserved areas like the Brazilian Amazon, where limited healthcare access left vulnerable populations at high risk. It highlighted that AI can help detect the disease earlier, provide faster and more accurate diagnoses, and ultimately reduce mortality. The study also stressed the need to adapt these models to local conditions to make them truly effective and ensure that more people can access quality healthcare. Foziya Ahmed Mohammed et al. [4] developed a hybrid model combining a SWIN Transformer and a CNN, supported by transfer learning, was developed to classify precancerous colposcopy images. The analysis was conducted on 898 high-quality images obtained from the IARC Colposcopy Image Bank, with normal and precancerous cases divided into training and testing groups. By integrating CNN-based local feature extraction with the SWIN Transformer's ability to capture global context, the model achieved more accurate classification. The consistently strong performance across cross-validation and independent test sets highlighted the model's potential to support early and reliable detection of precancerous cervical lesions.

Blanca Vazquez et al. [5] found that machine learning and deep learning had been widely applied to predict, diagnose, and assess the prognosis of cervical cancer, with most studies focusing on diagnostics. It noted that much of this research came from China, the United States, India, South Korea, and Japan, and that nearly half of the studies relied on deep learning with image-based data. While these AI approaches showed great potential for clinical use, the study emphasized that more validation and reproducibility were needed before they were reliably applied for early detection,

prognosis, and treatment of cervical cancer. Pooja Patre et al. [6] showed that deep learning had great potential for improving the detection and classification of cervical cancer, especially using CNNs to analyse cervical cytology images. They highlighted that model performance was measured through metrics like accuracy, sensitivity, specificity, and AUC, while challenges such as limited annotated datasets and inconsistent medical imaging remained. The study also pointed out that approaches like data augmentation, transfer learning, and semi-supervised learning had been explored to tackle these challenges and support future research and clinical applications.

Mariangel Rodríguez et al. [7] examined how deep learning can support automated cervical cell classification using both Pap smear and liquid-based cytology (LBC) images. By applying a new segmentation approach to isolate individual cells, a ResNet-50 model was trained and evaluated. The model showed excellent performance on LBC images, achieving high sensitivity, specificity, and accuracy, while results on Pap smear images were weaker due to noise and poorly defined cells. Overall, the findings underscored the promise of AI-driven screening with LBC and highlighted the importance of explainable AI to encourage clinical trust and adoption in resource-limited settings. Md. Sadekur Rahman et al. [8] explored deep learning techniques, including D-CNNs, transfer learning, and a new hybrid ensemble model called AZL, to automate the detection and classification of cervical cancer. After carefully preprocessing and augmenting the dataset, the AZL model, which combined AlexNet, ZfNet, and LeNet, achieved an impressive accuracy of 99.92%, outperforming individual models in precision, recall, and F1-score. These findings showed that ensemble deep learning can significantly improve diagnostic accuracy, helping pathologists detect cervical cancer more quickly and effectively, particularly in areas with limited resources.

Zongjian Lv et al. [9] introduced a hybrid data-processing strategy that combined CTGAN with traditional machine learning. The method used mRMR and XGBoost-RFE to refine the most meaningful features, SMOTE-ENN to correct class imbalance, and CTGAN to enrich the dataset with additional samples. After applying this strategy to the RFCC dataset, the LR model delivered strong performance, achieving 99.00% accuracy, 99.28% precision, 98.77% recall, and a 99.02% F1-score. Khandaker Mohammad Mohi Uddin et al. [10] developed a predictive model was designed using early screening information and individual health records to better understand key risk factors. Various machine learning techniques were applied, along with feature selection, dimensionality reduction, and oversampling methods, to manage missing data and class imbalance. The results showed that an ensemble model combining RF and MLP through hard voting delivered the best performance, highlighting its effectiveness in accurately predicting outcomes despite challenging data conditions.

Aymen M Al-Hejri et al. [11] developed a hybrid framework combining Vision Transformers (ViT) with ensemble CNNs to classify cervical cancer from Pap smear images. Tested on the Mendeley LBC and SIPaKMeD datasets with nine cell types, features from DenseNet201, Xception, and InceptionResNetV2 were fused and processed through a ViT-based encoder. The model achieved 97.26% accuracy on Mendeley LBC, 99.18% on SIPaKMeD, and 95.10% on the combined dataset, with Grad-CAM-based XAI techniques enhancing clinical interpretability. Eman Hussein Alshdaifat et al. [12] compared the accuracy of CT and MRI using a hybrid model that combined the Large Vision Model (LVM) with InternImage, while the Shark Optimization Algorithm (SOA) dynamically selected the best model weights. InternImage focused on deep tumour-specific features, and LVM captured fine-grained spatial patterns to detect early pathological changes. The model

was tested on the KAUH-CCTD and KAUH-CCMD datasets, classifying images as benign, malignant, or normal. It achieved 98.49% accuracy on CT images and 92.92% on MRI, highlighting the important role of imaging in treatment planning and monitoring patient progress.

3. PROPOSED SYSTEM

The system as shown in Fig. 2 presents a DEL based CVD-Net framework that automates the classification of histopathology cervical images into four categories: High squamous intra-epithelial lesion (HSIL), Low squamous intra-epithelial lesion (LSIL), Negative for Intraepithelial malignancy (NILM), Squamous cell carcinoma (SSC). The system begins with comprehensive image preprocessing to enhance visual quality and suppress noise, ensuring reliable feature learning.

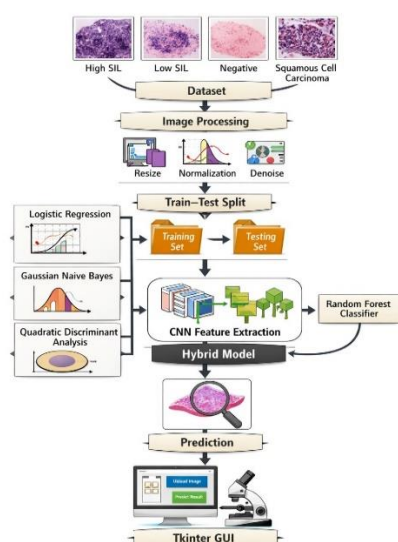


Fig. 2: Proposed system architecture.

While traditional machine learning classifiers such as LR, GNB, and QDA are implemented as baseline models for comparative evaluation, the core contribution lies in the proposed hybrid model. This model utilizes a CNN for automatic extraction of high-level spatial and textural features from medical images, eliminating dependence on handcrafted features. These deep features are subsequently classified using an ensemble RFC, which improves robustness,

generalization capability, and classification accuracy. The entire framework is integrated with a Tkinter-based graphical user interface, enabling user-friendly interaction and real-time cervical cancer prediction, thereby supporting accurate and early diagnosis.

Dataset: The dataset used in this study is categorized into four clinically significant classes: HSIL, LSIL, NILM, SSC. These categories correspond to progressive stages of cervical epithelial abnormality, as defined by medical standards. Normal images represent healthy cervical tissues with regular cellular structures, while HSIL, LSIL, NILM and SSC indicate low-grade and high-grade cervical intraepithelial neoplasia, respectively. The Cancer class includes malignant tissues exhibiting severe cellular deformation. Accurate class definition is critical, as the morphological differences between these stages are often subtle, making automated classification both challenging and clinically valuable.

Image Processing: Image preprocessing plays a vital role in improving the quality and interpretability of histopathology images. This step includes resizing all images to a uniform dimension to maintain consistency during model training. Normalization is applied to scale pixel intensities, ensuring numerical stability and faster convergence of deep learning models.

Train-Test Splitting: To ensure unbiased evaluation and robust learning, the dataset is divided into training and testing subsets using the train_test_split technique. The training set is utilized to learn model parameters and optimize weights, while the testing set is reserved exclusively for performance evaluation on unseen data. This approach prevents data leakage, reduces the risk of overfitting, and provides a realistic assessment of how the proposed system would perform in real-world clinical scenarios.

Existing LRC: LR is implemented as a baseline classifier due to its simplicity and interpretability. It models the relationship between input features and output classes using a linear decision boundary. Although it may struggle with complex nonlinear patterns, LR provides a fundamental benchmark to assess the effectiveness of more advanced classifiers in distinguishing between cervical cancer stages.

Existing GNB Classifier: The GNB classifier is employed as a probabilistic baseline model that assumes feature independence and Gaussian distribution. Despite its simplifying assumptions, it is computationally efficient and performs well on small datasets. Its inclusion helps evaluate how probabilistic learning compares with discriminative and ensemble-based methods for cervical cancer detection.

Existing QDA: QDA is applied to capture nonlinear class boundaries by estimating class-specific covariance matrices. This allows QDA to better model complex data distributions, particularly when classes overlap. In the context of CVD-Net, QDA is useful for differentiating intermediate stages such as CIN1 and CIN2, where subtle variations exist.

Proposed CNN Feature Extraction with RFC: The proposed hybrid model integrates the strengths of deep learning and ensemble learning to overcome the limitations of traditional approaches. A CNN is used to automatically extract hierarchical and discriminative features from raw histopathology images. These features are then classified using a RF ensemble, which improves robustness, reduces variance, and enhances overall predictive accuracy. This combination ensures reliable performance even in the presence of noisy or imbalanced data.

Prediction From Test image: The test unknown sample images is considered for prediction. Then, the deep features extracted by the CNN are provided as input to the RFC. RF constructs multiple decision trees using random subsets of features and training samples, and the

final classification is obtained through majority voting. This ensemble strategy minimizes overfitting, improves generalization, and enhances classification stability across all cervical cancer stages.

Tkinter Frontend: A graphical user interface is developed using the Tkinter library to facilitate user interaction with the proposed system. The frontend allows users to upload histopathology images, initiate preprocessing and classification, and view predicted results instantly. This interface enhances usability and demonstrates the practical applicability of the system in clinical decision support and academic environments.

Proposed CNN Feature Extraction with RFC

In CVD-Net, the proposed CNN feature extraction with RFC integrates deep learning and ensemble learning to accurately classify cervical tissue images into clinically significant diagnostic categories as shown in fig. 3. The input consists of training and testing datasets derived from pre-processed cervical histopathology images along with their corresponding labels. The CNN automatically learns high-level spatial, structural, and textural features from cervical images, reducing dependency on handcrafted features. These extracted deep features are then supplied to RFC, which performs robust classification by combining multiple decision trees. The output of the method is the reliable classification of cervical samples into HSIL, LSIL, NILM, or SCC.

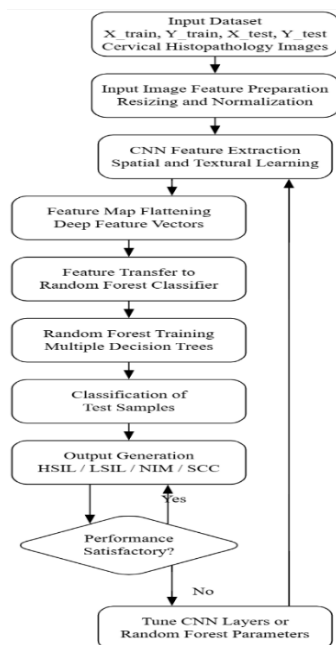


Fig. 3: Proposed CNN feature extraction with RFC.

Input Image Feature Preparation: In this step, pre-processed cervical histopathology images from the training and testing datasets are prepared as input to the CNN. These images are resized and normalized to meet the input requirements of the convolutional network while preserving diagnostically relevant tissue structures.

Convolutional Feature Extraction: The CNN processes the input images through multiple convolution and pooling layers to learn hierarchical feature representations. Lower layers capture basic patterns such as edges and textures, while deeper layers extract complex morphological features that distinguish different cervical cancer stages.

Feature Map Flattening and Selection: The high-level feature maps obtained from the CNN are flattened into feature vectors. These vectors represent compact and informative descriptors of cervical tissue characteristics, suitable for subsequent machine learning classification.

Feature Transfer to RF: The extracted CNN feature vectors from the training set are transferred to the RFC. This step bridges deep learning feature learning with ensemble-based decision-making, enabling effective handling of nonlinear feature interactions.

RF Training: The RFC constructs multiple decision trees using different subsets of features and samples. Each tree learns decision rules that separate cervical cancer classes, and ensemble voting enhances robustness and reduces overfitting.

Classification of Test Samples: CNN-extracted features from the testing dataset are passed through the trained RF model. Each decision tree predicts a class label, and the final classification is determined through majority voting across all trees.

Output Generation and Evaluation: The final output consists of predicted diagnostic categories for each cervical image, including HSIL, LSIL, NILM, or SCC. The model's performance is evaluated using testing data to ensure reliable and accurate cervical cancer detection.

4. RESULTS ANALYSIS

Fig. 4 demonstrates near-perfect classification performance. The model correctly classified all 45 HSIL, 45 LSIL, and 40 NILM samples with zero misclassifications, while in the SCC class, 26 samples were correctly identified with only 4 misclassified as HSIL. The ROC-AUC values achieved are 1.00 for HSIL, LSIL, NILM, and SCC, with curves reaching the top-left corner, indicating perfect classification with maximum true positive rate and zero false positive rate, thereby confirming the superior discriminative capability of the proposed CNN-RFC model.

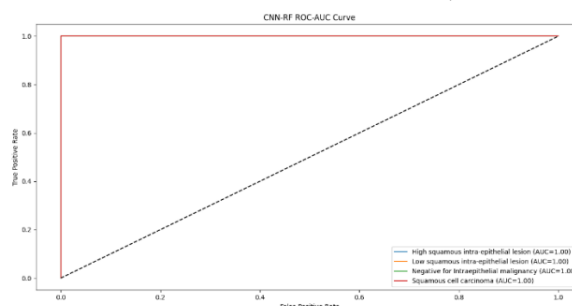
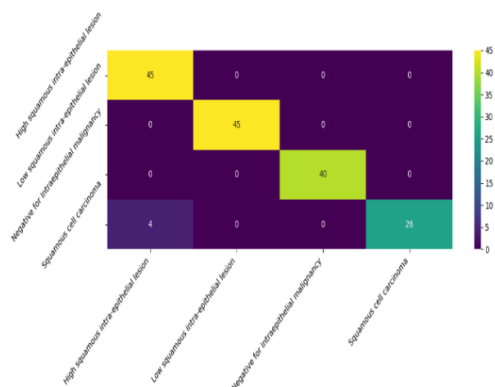


Fig. 4: Confusion matrix and ROC curve of proposed CNN-RFC model.

Comparative analysis

Table 1 shows the overall performance comparison of different CVD methodologies in terms of Accuracy, Precision, Recall, and F1-Score. The proposed CVD-Net with CNN-RF model achieves the highest performance with an Accuracy of 0.9750, Precision of 0.9796, Recall of 0.9667, and F1-Score of 0.9715, indicating superior classification capability and balanced predictive performance. Among the conventional models, LR demonstrates strong results with an Accuracy of 0.8938, Precision of 0.8907, Recall of 0.8778, and F1-Score of 0.8802. GNB shows moderate performance with an Accuracy of 0.6500, Precision of 0.6715, Recall of 0.6569, and F1-Score of 0.6471. Meanwhile, QDA records the lowest metrics, with an Accuracy of 0.5375, Precision of 0.6061, Recall of 0.4958, and F1-Score of 0.4771. Overall, the comparison clearly highlights the effectiveness of the proposed hybrid CNN-RF model over the traditional machine learning approaches.

Table 1 Overall comparison table of CVD methodologies.

Model	Acc urac y	Pre cisi on	Re ca ll	F1- Sco re
LR	0.8938	0.8907	0.8778	0.8802

GNB	0.6500	0.6715	0.6569	0.6471
QDA	0.5375	0.6061	0.4958	0.4771
Proposed CVD-Net with CNN-RF	0.9750	0.9796	0.9667	0.9715

Fig. 5 Prediction results of the proposed Hybrid CNN-RF model for a test histopathology image of the HSIL class. The left panel shows the original cervical cell image, while the right panel presents the predicted output, where the model correctly classifies the sample as HSIL, indicating significant epithelial abnormalities. The XAI analysis enhances interpretability by confirming the image as cervical tissue (is_cervical: True) and identifying the image type as cervical. It also detects the presence of disease (disease_present: True) with severity classified as severe. Additionally, the affected region is identified as the transformation zone, which is highly susceptible to precancerous changes. These findings demonstrate the model's accuracy and its ability to provide clinically meaningful explanations for reliable decision-making.



Fig. 5: Prediction results on HSIL.

Telegram Bot Results

Fig. 6 shows the Telegram Bot interface used for uploading and analysing cervical histopathology images within the proposed Hybrid CNN–RF framework. The chat window displays the initialization command “/start”, followed by the bot’s response requesting the user to “Send a cervical leaf image” and listing its detection capabilities: Cervical disease detection, Severity assessment, and AI model prediction.



Fig. 6: Telegram bot with image input.

A cervical cytology image is then uploaded by the user, and the bot responds with the status message “Analysing...”, indicating that the image is being processed by the backend model. This figure demonstrates the real-time integration of the trained hybrid model with a Telegram Bot interface, enabling automated image submission, instant analysis, and user-friendly remote access for cervical cancer screening and decision support.

5. CONCLUSION

The experimental results clearly demonstrate that the proposed CNN-RF hybrid model significantly outperforms the traditional machine learning classifiers across all

evaluation metrics. The model achieved the highest Accuracy of 0.9750, along with Precision of 0.9796, Recall of 0.9667, and F1-Score of 0.9715, indicating excellent classification capability and balanced predictive performance. In comparison, LR achieved an Accuracy of 0.8938, while GNB and QDA recorded lower accuracies of 0.6500 and 0.5375, respectively. The macro average values further confirm the superiority of the proposed model, with 0.98 Precision, 0.97 Recall, and 0.97 F1-Score, compared to 0.89, 0.67, and 0.61 macro precision values for LR, GNB, and QDA, respectively. Class-wise analysis also highlights the robustness of the CNN-RF model, achieving perfect recall of 1.00 for HSIL, LSIL, and NILM classes, and 0.87 for SCC, which was the most challenging category. Similarly, precision reached 1.00 for LSIL, NILM, and SCC, with 0.92 for HSIL. The micro average score of 0.9750 confirms its overall reliability in multi-class classification. These results validate that combining deep feature extraction using CNN with ensemble classification through RF enhances diagnostic accuracy and reduces misclassification, especially in critical cervical cancer detection tasks.

REFERENCES

- [1] Alzakari SA, Alhussan AA, Towfek SK, Metwally M, Salem DA. Cervical Cancer Detection Using Deep Neural Network and Hybrid Waterwheel Plant Optimization Algorithm. *Bioengineering (Basel)*. 2025 Apr 30;12(5):478. Doi: 10.3390/bioengineering12050478. PMID: 40428097; PMCID: PMC12108694.
- [2] Aishwarya N. Kumar and Meenakshi Sundaram A., "A Hybrid Ensemble Learning Approach for Cervical Cancer Detection: Combining Multiple CNN Models for Enhanced Diagnostic Accuracy," *Journal of Advances in Information Technology*, Vol. 16, No.

- 6, pp. 893-903, 2025. Doi: 10.12720/jait.16.6.893-903
- [3] De Lima, R.C.; Quaresma, J.A.S. Exploring Deep Learning Model Opportunities for Cervical Cancer Screening in Vulnerable Public Health Regions. *Computers* **2025**, *14*, 202.
- [4] Mohammed FA, Tune KK, Mohammed JA, Wassu TA, Muhie S. Early Cervical Cancer Diagnosis with SWIN-Transformer and CNNs. *Diagnostics* (Basel). 2024 Oct 14;14(20):2286. doi: 10.3390/diagnostics14202286. PMID: 39451609; PMCID: PMC11506970.
- [5] Banda Saikumar. (2025). Integrating azure network rules for storage account through terraform in CI/CD pipelines: automating storage account access restrictions to public IP. *Journal of Science & Technology*, 10(2), 15–22. <https://doi.org/10.46243/jst.2025.v10.i02.pp15-22>.
- [6] Reddy, S. K. R. (2024). Designing Blockchain Architecture to Transform Loyalty Rewards into Cryptocurrency Investments.
- [7] Patre P, Verma D. Deep dive into deep learning methods for cervical cancer detection and classification. *Rep Pract Oncol Radiother*. 2025 Aug 7;30(3):396-416. Doi: 10.5603/rpor.106148. PMID: 40919255; PMCID: PMC12413234.
- [8] Rodríguez, M.; Córdova, C.; Benjumeda, I.; San Martín, S. Automated Cervical Cancer Screening Using Single-Cell Segmentation and Deep Learning: Enhanced Performance with Liquid-Based Cytology. *Computation* **2024**, *12*, 232.
- [9] Abu Hanzala*, Tanjila Akter, Md. Sadekur Rahman. Department of Computer Science and Engineering, Daffodil International University, Daffodil Smart City, Birulia, 1216, Dhaka, Bangladesh
- [10] Mengdi Tang, Hua Chen, Zongjian Lv and Guangxing Cai. Diagnosis of Cervical Cancer Based on a Hybrid Strategy with CTGAN. *Electronics* **2025**, *14*, 1140.
- [11] Khandaker Mohammad Mohi Uddin, Abdullah Al Mamun, Aditi Chakrabarti, Rafid Mostafiz, Samrat Kumar Dey. An ensemble machine learning-based approach to predict cervical cancer using hybrid feature selection. Volume 4, issue 3 September 2024, 100169
- [12] Todupunuri, A. (2025). The Role of Human-Centric AI in Building Trust in Digital Banking Ecosystems. *SSRN Electronic Journal*. <https://doi.org/10.2139/ssrn.5120605>
- [13] Kalae, U. K. (2020). Developing scalable Power BI dashboards for enhanced data analysis and strategic business decision-making. *International Journal of Enhanced Research in Science, Technology & Engineering*, 9(3), 8–15.
- [14] Al-Hejri AM, Al-Tam RM, Sable AH, Almuhaya B, Alshamrani SS, Alshmrany KM. A hybrid vision transformer with ensemble CNN framework for cervical cancer diagnosis. *BMC Med Inform Decis Mak*. 2025 Nov 5;25(1):411. Doi: 10.1186/s12911-025-03250-x. PMID: 41194237; PMCID: PMC12590815.
- [15] Alshdaifat EH, Sindiani AM, Alhatamleh S, Abu Mhanna HY, Madain R, Amin M, Malkawi M, Jaradat A, Akhdar HF, Gharaibeh H, Maashey F, Alghulayqah L. Enhancing cervical cancer diagnosis with ensemble learning and shark optimization algorithm: comparative study of CT and MRI in cervical cancer diagnosis. *Front Oncol*. 2025 Oct 10; 15:1608386. Doi: 10.3389/fonc.2025.1608386. PMID: 41142601; PMCID: PMC12550778.