



Research Article

An advanced framework for cervical cancer detection using optimized ensemble learning

Rufina Sherin John ROBERT¹, Tamilazhagan SENTHIL¹, Manikandan RAMACHANDRAN¹,
Vidivelli SOUNDARARAJAN¹, Magesh SANKAR², Mohammad KHISHE^{3,4,*}

¹School of Computing, SASTRA Deemed University, Tamil Nadu, 613401, India

²Department of Computer Science and Engineering, Dr. M. G. R. Educational and Research Institute, Chennai, 600095, India

³Applied Science Research Center, Applied Science Private University, Amman, 11937, Jordan

⁴Department of Electrical Engineering, Imam Khomeini Naval Science University of Nowshahr, Nowshahr, 4651763836, Iran

ARTICLE INFO

Article history

Received: 05 August 2024

Revised: 03 October 2024

Accepted: 27 May 2025

Keywords:

Cervical Cancer Detection;
Convolutional Features; Ensemble
Learning; Neural Networks;
Machine Learning

ABSTRACT

Cervical cancer remains the most common and life-threatening malignancies around the world. As with many diseases, timely and accurate intervention is critical in this case if one wishes to reduce the mortality rate associated with this medical condition. The purpose of this research is to develop an innovative computational framework to detect cervical cancer using both quantitative and data-driven approaches. The primary data set consists of the demographics, lifestyle choices, and medical history of 980 patients, which was collected at the Hospital Universitario de Caracas in Venezuela. For reproducibility purposes and computational efficiency, all model testing was performed on Google Colab, a cloud platform suitable for big data. The ensemble learning methodology integrated a convolutional neural network (CNN) with several machine learning classifiers, including Random Forest, Decision Tree, Logistic Regression, K-Nearest Neighbors, Extreme Gradient Boosting, Support Vector Machine, Extra Tree Classifier, Categorical Boosting, Gradient Boosting, and Adaptive Boosting. This paper also looks at the impact of model performance on original and engineered feature selection. After extensive comparative evaluation, the proposed framework showed an extraordinary predictive performance of 98.96% classification accuracy, thus confirming the value of the framework. This showcases the enormous potential of machine learning in revolutionizing cervical cancer diagnosis. This study seeks to refine diagnostic precision in order to enable timely detection while minimizing misdiagnosis risks. Through integrated datasets, the framework deepens understanding of the influencing factors and enhances the accuracy of the predictions, providing more precise forecasts. This technology seeks to improve patient prognosis by providing more timely and effective cervical cancer diagnosis and treatment.

Cite this article as: Robert RSJ, Senthil T, Ramachandran M, Soundararajan V, Sankar M, Khishe M. An advanced framework for cervical cancer detection using optimized ensemble learning. Sigma J Eng Nat Sci 2026;44(4):2458–2472.

*Corresponding author.

*E-mail address: m_khishe@alumni.iust.ac.ir

This paper was recommended for publication in revised form by
Editor-in-Chief Ahmet Selim Dalkilic



INTRODUCTION

Cervical cancer persists as a major public health issue, particularly impacting women in their childbearing age. This cancer develops from the cervix, which is the cylindrical-shaped lower part of the uterus leading into the vagina. The primary driver of cervical cancer is a long-term infection with high-risk HPV strains, a virus that is sexually transmitted. While most of the HPV infections resolve on their own and are non-threatening, some are persistent and resolve more slowly [1, 2]. These precancerous lesions, over time, may lead to invasive cervical cancer.

One of the most ancient and prevalent viral infections in humans is HPV, with infected individuals all over the globe [3]. As stated by the WHO, cervical cancer is the fourth most common cancer amongst women in the world. New case counts for the disease reached approximately 604,000 with 342,000 deaths in 2020 alone. Of great concern, almost 90% of the new cases and deaths were from low-middle income nations. This is a phenomenon driven by a lack of robust healthcare systems and the absence of integrated, tailored cancer prevention programs.

HPV types 16 and 18 are the most oncogenic strains, precipitating approximately 70% of cervical cancers. The risk of infection and disease advancement may increase due to various factors: early onset of sexual activity, multiple sexual partners, long-term contraceptive use, smoking, and HIV co-infection. Disease advancement at the molecular level is defined by the integration of HPV DNA into the host genome, resulting in the overexpression of viral oncogenes E6 and E7. E6 and E7 oncoproteins hijack essential tumor suppressors such as p53 and retinoblastoma (Rb), leading to uncontrolled cell division and subsequently, malignant transformation [4].

Cervical screening and HPV vaccination have shown significant proactive results. Precancerous changes can be detected early via Pap smear cytology along with DNA HPV testing. Developed countries have reported a significant decline in the incidence and mortality of cervical cancer due to proactive screening programmes [5]. In resource-restricted countries, earlier diagnosis is rare, which leads to poor outcomes and higher mortality. To address this gap, global health strategies need to prioritize the strengthening of healthcare systems, the launching of mass education HPV vaccination campaigns, and the elevation of awareness concerning risk factor reduction [6,7]. The burden of cervical cancer and women's health inequities on a global scale can only be alleviated through effective, affordable, and culturally appropriate policies and programs. The stigma of cervical cancer and poor awareness among women deters them from seeking preventive care or early diagnosis. It is essential, therefore, to counter these hindrances with innovative, cost-effective solutions [8]. Acetic acid visual inspection is much less expensive than Pap smears, and luminations have shown promise in detecting cervical abnormalities, especially well-suited in resource-limited settings [9].

Mobile health clinics and outreach programs bring screening services to the underserved, remote areas, thus increasing access. Digital health technologies, from telemedicine to AI-supported diagnoses, can benefit from a complete turnaround by allowing patients to leap over the barrier of distance to reach specialist care. These offer the possibility of remote assessments with early detection of conditions in places lacking access to healthcare services. All education and awareness work about the value of protection against cervical cancer and getting proper Pap tests should be targeted at educating women. Community health workers can also serve as important channels for such education efforts and encouragement of community members to participate in screening programs.

Conventional cervical cancer screening methods comprise Pap smears, in addition to HPV tests, which call for manual interpretation. Certainly time-consuming, prone to human error, and ineffective in low-resource contexts. Vaccinating against HPV is another essential preventive step [10]. Cervarix and Gardasil vaccinations defend against most widespread speculative-risk HPV varieties. The implementation of these vaccinations has yielded positive results in reducing the occurrence of HPV infections and cervical cancer. As per the WHO, girls between the ages of 9 and 14 ought to receive a routine HPV vaccination prior to starting sexual activity [11].

Early signs of the disease do not present themselves. With the ever-changing nature of data, It Is pertinent to understand that cervical cancer screening requires HPV and PAP smear testing and regular follow-up visits so that it can be addressed in as timely a manner as possible [40]. The more advanced stages of the disease are associated with increased discomfort in the pelvic area, along with pain during intercourse. Each Individual patient has their treatment plan, which may include surgical treatment combined with radiotherapy or plasma chemotherapy in more advanced patients. Recently, some of the more specific and compelling recurrent or metastatic cervical cancer treatment options, such as immunotherapy, have emerged with much more favorable and efficient regard for achieving functional outcomes. As with any complicated prediction, there is always some stage and combination of treatments that changes it. For cervical cancer, it is highly dependent. However, early detection increases the five-year survival chance of a patient to roughly ninety-two percent [12, 13].

Comparably, the chances of survival with cancer detected at an advanced stage drop sharply to seventeen percent. In the United States, the American Cancer Society data shows that 66 percent of women with cervical cancer are expected to live beyond five years post-diagnosis. These figures demonstrate the importance of effective treatment and intervention to reduce mortality and increase survival rates for patients diagnosed with cervical cancer. Early detection is made possible by combining approaches utilizing ensemble methods, ensuring improved confidence in diagnostics forming within the disjunctions of data

collected through mechanisms such as HPV tests and Pap smears [14]. Feature selection is so vital as it emphasizes explanatory factors to reasonably predict diagnosis of cervical cancer.

Machine Learning models deal with vast volumes of patient data. They are able to recognize very subtle patterns that a human being might overlook, thus leading to more excellent detection rates. An ML-based approach can lead to greater diagnostic accuracy and fewer instances of false positives and negatives, with a view to furthering the clinical decision support system towards timely and effective interventions. Ensemble techniques include bagging, boosting, and stacking, depending on the desired application to enhance performance. The ensemble learns from the dataset and evaluates the model's performance in terms of accuracy and F1 score. Machine Learning models deployed in automated systems provide high-quality diagnostic support to those areas of the world with limited access to expert pathologists. Once trained and validated, this model would now provide risk of cervical cancer in consideration of new patients and recommendations for any follow-up testing. It is possible to incorporate machine learning into the existing healthcare structures in the fight against cancer. Thus, this may help clinicians determine the prognosis and design treatment plans personalized to the patient [41, 42]. A highly reliable and corroborated method toward an enhanced diagnosis enables timely intervention, finalizing better patient outcomes. Continuous monitoring and retraining imply that the model remains working well over time.

Related Works

These recent advances in cervical cancer (CC) detection and diagnosis, as well as active research on the use of technology, have identified the emerging areas of technology like computer-aided diagnosis (CAD) and advanced diagnostic support systems [15]. Researchers have developed automated diagnosis of cervical cancer through individual and ensemble models based on machine learning, improving accuracy of machine learning when deploying models. In these studies, multiple classifiers were tested, with genetic preprocessing and feature selection in genetic programming producing excellent results that improved diagnostic accuracy. Several studies have contributed to increasing cervical cancer detection and understanding the link between this cancer and HPV. For instance, after Munoz et al. (2006) researched worldwide HPV prevalence, they helped forge a crucial association between HPV infection and cervical cancer while simultaneously aiding vaccine development. They furthered the evaluation of various screening strategies, too. According to Arbyn et al. (2012), HPV DNA testing is much more sensitive (90–95%) than cytology (70–80%). However, trials utilized by Ronco et al. (2014) in Europe have confirmed that HPV testing has superior early detection sensitivity of 94 and specificity of 98 percent when compared to cytology. We identify broad

disparities in incidence and outcome of cervical cancer that compare high-income with treatment inequalities amongst low to middle-income countries as part of Bruni et al. (2020) global cancer registry data analysis. Newer treatment modalities in Marth et al.'s (2017) clinical trial assessment of targeted therapies for advanced cervical cancer have shown promise for better response rates. Yu et al. (2021) extend diagnosis capabilities beyond the scope of both clinical images and the data to machine learning models.

Extensive evaluations of screening strategies have also been done. HPV DNA testing has much higher sensitivity (90 – 95%) than cytology (70 – 80%), as shown by Arbyn et al. (2012). However, Ronco et al. (2014) trials in Europe confirmed that HPV testing performs better than cytology in early detection, with sensitivity of 94% and specificity of 98%. This is from Bruni et al, 2020. An international nagging cancer registry study pointed out the sharp differences in the incidence and outcomes of cervical cancer in relation to their lower- and middle-income peers within high income countries. It emphasized the lack of timeliness in the provision of healthcare services, starkly highlighting the divide in the care gap based on economic classifications. In the clinical trial conducted by Marth et al. (2017), newly emerging treatment options for aggravated invasive cervical cancer showed better response rates, giving hope for further treatment. Making a further contribution in the area, Yu et al. (2021) applied machine learning algorithms to clinical images and patient data which greatly improved the accuracy of the diagnosis and the possibilities of earlier detection.

MATERIALS AND METHODS

The following part explains the methodology phases, the suggested technique, and the dataset utilized to detect carcinoma of the cervix. The model workflow is shown in Figure 1. It also offers an overview of machine learning classifiers used in this study.

Dataset for Experiments

The dataset contains information related to cervical cancer screening outcomes and risk factors. This dataset is significant in research and analytical works, particularly in assessing and forecasting the occurrence of cervical cancer. It is organized in tables and saved as a CSV (Comma Separated Values) file, which provides easy access and allows it to be used in data-centric studies. The dataset includes various attributes (features) categorized into personal, demographic, and medical risk factors. The size of the dataset is 99.7 KB. Data has 36 variables comprising historical details and risk factors regarding a patient.

Demographic variables vary among women (age, reproductive history, sexuality), as well as medical variables (history of smoking and contraceptive use (both hormonal contraceptives and intrauterine devices (IUDs))). It also records previous diagnoses of other sexually transmitted

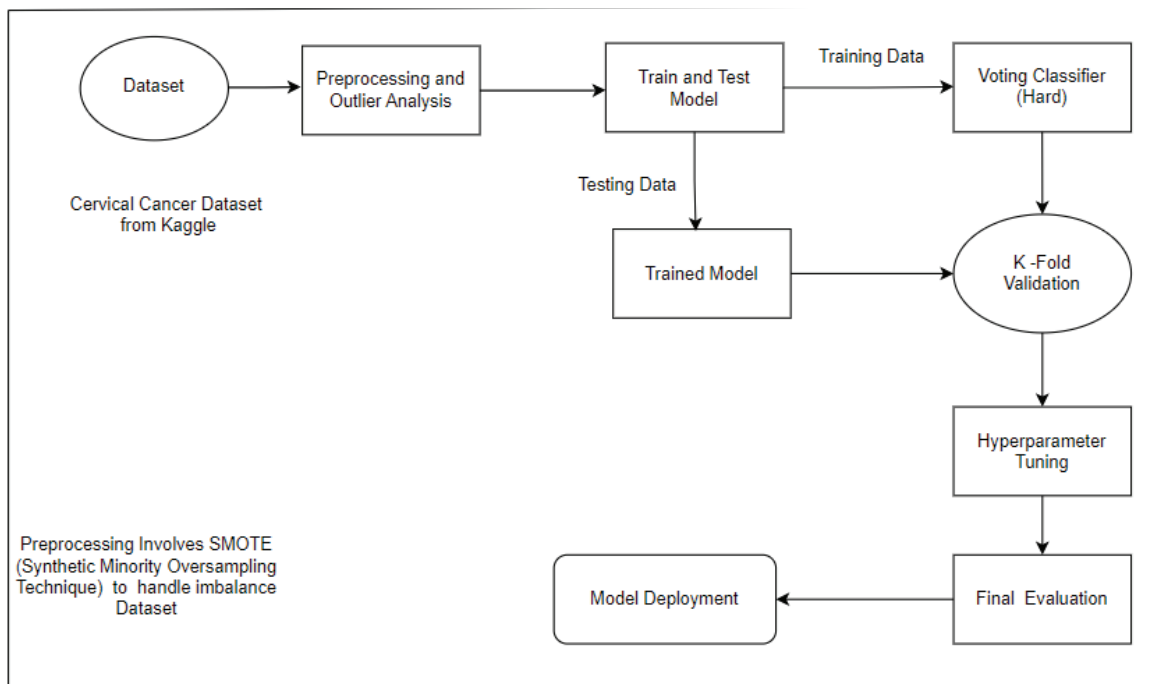


Figure 1. Proposed model workflow diagram.

infections, such as condylomatosis, syphilis, genital herpes, pelvic inflammatory disease, and vulva-perineal and vaginal condylomatosis. It also records less common infections like molluscum contagiosum, hepatitis B, HIV and HPV. The clinical parameters record the history of previously diagnosed STIs along with other relevant diseases like cervical or HPV cancer, CIN, and any other herpetic involvement. Moreover, the dataset contains biopsy results, which are necessary for determining relationships and trends that might be associated with or indicate the possibility of cervical cancer. Cervical cancer can be detected earlier, and its risks can be assessed through epidemiological research utilizing these datasets, which are gathered from systematic healthcare surveys or medical documents. This dataset is often derived from medical records or health surveys. It is used in medical research, particularly in studies focusing on cervical cancer prevention and early detection. Missing values are regular occurrences within real-world datasets, and utmost care should be taken to handle them efficiently to vouch for the validity of a data-based endeavor. The treatment of missing values may vary according to the nature of their manifestation, extent, and impact on the analysis. In case of missing values, replace them with the mean or median of the column and in case of categorical variables, replace missing values with the most frequent category.

Preprocessing and Machine Learning Models for Cervical Cancer Prediction

Supervised learning falls under machine learning, and it concerns training algorithms on a labeled dataset so that

new data points can be accurately classified. Classification tasks can be defined in terms of two main categories: binary classification, in which a model predicts one of two possible outcomes, and multi-class classification, in which a prediction can belong to multiple classes. For practical purposes, a well-known Python-based framework for machine learning and data mining – sci-kit-learn – provides an extensive collection of tools. The library is highly efficient for binary and multi-class classification problems with tools for classification such as decision trees, support vector machines (SVMs), and logistic regression [43].

Several preprocessing activities were performed before training machine learning models, with the objective of ensuring data quality and reliability. First, a missing value analysis was conducted, in which the original records in the dataset with missing values were not removed, as removing partial records results in data loss and possible biases; the treatment instead was a mean substitution. In this scenario, the mean substitution method computes the mean of a given numerical column, and imputes it into each missing value for that column. The Synthetic Minority Over-sampling Technique (SMOTE) is applied to medical datasets to tackle their class imbalance problem by oversampling the minority class through the creation of artificial samples. The SMOTE methodology generates artificial instances by making new points from existing ones. It is preferable to fix data with shared values that may enhance predictive power of models used when identifying a case from minority class; thus, analytical models have a substantive advantage when spotting instances belonging to the underrepresented class.

In addition, since a majority of machine learning algorithms require numerical input, it was necessary to convert categorical/non-numeric columns into numeric format. We utilized encoding techniques depending on the type - label encoding or one-hot encoding, to ensure that machine learning models can process and learn from all features [44, 45]. These preprocessors were imperative for model performance, fairness, and creating a robust cervical cancer detection framework.

Random Forest Classifier

Random Forest (RF) is a classic form of ensemble learning in classification and regression. It works by building many decision trees, each constructed from a random sample of features and the training data. In classification, RF combines the predictions by majority voting, and in regression, it averages the predictions from all the trees. This combination approach improves the precision of predictions and, on top of that, reduces overfitting [23]. Besides its accuracy and adaptability, RF is particularly useful when feature importance needs to be evaluated as well as when incomplete data sets are present. These factors make RF a preferred choice in numerous sectors and their machine learning applications.

Decision Tree

Along with classification problems, decision trees can also be employed for regression problems. A machine learning method, decision trees are quite powerful, as well as interpretable. Furthermore, they follow a structured hierarchy resembling trees where features serve as internal decision nodes, the branches illustrate particular featured decision rules and the leaves present the resulting outcomes or predictions. Such a model is easily visualized and comprehended, therefore supporting the effective and practical usage of decision trees. This systematic approach helps decision trees to systematically split data and make predictions based on a predetermined set of conditions. In decision trees, the patterns are revealed by sequentially splitting data into smaller groups, and these can be used for trend analysis. Despite their simple interpretation and practice, overfitting is relatively common when the data under consideration is complex, and risk of overfitting is thus relatively high [24, 25]. However, they can overfit, mainly when dealing with intricate datasets. Tightening up and figuring out the maximum depth are two strategies to help resolve this problem.

Logistic Regression

Logistic regression analysis is one of the statistical methods explicitly used in binary classification problems. It calculates the likelihood of an event occurring either above or below a threshold value based on one or more explanatory variables, which in this case are considered predictors. In this, a logistic function is applied to transform a linear combination of input variables into probability values between 0 and 1. As a model with good interpretability, logistic

regression allows it to establish a relationship between predictors and outcomes [26]. Moreover, it is computationally efficient and is optimal when the relation between independent and dependent variables follows a linear pattern. However, it can encounter difficulties in dealing with non-linear ties and is also vulnerable to multicollinearity among the predictors [27].

K-nearest neighbors

K-Nearest Neighbors (KNN) is a popular machine learning algorithm used for both supervised classification and regression tasks. Its core principle is built on the belief that data outcomes tend to be similar based on their spatial proximity to one another. KNN assigns labels to instances based on the most prevalent class of the k nearest neighbors (for classification) or computes the mean of their values (for regression). In finding closeness, distance metrics are used, with Euclidean distance being the most prevalent. KNN is described as a non-parametric method in that it makes no assumptions about the data structure; this property grants flexibility across numerous areas of application [28]. Although KNN is straightforward and efficient, it lacks scalability due to the ever-increasing amount of data. The computational work required to calculate distances for each prediction makes this a burdensome task.

Extreme Gradient Boosting

Regression and classification are combined using XGBoost, an optimized gradient-boosting technique. It is well known for its effectiveness and speed, and is one of the methods for collaborative learning. Following one another. XGBoost generates multiple decision trees that correct each other's mistakes. Gradient descent optimization techniques are used to minimize the loss function, resulting in a highly accurate prediction model [29]. Aside from giving features relevant to scores, XGBoost can manage missing values internally and simulate complex interactions between variables. Large datasets may necessitate extensive computer training, and hyperparameters may need to be finely tuned to minimize overfitting [30].

Support Vector Machine

Support Vector Machine (SVM) attempts to achieve the best classification for a given dataset by analyzing the optimum separation of features using hyperplanes, finding multiple areas classified as one type in regions separated from the other type. Dissimilar tasks and core duties can be managed effectively through SVM. SVM is useful in the case of classification of complex and high-dimensional datasets because it incorporates nonlinear relationships within the data using various types of kernels, including the radial basis function (RBF), polynomial, and linear kernels [31]. "In space, the SVM tries to find the appropriate optimal separation boundary which separates the classes or groups accurately." In simpler terms, an algorithm class overlap refers to where a specific portion of a cluster gets modified to try to fit into another higher-defined cluster. It

attempts to enhance classification accuracy. It is designed to reduce overfitting, which is a common problem in systems with high-dimensional spaces. On the downside, SVM requires intensive computations and is highly susceptible to the selection of kernel and its hyperparameters.

Extra Tree Classifier

The Extra Trees Classifier is multiformed as an ensemble learner, which is used for classification and regression tasks. Just like Random Forests, it builds many decision trees during the training process. However, in deeper random forests, Extra Trees randomly take both the split and the features for each split selection, which adds more diversity to the model. This greater level of variability improves the model's strength. It reduces the variance, making it work more optimally on complex datasets. Stronger model robustness and decreased variance are a result of this additional degree of unpredictability [32]. Compared to standard decision trees, this method is less prone to overfitting. It is capable of handling large, highly dimensional datasets. It works remarkably well. The best results with Extra Trees require careful adjustment of hyperparameters, but they offer excellent accuracy and computational efficiency.

Categorical Boosting

CatBoost, along with other gradient-boosting algorithms, are designed to handle categorical features well. It is well known for its ability to perform classification and regression tasks on categorical data without necessitating extensive preparation. CatBoost automatically encodes categorical features and applies sophisticated techniques to reduce overfitting and improve generalization. It is resistant to overfitting, handles missing values internally, and can be trained on massive datasets efficiently. CatBoost is also relatively efficient, often requiring fewer hyperparameter adjustments than other boosting algorithms. Due to its versatility and user-friendliness, it is a favorite among the variety of functions in machine learning applications [33].

Gradient Boosting

Gradient boosting is an advanced ensemble method utilized for classification and regression tasks, and it is particularly powerful. This method builds models in a sequence such that every new model improves on the errors of previous models. Complex datasets require sophisticated modeling, and gradient boosting achieves this by minimizing residual errors to increase predictive accuracy. Gradient Boosting minimizes a loss function with gradient descent to get the best performance of its base learner, decision trees. It is an incredibly accurate predictive model [33] which is obtained from this iterative process. Gradient Boosting is a versatile method that can handle a variety of data types and distributions. It is particularly effective with big data and can detect intricate patterns. However, hyperparameter tuning is necessary to mitigate overfitting; this step is, however, highly resource-intensive, particularly with large data sets.

Adaptive Boosting

Besides being a general ensemble learning technique, AdaBoost has proven to be very effective for classification tasks. Combining multiple weak learners, such as decision trees with a small number of splits (stumps), improves the predictive performance. It is an algorithm where weights are assigned to training samples. The weights are iteratively updated according to the errors generated by the previous classifiers (who have been trained using weights); more weights are assigned to the training samples that are misclassified. Improvement in overall accuracy is achieved by combining output of weak learners and deriving a final predictive model [34, 35]. Besides, it is straightforward to implement, and it can produce very effective weak classifiers. It, however, is sensitive to noisy data and outliers and suffers from efficiency issues when base learners are incredibly complex.

Approach for Handling Numerical Data Features

There are multiple crucial steps in the suggested methodology for processing numerical data aspects. Data preprocessing is first carried out to manage missing values and outliers for quality and consistency. To find the most pertinent features and maybe improve model performance, feature selection techniques are used. The optimal performing model is identified by training and evaluating machine learning methods. Lastly, hyperparameter adjustment is done to maximize the chosen model's accuracy and endurance [36, 37].

Algorithm 1. Detection of cervical cancer using voting classifier

Procedure: Initialise classifiers, Train data X, labels y

1. Split the data into training and testing sets.
 2. Train each base classifier (RF) on the training data.
 3. for each trained base classifier to make predictions on the test set.
 4. **for** i in 1 to len (X_{test}) **do**
 5. $\text{pred}_{\text{RF}} \leftarrow \text{RF.predict}(X_{\text{test}}[i])$
 6. $\text{ensemble_vote} \leftarrow \text{Majority_vote}$
 7. Append ensemble_vote to predictions
 8. **end for**
 9. **return** predictions
-

The predictive probabilities for individual test samples are computed as the sum of predictions across all n Random Forest models RF_i undergoes Hard voting criterion [38, 39].

$$\hat{y} = \underset{c}{\operatorname{argmax}} \sum_{m=1}^n I(\hat{y}_m = c) \quad (1)$$

In this expression (1), $\hat{y}$ denotes the predicted class for a sample. $\hat{y}_m$ specifies the predicted class by classifier m for a particular sample m . $I(\hat{y}_m = c)$ this function 1 if $\hat{y}_m$ equals to c , and 0 otherwise.

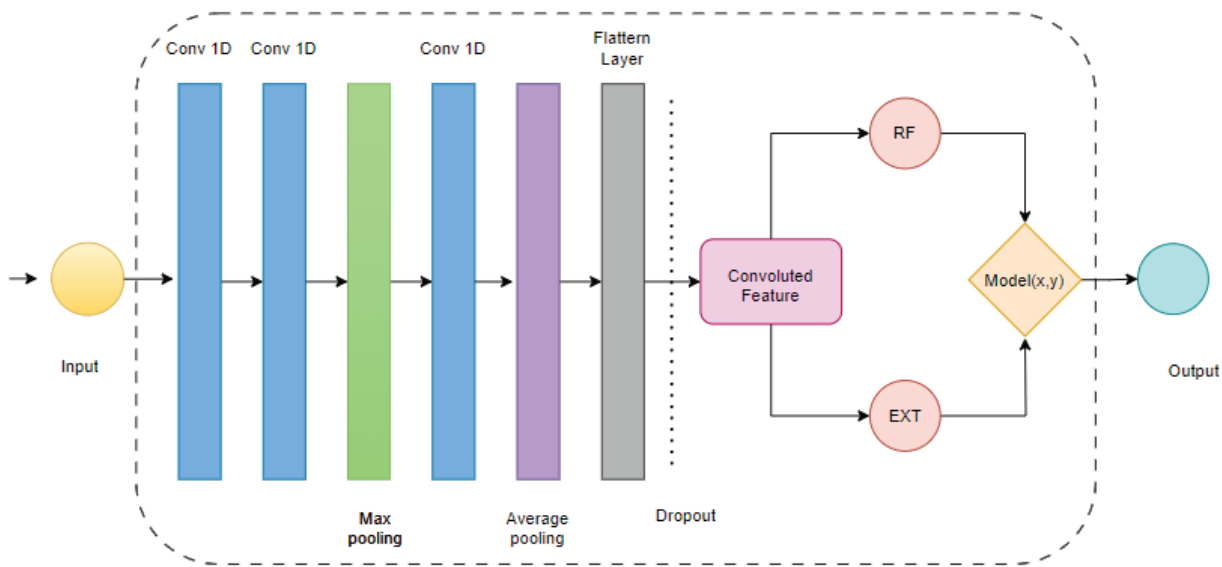


Figure 2. Cervical cancer detection model.

A Keras based model utilises a CNN (Convolutional Neural Network) architecture to handle one-dimensional input data. The dataset informs the input layer which serves as the model's starting point. The model consists of three sequential convolutional layers with 32, 64, and 128 filters respectively. Each layer contains a kernel of size three with non-linear activation through a ReLU function. To manage overfitting, increase the spatial dimensionality of the feature maps, and mitigate overfitting, each convolutional layer is followed by a 2x2 max-pooling layer. After the convolutional processes, the intermediate 3D feature maps are converted to a 1D vector. This vector is processed by a dense layer featuring 128 ReLU neurons. A crucial component for regulating overfitting is a 0.5 dropout layer prior to the final output layer. This regulates additional overfitting while preserving essential model performance features.

Evaluation Metrics

To provide a complete assessment of our ensemble technique, we apply a set of metrics that are varied in nature. While general model evaluation is usually conducted by measuring accuracy, class-specific performance is captured using precision and recall—two key metrics that determine how well the model captures true positives while mitigating the pitfalls of false positives and true negatives. The F1 score, best suited for imbalanced datasets, is the metric that balances precision and recall. In addition to measuring class distinction errors, the Receiver Operating Characteristic - Area Under the Curve (ROC AUC) metric is also used. Classification becomes more robust with higher values of the ROC AUC while greater discrimination of positive and negative instances indicates better performance. These capabilities make the metric very important for classification tests.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (3)$$

Expression 2 shows the whole model as overall correct or to have ratio of correctly predicted to total observations, just accurate. As precision, expression (3) is the correctness of optimistic prediction only.

$$\text{Recall} = \frac{TP}{TP + FN} \quad (4)$$

$$\text{F1 score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

$$\text{ROC} = \int_0^1 TPR(FPR)d(FPR) \quad (6)$$

TP = True Positives , TN = True Negatives

FP = False Positives , FN = False Negatives

Recall reflects the level of fulfilment of positive assertions made, and it is regarded as a useful metric, as it was presented in Equation (4). It evaluates the efficiency of a given set model in identifying real positive instances while ensuring that none are overlooked. F1 Score, defined as the result of harmonic function of Precision and Recall, offers some balance between two ratios discussed above. This is expressed in Equation (5) in addition to allowing for a compromise between precision, which seeks to reduce false positives, and recall, which aims to reduce false negatives.

Table 1. Summary of the research work

Ref	Models	Dataset	Performance Metric
[16]	ML and Deep Learning Algorithms + Artificial Neural Network	UCI database Cervical Cancer	93% accuracy
[17]	CNN + Fuzzy distance-based ensemble	SIPaKMed	98.58% accuracy with AUROC
[18]	CNN + transfer learning with 2 and 7-class classification	Harvel dataset	94.5% accuracy
[19]	Inceptionv3 + transfer learning	Harvel dataset	98.23% accuracy
[20]	Ensemble classification + k-means clustering	SIPaKMed	98.27% accuracy
[21]	lasso regression + SMOTE + DT for classification	Universitario de caracas dataset	98.72% accuracy
[22]	LWNB + RF + Logistic regression	Hospitals of Jordan	92.3% accuracy

The ROC curve objectively captures a classifier's performance. The TPR incorporates the FPR, and the relationship between them is defined by trade-off. As one goes up, so must the loss of the other, as illustrated by Equation (6). This is one of the essential aspects when it comes to assessing model performance and seeing how well the model breaks down classes.

RESULTS

A number of evaluations were done to detect cervical cancer using both distinct and intricate features in machine learning models. When results were analyzed, an ensemble of the top models was used in additional tests. Both feature sets were subjected to this ensemble. The models' performance was then assessed by comparing the outcomes. The analysis must check whether there is an equal weight of true positive and true negative predictions. Sensitivity (recall) and specificity were evaluated so that the model performs well in classifying cancerous and non-cancerous patients. Also, it is necessary to determine if the model suffers from overfitting. High accuracy values indicate that the model fits training instances closely and does not acquire generalized patterns. Cross-validation and unseen test sets are used to determine if the model is robust across different datasets. Incorporating such discussions would lend strength to the analysis and enhance the model's credibility in applications for the real world of clinical features (Table 1).

Table 2. Experimental setup for the model

Language	Python 3.10.12
RAM	8 GB
OS	64-bit Windows 11
CPU	Inter(R) core(TM) 11th gen i3@3.00GHz
GPU	Intel UHD Graphics
TensorFlow version	2.15.0

Experimental Setup

The system runs on a 64-bit Windows 11 OS and features an 11th gen Intel Core i3 processor at 3.00GHz, coupled with 8 GB of RAM. It utilizes Intel UHD Graphics and is equipped with TensorFlow version 2.15.0, making it suitable for fundamental machine learning tasks and development in Python 3.10.12. The following details about the experimental setup are given in Table 2.

Comparing Performance of Various Machine Learning Models on Original and Convolved Features

In general, there is a perceived improvement in model accuracy when convoluted features are used. With convoluted features, both SVM and logistic regression, which did pretty poorly with the original features as given in Table 3, significantly improved, and both achieved 90.35% accuracy. With convoluted features, models with excellent accuracy with original features, like Random Forest and Extra Trees Classifier, performed somewhat deeper, scoring about. These minor differences, however, show the resilience of these models to various feature sets. XGBoost, Extra Trees Classifier, and AdaBoost all show great accuracy and efficient feature convolution used to achieve the

Table 3. Accuracy of models with original features

Model	Accuracy with original features
LR	65.56%
DT	89.25%
RF	93.38%
KNN	86.50%
SVM	64.18%
XGB Model	90.35%
ETC	93.66%
Ada Boost	74.38%
Cat Boost	88.98%
Gradient Boost	83.19%

highest performance using convoluted data, as shown in Table 4. Significant progress was also demonstrated via gradient boosting. Convoluted features, in general, greatly improved model performance, particularly for models that did not perform as well with the original features at first. This is shown in Figure 3, which shows how feature engineering can be used to improve model accuracy.

Based on the performance comparison, the Extra Trees Classifier (ETC) performs with the best accuracy, with the original features at 93.66%, as shown in Figure 4.

However, with convoluted features, AdaBoost achieved the highest accuracy at 90.90%, as shown in Figure 5.

In this analysis, the feature engineering was paramount to improving model performance through better quality

Table 4. Accuracy of models with convoluted features

Model	Accuracy with Convoluted features
LR	90.35%
DT	92.28%
RF	92.28%
KNN	90.35%
SVM	90.35%
XGB Model	90.63%
ETC	92.01%
Ada Boost	90.90%
Cat Boost	90.50%
Gradient Boost	91.73%

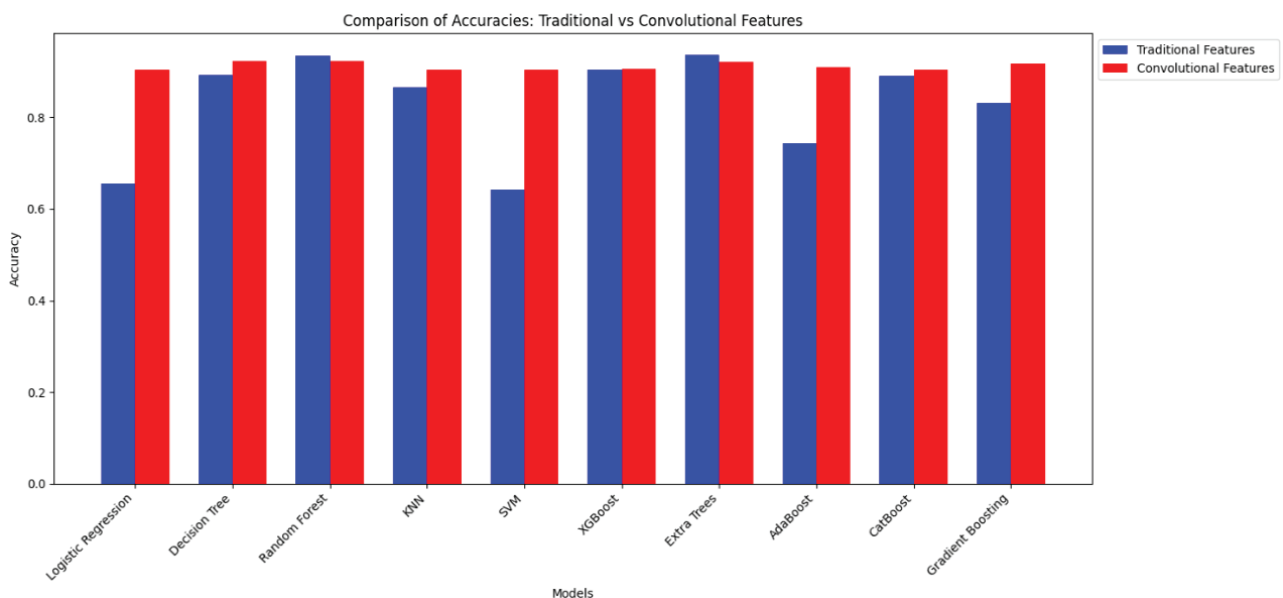


Figure 3. Comparison of original and convoluted features.

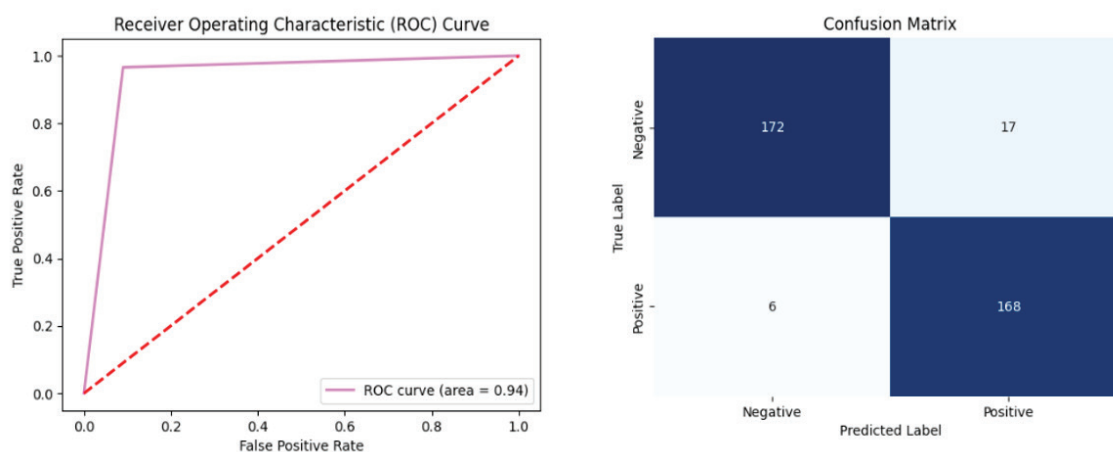


Figure 4. Extra tree classifiers with original features.

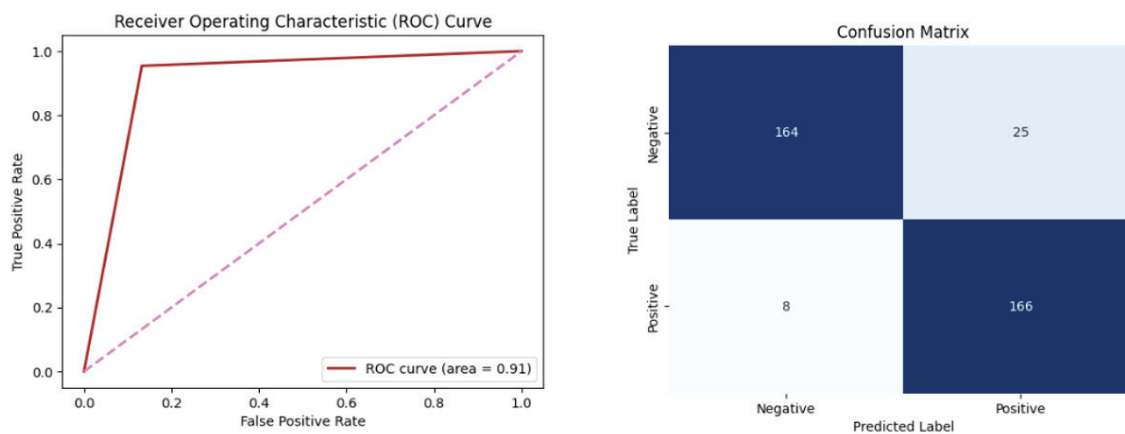


Figure 5. AdaBoost with convoluted features.

input features. The utilized data was initially derived from patient records and a dataset of raw patient record features. To identify and construct more informative patterns from the original dataset, which involved applying a number of feature engineering methods. The model constructed features were compared to the raw features in their original form, statistical transformations, such as mean, standard deviation (SD), and range, to estimate variability and distribution characteristics of continuous variables from which they were initially derived. For categorical features, for example, we applied acceptably appropriate encoding techniques, including one-hot and label encoding, to transform textual features into numbers that could be directly imported into the model. Next, some interaction features were constructed through multiple relevant original features (e.g., creating a variable to indicate age of pregnancies). Every feature was transformed, or not transformed raw features, into a consistent set of features, which reflects real-world impact of clinical settings on approximating risk for developing cervical cancer. Raw and feature-engineered features combined produced a more informative dataset, which enhanced contribution of the features for training the ensemble model; the engineered features provided a feature set that explained extreme amounts of variance in validation and test model accuracy.

Performance of Ensemble Model

Effectiveness of ensemble learning as an approach is often complicated for RF models in combination with other models, as the integration of intricate features can undermine performance results. Results are dependent on the value of additional features in improving generalization and accuracy of prediction. Enhanced feature complexity may improve the decision boundaries. Nevertheless, issues can arise as a result of overfitting or excessive processing costs in relation to how a dataset and model interact. RF combined with Decision Tree (DT) and RF merged with Extra Trees Classifier (ETC) yield a notable maximum accuracy

of 92.28%, but the latter's overfitting undermines the former model's accuracy. By combining advantages of both RF and the secondary model, these combinations suggest the possibility of a marginal improvement over individual models in Table 5. Additional combinations exhibit competitive performance as well. Gradient Boosting (GB) and RF together yield accuracy of 92.01%, which is not far behind the best combinations. This implies that RF's strong feature selection can be enhanced by GB's capacity to strengthen weak models. The RF and Logistic Regression (LR), as well as RF and Support Vector Machine (SVM), achieve accuracy of 91.46%. RF may still produce high results, but it is marginally lower than other ensembles. These combinations show how RF can improve the decision boundaries of more complicated models like SVM and simpler models like LR. RF with AdaBoost and RF with CatBoost combinations have somewhat lower accuracy. The comparison of the complete ensemble model accuracies is shown in Figure 6. They nonetheless gain from RF's feature selection capabilities, which provide the boosting techniques more stability.

The hybrid ensemble learning technique integrates convolutional neural networks (CNN) with other machine learning models such as Random Forest and Extra Trees

Table 5. Accuracy of ensemble model using convoluted features

Model	Accuracy
RF + LR	91.46%
RF + DT	92.28%
RF + SVM	91.46%
RF + XGB Model	90.90%
RF + ETC	92.28%
RF + Ada Boost	91.18%
RF + Cat Boost	91.10%
RF + Gradient Boost	92.01%

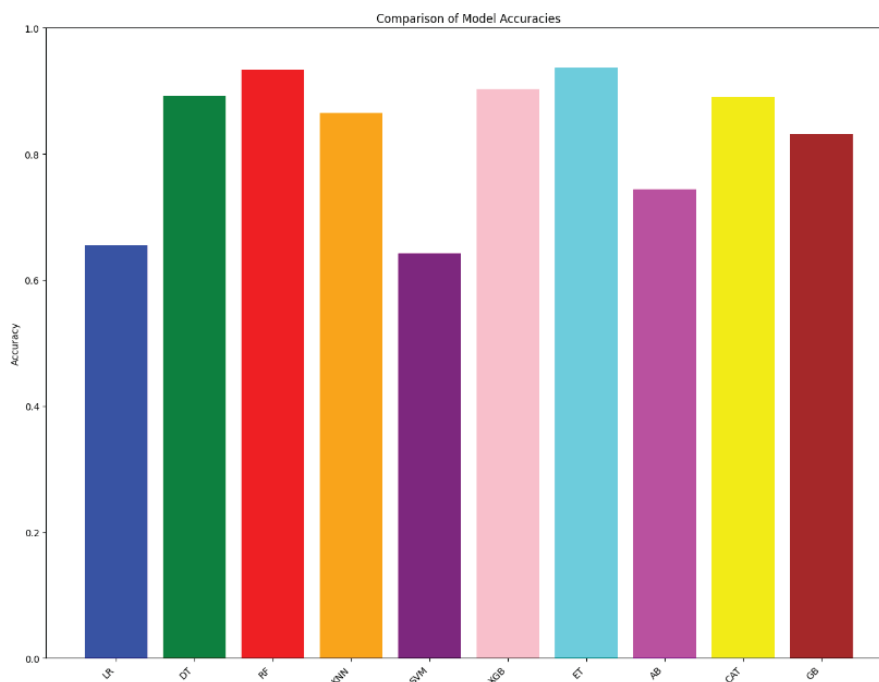


Figure 6. Comparison of ensemble model accuracies.

Classifier. This is done to take advantage of the strengths of each algorithm. Although the machine learning models may seem somewhat redundant because they rely on decision trees and an ensemble method, they differ enough to justify inclusion in the study. Random Forest utilizes multiple decision trees and bootstrap sampling. Random Forest narrows down to get the best split from among a randomly selected subset of features, while introducing randomness to reduce overfitting, while still having a high degree of accuracy. The Extra Trees (Extremely Randomized Trees) classifier has even more randomness by using the entire dataset (no bootstrapping) and selecting the split points randomly rather than selecting the best split. This makes Extra Trees computationally more efficient, and even more importantly, discovery of patterns resistant to noise is better. The intention of including both models is to maximize diversity in the base learners. Maximizing diversity in the base learners is a significant component of ensemble learning. Academics have shown that when learners are carefully chosen, diversity can yield models that generally perform better. Many conceivable attributes restrict individual models, and subtle differences like these in data sampling and feature splits are able to garner totally different contributions and insights to the final prediction.

Hyperparameter Tuning

This study combines the Random Forest and XGBoost classifiers to show how ensemble learning approaches can be used. It involves creating a Voting Classifier and specifying parameter distributions for both classifiers [40, 41]. Hyperparameter tuning is carried out using

RandomizedSearchCV, optimizing performance of the ensemble model. This search yields the optimal parameters and estimator. The accuracy of the Voting Classifier, once it has been optimized, is then determined by testing it on a different dataset [42]. Ultimately, a thorough classification report that offers insights into the model's performance in several classes is produced. Robust prediction performance is achieved by utilizing the capabilities of many classifiers in this strategy [43, 44].

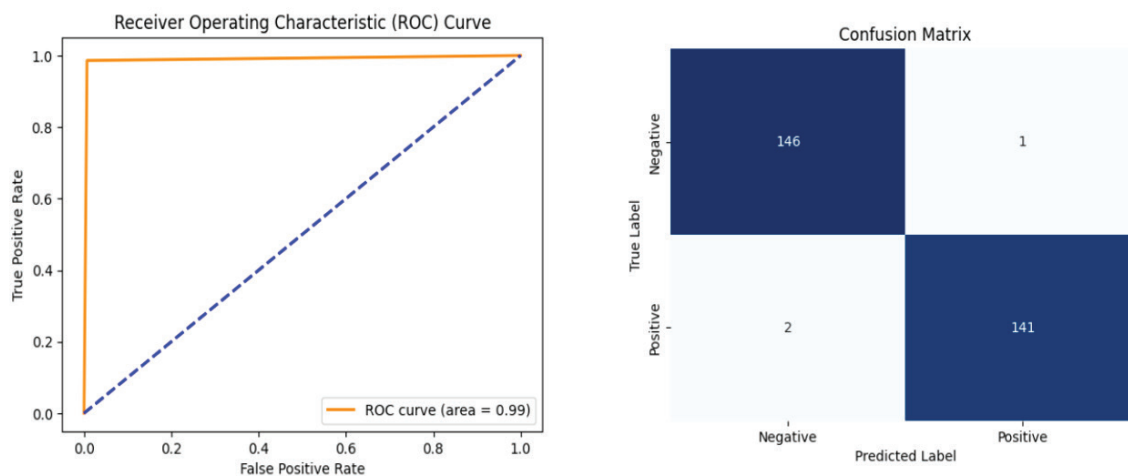
Performance of K-Fold Validations

In order to comprehensively assess differences in model performance, we conducted a statistical significance test through a paired t-test on the evaluation metrics containing accuracy, precision, recall, and F1-score from the 5-fold cross-validation. The t-test evaluates the effectiveness of proposed Hybrid Ensemble against the baseline classifiers. These performance differences were statistically significant ($p < 0.05$), which strengthens the claim that our proposed hybrid model outperformed each base model. This also reinforces the accuracy and dependability of the comparative analysis.

The functioning of the model is evaluated with the help of 5-fold cross-validation over different data sets. The entire dataset is split into five equal parts. One part is preserved for validation in each fold, while the other four parts are used in training. This ensures consistency and reliable assessment of performance. Table 6 presents that the ensemble model achieved a mean accuracy of 97.10%. Figure 7 illustrates that in fold 3, the averaged metrics of accuracy score, precision, recall, and F1 score were reported as 0.989%, 0.992%, 0.993%, and 0.991%, respectively.

Table 6. K-fold cross-validation results for the proposed ensemble model

Fold	Accuracy	Precision	Recall	F1 - Score	ROC - AUC
1	97.59%	97.50%	98.10%	97.59%	0.98
2	95.86%	96.30%	96.10%	96.30%	0.96
3	98.96%	99.25%	99.30%	99.15%	0.99
4	97.24%	97.55%	97.32%	97.30%	0.97
5	95.86%	95.52%	95.50%	95.58%	0.96
Avg.	97.10%	97.22%	97.26%	97.28%	0.97

**Figure 7.** Best outcome of ensemble model (fold 3).

DISCUSSION

Cervical cancer is an emergent health issue and world-wide concern that needs amelioration through early detection techniques to improve the outcomes of patient treatment. In fact, the present study focused on this problem by suggesting a new framework of ensembles to integrate feature extraction for classification through different machine-learning strategies, random forests, decision trees, logistic regressions, and boosting techniques. Machine Learning models are creating additional opportunities in cervical cancer detection apart from aiding in early diagnosis, enhancing screening accuracy, and supporting clinical decision-making.

1. These models analyze Pap smear images to recognize abnormal cervical cells. This reduces manual errors and promotes early detection.
2. Logistic regression and ensemble models are used for risk stratification.
3. AI-based applications and cloud-based Machine Learning allow screening in remote areas where expert pathologists cannot be found.

4. ML models trained on smartphone microscopy images provide an inexpensive and scalable solution to diagnosis.

Regardless of the advances made in cervical cancer detection by ML-based approaches, there are many hurdles to be resolved in order to ensure successful world deployment. The issue of biased predictions posed by data quality and imbalanced data is one of the real challenges in medical datasets. The ethical and privacy concerns subsist due to the very sensitive nature of medical data, which needs to be acknowledged within strict protocols. The quality issues of bias in training data lead to one or the other demographic being unfairly predicted, raising ethical concerns in health-care AI. For accessibility in underdeveloped and remote areas, optimizing architectures for the edge remains relevant. A model produced under study could show strong performance with accuracy as high as 98.96% and was further validated through K-fold cross-validation steps, ensuring those benefits of robustness and reliability. These findings further establish the base of ensemble methods that hold promise towards such complementary strengths between deep learning and machine learning in clinical diagnosis. However, a few limitations remain, for example, the need for interpretability and the fact that it should be

applicable in resource-constrained environments, which now mark out their further usage in the actual scenario outside settings.

CONCLUSION

Cervical cancer maintains its status as one of the most prevalent and lethal cancers in women, continually contributing to cancer mortality rates globally. Decreasing death rates and enhancing patient management outcomes depends on accurate and timely diagnosis. In response to this challenge, we propose a new computational framework for increasing the accuracy of cervical cancer diagnosis by merging features from various fields of data and numerical analytics. The significant results of our research are:

1. The ensemble learning framework synthesized a convolutional neural network (CNN) with multiple classifiers and achieved predictive accuracy of 98.96%.
2. Feature engineering and feature selection were also integral in improving model performance.
3. Robustness testing through K-fold cross-validation ensures reliability.
4. Realization of advanced accuracy rates for cervical cancer detection.
5. A firm comparative analysis indicates that our proposed framework was successful in improving cervical cancer diagnosis capability.

By computing multiple sources of data, the proposed framework achieves better prediction accuracy and enhances understanding of the underlying factors related to the diagnostic ability. The results from this research work thus support the constructive deployment of machine learning in revolutionizing medical diagnostics towards making timely and accurate diagnoses of cervical cancer.

The findings illustrate the necessity of using deep learning approaches under machine learning to address complex healthcare challenges. Therefore, for future enhancement of the detection of cervical cancer with an ML-based system, there are several integral routes forward. Nevertheless, the study does have some limitations. The model was tested only on one dataset; it was not tested with any external datasets to evaluate the generalizability of its predictions, and the real-time deployment of the model in a clinical situation and its evaluation are not the intent of this study. These are essential next steps to evaluate its functionality in real-world health care workflows. The advancement of this line of work has clear next steps planned, including the validation of the model on external and multi-center datasets, conducting real-world clinical trials, and deploying the system for real-time use in clinical situations. The implementation of explainable AI techniques, SHAP and LIME, improves the model's interpretability, which could enhance clinicians' trust in the decision-support system.

Additionally, several goals could lead to revitalizing implementation, including the use of lighter-weight architectures for low-resource settings and the implementation

of system interfaces with electronic health records (EHRs) through standardized application programming interfaces (APIs), which could enhance real-time decision support and accessibility. Transfer learning is an additional possibility that could improve model generalizability and training time. All these aspects would benefit the advancement of machine learning-based cervical cancer detection systems designed to be clinically relevant, valid and accurate, interpretable, and accessible.

NOMENCLATURE

CNN	Convolutional Neural Network
HPV	Human Papillomavirus
WHO	World Health Organization
RF	Random Forest
DT	Decision Tree
LR	Logistic Regression
KNN	K-Nearest Neighbors
XGB	Extreme Gradient Boosting
SVM	Support Vector Machine
ETC	Extra Tree Classifier
CatBoost	Categorical Boosting
GB	Gradient Boosting
AdaBoost	Adaptive Boosting
AUROC	Area Under Receiver Operating Characteristic Curve
EHR	Electronic Health Record
CAD	Computer-Aided Diagnosis
SMOTE	Synthetic Minority Over-sampling Technique

Greek Symbols

α (alpha)	Learning rate
β (beta)	Coefficients in regression models
Σ (sigma)	Summation
μ (mu)	Mean of a distribution
σ (sigma)	Standard deviation
λ (lambda)	Regularization parameter

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available at the following URL: <https://www.kaggle.com/datasets/ranzeet013/cervical-cancer-dataset>

CONFLICT OF INTEREST

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

STATEMENT ON THE USE OF ARTIFICIAL INTELLIGENCE

Artificial intelligence was not used in the preparation of the article.

REFERENCES

- [1] Meza Ramirez CA, Greenop M, Almoshawah YA, Martin Hirsch PL, Rehman IU. Advancing cervical cancer diagnosis and screening with spectroscopy and machine learning. *Expert Rev Mol Diagn* 2023;23:375–390. [\[CrossRef\]](#)
- [2] Chauhan NK, Singh K. Performance assessment of machine learning classifiers using selective feature approaches for cervical cancer detection. *Wirel Pers Commun* 2022;124:2335–2366. [\[CrossRef\]](#)
- [3] Singh SK, Goyal A. Performance analysis of machine learning algorithms for cervical cancer detection. In: *Research Anthology on Medical Informatics in Breast and Cervical Cancer*. IGI Global; 2023. p. 347–370. [\[CrossRef\]](#)
- [4] Chen H, Liu J, Wen QM, Zuo ZQ, Liu JS, Feng J, et al. CytoBrain: cervical cancer screening system based on deep learning technology. *J Comput Sci Technol* 2021;36:347–360. [\[CrossRef\]](#)
- [5] Park YR, Kim YJ, Ju W, Nam K, Kim S, Kim KG. Comparison of machine and deep learning for the classification of cervical cancer based on cervicography images. *Sci Rep* 2021;11:16143. [\[CrossRef\]](#)
- [6] Pacal I, Kılıcarslan S. Deep learning-based approaches for robust classification of cervical cancer. *Neural Comput Appl* 2023;35:18813–18828. [\[CrossRef\]](#)
- [7] Akter L, Islam MM, Al-Rakhami MS, Haque MR. Prediction of cervical cancer from behavior risk using machine learning techniques. *SN Comput Sci* 2021;2:177. [\[CrossRef\]](#)
- [8] Sellamuthu Palanisamy V, Athiappan RK, Nagalingam T. Pap smear based cervical cancer detection using residual neural networks deep learning architecture. *Concurr Comput Pract Exp* 2022;34. [\[CrossRef\]](#)
- [9] Kalbhor M, Shinde S, Popescu DE, Hemanth DJ. Hybridization of deep learning pre-trained models with machine learning classifiers and fuzzy min-max neural network for cervical cancer diagnosis. *Diagnostics (Basel)* 2023;13:1363. [\[CrossRef\]](#)
- [10] Fekri-Ershad S, Alsaffar MF. Developing a tuned three-layer perceptron fed with trained deep convolutional neural networks for cervical cancer diagnosis. *Diagnostics (Basel)* 2023;13:686. [\[CrossRef\]](#)
- [11] Weegar R, Sundström K. Using machine learning for predicting cervical cancer from Swedish electronic health records by mining hierarchical representations. *PLoS One* 2020;15. [\[CrossRef\]](#)
- [12] Kalbhor MM, Shinde SV. Cervical cancer diagnosis using convolution neural network: feature learning and transfer learning approaches. *Soft Comput* 2023;1–11. [\[CrossRef\]](#)
- [13] Kalbhor M, Shinde S, Joshi H, Wajire P. Pap smear-based cervical cancer detection using hybrid deep learning and performance evaluation. *Comput Methods Programs Biomed Imaging Vis* 2023;11:1615–1624. [\[CrossRef\]](#)
- [14] Kumawat G, Vishwakarma SK, Chakrabarti P, Chittora P, Chakrabarti T, Lin JCW. Prognosis of cervical cancer disease by applying machine learning techniques. *J Circuits Syst Comput* 2023;32:2350019. [\[CrossRef\]](#)
- [15] Kalbhor M, Shinde S, Popescu DE, Hemanth DJ. Hybridization of deep learning pre-trained models with machine learning classifiers and fuzzy min-max neural network for cervical cancer diagnosis. *Diagnostics (Basel)* 2023;13:1363. [\[CrossRef\]](#)
- [16] Kaushik M, Joshi RC, Kushwah AS, Gupta MK, Banerjee M, Burget R, et al. Cytokine gene variants and socio-demographic characteristics as predictors of cervical cancer: a machine learning approach. *Comput Biol Med* 2021;134:104559. [\[CrossRef\]](#)
- [17] Mehmood M, Rizwan M, Gregus ML, Abbas S. Machine learning assisted cervical cancer detection. *Front Public Health* 2021;9:788376. [\[CrossRef\]](#)
- [18] Pramanik R, Biswas M, Sen S, de Souza Júnior LA, Papa JP, Sarkar R. A fuzzy distance-based ensemble of deep models for cervical cancer detection. *Comput Methods Programs Biomed* 2022;219:106776. [\[CrossRef\]](#)
- [19] Lin H, Hu Y, Chen S, Yao J, Zhang L. Fine-grained classification of cervical cells using morphological and appearance based convolutional neural networks. *IEEE Access* 2019;7:71541–71549. [\[CrossRef\]](#)
- [20] Dong N, Zhao L, Wu CH, Chang JF. Inception v3 based cervical cell classification combined with artificially extracted features. *Appl Soft Comput* 2020;93:106311. [\[CrossRef\]](#)
- [21] Win KP, Kitjaidure Y, Hamamoto K, Myo Aung T. Computer-assisted screening for cervical cancer using digital image processing of pap smear images. *Appl Sci* 2020;10:1800. [\[CrossRef\]](#)
- [22] Tanimu JJ, Hamada M, Hassan M, Kakudi HA, Abiodun JO. Machine learning method for classification of cervical cancer. *Electronics* 2022;11:463. [\[CrossRef\]](#)
- [23] Al-Batah MS, Alzyoud M, Alazaidah R, Toubat M, Alzoubi H, Olaiyat A. Early prediction of cervical cancer using machine learning techniques. *Jordan J Comput Inf Technol* 2022;8. [\[CrossRef\]](#)

- [24] Ramzan Z, Hassan MA, Asif HM, Farooq A. A machine learning-based self-risk assessment technique for cervical cancer. *Curr Bioinform* 2021;16:315–332. [\[CrossRef\]](#)
- [25] Sun L, Yang L, Liu X, Tang L, Zeng Q, Gao Y, et al. Optimization of cervical cancer screening: a stacking-integrated machine learning algorithm based on demographic, behavioral, and clinical factors. *Front Oncol* 2022;12:821453. [\[CrossRef\]](#)
- [26] Jia AD, Li BZ, Zhang CC. Detection of cervical cancer cells based on strong features CNN-SVM network. *Neurocomputing* 2020;411:112–127. [\[CrossRef\]](#)
- [27] Ji M, Zhong J, Xue R, Su W, Kong Y, Fei Y, et al. Early detection of cervical cancer by fluorescence lifetime imaging microscopy combined with unsupervised machine learning. *Int J Mol Sci* 2022;23:11476. [\[CrossRef\]](#)
- [28] Sellamuthu Palanisamy V, Athiappan RK, Nagalingam T. Pap smear based cervical cancer detection using residual neural networks deep learning architecture. *Concurr Comput Pract Exp* 2022;34. [\[CrossRef\]](#)
- [29] Waly MI, Sikkandar MY, Aboamer MA, Kadry S, Thinnukool O. Optimal deep convolution neural network for cervical cancer diagnosis model. *Comput Mater Contin* 2022;70. [\[CrossRef\]](#)
- [30] Vargas-Cardona HD, Rodriguez-Lopez M, Arrivillaga M, Vergara-Sanchez C, Garcia-Cifuentes JP, Bermúdez PC, et al. Artificial intelligence for cervical cancer screening: scoping review, 2009–2022. *Int J Gynaecol Obstet* 2024;165:566–578. [\[CrossRef\]](#)
- [31] Suguna C, Balamurugan SP. An extensive review on machine learning and deep learning based cervical cancer diagnosis and classification models. *J Comput Theor Nanosci* 2020;17:5438–5446. [\[CrossRef\]](#)
- [32] Singh J, Tan KT, Sahu D, Upreti K, Kshirsagar PR, Elangovan M. An investigation on machine learning models in classification and identifications cervical cancer using MRI scan. In: 2024 International Conference on E-mobility, Power Control and Smart Systems (ICEMPS). 2024. p. 1–5. [\[CrossRef\]](#)
- [33] Nour MK, Issaoui I, Edris A, Mahmud A, Assiri M, Ibrahim S. Computer aided cervical cancer diagnosis using gazelle optimization algorithm with deep learning model. *IEEE Access* 2024. [\[CrossRef\]](#)
- [34] Munshi RM. Novel ensemble learning approach with SVM-imputed ADASYN features for enhanced cervical cancer prediction. *PLoS One* 2024;19. [\[CrossRef\]](#)
- [35] Mustafa WA, Ismail S, Mokhtar FS, Alquran H, Al-Issa Y. Cervical cancer detection techniques: a chronological review. *Diagnostics (Basel)* 2023;13:1763. [\[CrossRef\]](#)
- [36] Chadaga K, Prabhu S, Sampathila N, Chadaga R, KS S, Sengupta S. Predicting cervical cancer biopsy results using demographic and epidemiological parameters: a custom stacked ensemble machine learning approach. *Cogent Eng* 2022;9:2143040. [\[CrossRef\]](#)
- [37] Xue P, Ng MTA, Qiao Y. The challenges of colposcopy for cervical cancer screening in LMICs and solutions by artificial intelligence. *BMC Med* 2020;18:169. [\[CrossRef\]](#)
- [38] Dweekat OY, Lam SS. Cervical cancer diagnosis using an integrated system of principal component analysis, genetic algorithm, and multilayer perceptron. *Healthcare (Basel)* 2022;10:2002. [\[CrossRef\]](#)
- [39] Uddin KMM, Al Mamun A, Chakrabarti A, Mostafiz R, Dey SK. An ensemble machine learning-based approach to predict cervical cancer using hybrid feature selection. *Neurosci Inform* 2024;4:100169. [\[CrossRef\]](#)
- [40] Devi S, Gaikwad SR, Harikrishnan R. Prediction and detection of cervical malignancy using machine learning models. *Asian Pac J Cancer Prev* 2023;24:1419. [\[CrossRef\]](#)
- [41] Glučina M, Lorencin A, Anđelić N, Lorencin I. Cervical cancer diagnostics using machine learning algorithms and class balancing techniques. *Appl Sci* 2023;13:1061. [\[CrossRef\]](#)
- [42] Mathivanan SK, Francis D, Srinivasan S, Khataavkar V, P K, Shah MA. Enhancing cervical cancer detection and robust classification through a fusion of deep learning models. *Sci Rep* 2024;14:10812. [\[CrossRef\]](#)
- [43] Xiao T, Wang C, Yang M, Yang J, Xu X, Shen L, et al. Use of virus genotypes in machine learning diagnostic prediction models for cervical cancer in women with high-risk human papillomavirus infection. *JAMA Netw Open* 2023;6. [\[CrossRef\]](#)
- [44] Chikaraddi AK, Kanakaraddi S, Kalwad P, Kadole A, Budni S. Cervical cancer diagnostics using machine learning algorithms. In: 2024 5th International Conference for Emerging Technology (INCET). 2024. p. 1–6. [\[CrossRef\]](#)
- [45] Arora M, Dhawan S, Singh K. Cervical cancer diagnosis and prediction: an application of machine learning techniques. In: *Computational Intelligence in Analytics and Information Systems*. Apple Academic Press; 2023. p. 213–226. [\[CrossRef\]](#)