



Augmenting Cervical Cancer Analysis with Deep Learning Classification and Topography Selection Using Artificial Bee Colony Optimization

K. Ramu¹ · Arun Ananthanarayanan² · P. Joel Josephson³ · N. R. Rejin Paul⁴ · Praveen Tumuluru⁵ · Ch. Divya⁶ · Sanjay Kumar Suman⁷

Received: 7 March 2024 / Accepted: 5 June 2024

© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd. 2024

Abstract

According to the research and study, cervical cancer has risen to develop the fourth most communal malignancy to strike women. Five different forms of gynaecologic cancer affect the feminine generative organism. The cervix, the lower portion of the body that joins the vagina and the uterus, is where cervical cancer develops in a woman. Cancers, in general, are abnormal alterations in cell development that take place within the human body. Additionally, aberrant cell alterations in the uterine lining or at the womb's opening have been linked to cervical cancer. Additionally, the Artificial Bee Colony (ABC) approach's enhancement of the topography selection process is taken into consideration. This work suggests a novel approach for better identifying the risk factors for cervical cancer in females by combining an evolutionary technique for topography selection with a deep learning model. The lack of specificity regarding the timeframe or demographic affected might limit the study's applicability and generalizability. To create an improvised topography selection, a deep learning method known as LSTM is paired with an evolutionary computation method known as ABC. The model's accuracy is found to be 98.68% when compared to previously used models like SVM-PCA and SVM-BC. Comparing the implemented model to other models, it provided the highest level of accuracy.

Keywords ABC · LSTM · SVM-PCA · ML

Introduction

According to figures based on the literature, cervical cancer is one of the deadliest tumors, ranking fourth among prevalent malignancies in women and seventh overall. Each

year, more than 500,000 of these instances are reported worldwide [1]. Additionally, it can be deduced from many statistical studies that this cancer is more prevalent in less developed parts of the world, more so than other types of cancer that are more widespread [2]. Finding the best subset

✉ K. Ramu
k.ramu147@gmail.com
Arun Ananthanarayanan
paatu.84@gmail.com
P. Joel Josephson
joeljosephsonp@gmail.com
N. R. Rejin Paul
nrrejinpaul@gmail.com
Praveen Tumuluru
Praveenluru@gmail.com
Ch. Divya
divyachavali17@gmail.com
Sanjay Kumar Suman
prof.dr.sanjaykumarsuman@gmail.com

¹ IBM KYNDRYL LLC, Chennai, TN, India
² Department of ECE, Dayanand Sagar University, Bangalore, Karnataka, India
³ Department of ECE, Malla Reddy Engineering College, Secunderabad, Telangana, India
⁴ Department of CSE, R.M.K College of Engineering and Technology, Chennai, TN, India
⁵ Department of CSE, Koneru Lakshmaiah Education Foundation, Vaddeswaram, AP, India
⁶ Department of AI&DS, St. Martin's, Engineering College, Secunderabad, Telangana, India
⁷ Department of ECE and R&D, St. Martin's Engineering College, Secunderabad, Telangana, India

of topographies from the vast pool of potentially obtainable topographies is the process of topography selection. Filter, wrapper, and embedding methods are a few of the topography selection techniques. Filter-based topography selection techniques are extremely scalable for huge datasets, computationally quick, and simple to understand [3]. In their implementation, wrapper- and embedded-based topography selection techniques make use of certain machine learning algorithms. The accuracy of the model is improved by employing swarm-based topography selection approaches in order to single out the most important risk factors and topographies from among the numerous risk variables [4]. When there is a large space for topography selection, or when the solution space is large, bio-inspired algorithms are better equipped to search for optimal and nearly optimal solutions [5].

Various health-related companies use certain computing approaches, such as classification, clustering, etc. through machine and deep learning algorithms to improve medical diagnosis. Due to a shortage of medical equipment, it is impossible for people in developing nations and those who are economically underprivileged to use a medical diagnostic system [7]. This is the driving force behind computer-aided screening methods, which aid in the early diagnosis of the illness and lengthen the life expectancy of women. Values must be subjected to multivariate analysis for cervical cancer screening [8]. The available dataset's solution space is largely constrained, making it unlikely to produce accurate results. Methods for identifying approaches to generate samples for applying screening procedures and generating greater convergence to solutions present challenges [9].

Contribution of the Work

- Cervical cancer ranks as the fourth most common malignancy affecting women, underscoring the need for improved detection methods.

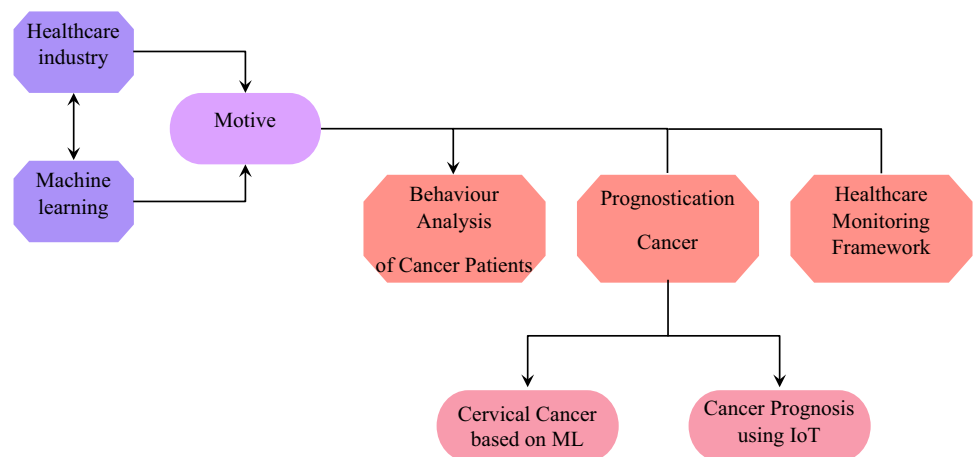
- The study proposes a novel approach, combining evolutionary topography selection with deep learning (LSTM) and Artificial Bee Colony (ABC) optimization.
- Results indicate a significant accuracy boost to 98.68%, outperforming conventional models like SVM-PCA and SVM-BC, thus advancing cervical cancer risk factor identification.

The ABC algorithm is a bio-inspired topography selection method that lowers the classifier miss rate. Any numerical issue can be optimized using the ABC algorithm, a population-based stochastic optimization technique [11]. The ABC algorithm's usage of the honey bees' food-foraging activities as a topography selection strategy is described in the proposed study. Solutions are used in this manner to refer to the honey bees' food source. To effectively utilize food supplies, the bees can be dispersed over a range of distances [12]. This ABC technique requires three key elements in order to have a minimal model for choosing the foragers: food supplies, hired foragers, and jobless foragers [13]. Recruitment of a rich food source, which produces positive feedback, and abandonment of a food source, which produces negative feedback, are the two primary actions that are associated with the self-organizing of bees [15] (Fig. 1).

In this study, the dataset's contributing attributes to the causes of cervical cancer are chosen using the ABC method. These types of bees—the spectator bee, the employed bee, and the scout bee are considered by this algorithm [16]. Every colony contains an equal number of workers and observers. Every bee in the workforce travels to the food sources and returns to the hive to alert the spectator bee by dancing in the dance area. By watching the dancing steps of a working bee, the spectator bee chooses the food source. An employed bee becomes a scout bee and begins searching for a new food source when it abandons its current one [18].

It is used to analyse the risk factors in the classification of cervical cancer. LSTM is an expanded version of RNN.

Fig. 1 General machine learning based cervical cancer analysis



Short-term memory is not a problem with this strategy. An input gate, an output gate, and a forget gate are among its three gates [20]. Each layer of the two-hidden-layer neural network has 100 nodes. These cell units get the activation signals from a variety of sources. The designed multipliers regulate how the cell is activated [21]. The rest of the networks' memory cell contents cannot be changed continuously thanks to an LSTM gate. The topographies chosen from the ABC module are supplied to the LSTM as input in this work [23].

The ABC algorithm is used in this study to choose characteristics optimally, increasing the classifier's accuracy while doing so. The outcome demonstrates that the suggested system selects optimal topographies using the ABC methodology and achieves higher accuracy. It is assumed that the classifier's accuracy is negligible in the absence of a topography selection strategy. The results of the experimental study unequivocally demonstrate the value of the topography selection strategy in categorization [24–26].

The objective is to construct an intelligence-driven model by integrating Support Vector Machine, Genetic Algorithm, Backpropagation Network techniques, and a filter-based topography selection strategy. This aims to enhance accuracy and efficiency in complex data analysis, particularly in pattern recognition and classification tasks.

The subsequent section provides an outline for the remaining content of the paper. In part 2 of the document, you will find a brief description of the related work. This section provides an overview of the existing research and studies that are relevant to the topic at hand. Moving on to "[The Objective of the Work](#)", you will find a detailed explanation of the methodology employed in this study, as well as the theoretical foundations that underpin the methods used. This section aims to provide a clear understanding of the approach taken and the principles guiding the research process. In "[The Proposed Work](#)", we will discuss the simulation results and analysis. In the concluding section of this research paper, titled "key findings," we aim to provide a concise summary of the most significant outcomes.

Previously Done Work

Diagnostic and prognostic strategies for cervical cancer, including those based on Pap smear imaging, clinical analysis, gene expression analysis, and other relevant screening methods, have been the focus of numerous researchers. Screening approaches for cervical cancer detection and prognosis are discussed in full here. In this section, we also discuss the study's findings and drawbacks in light of the approaches used to investigate the topic (Table 1).

The research gap lies in the absence of comprehensive exploration into the combined application of deep learning

classification and topography selection techniques in augmenting cervical cancer analysis. While the study demonstrates promising results in terms of accuracy, there remains a need for further investigation into the scalability, robustness, and real-world applicability of these methods across diverse patient populations and clinical settings.

The Objective of the Work

- Develop a comprehensive model integrating Support Vector Machine, Genetic Algorithm, and Backpropagation Network methodologies.
- Incorporate a filter-based topography selection strategy to enhance model accuracy and efficiency.
- Aim for an intelligence-based approach to effectively analyze and interpret complex data, particularly in the context of pattern recognition and classification tasks.

The Proposed Work

The suggested study employs a deep learning technology called long short-term memory with ABC (LSTM-ABC) to aid in the identification of cervical cancer.

The cervical cancer screening model was trained and tested on a dataset comprising colposcopy images sourced from Intel and Smartphone ODT's public cervical screening dataset. It includes various cervix types, with experts classifying raw images based on visible transition zones, encompassing three types of cervical pre-cancerous transformation zones.

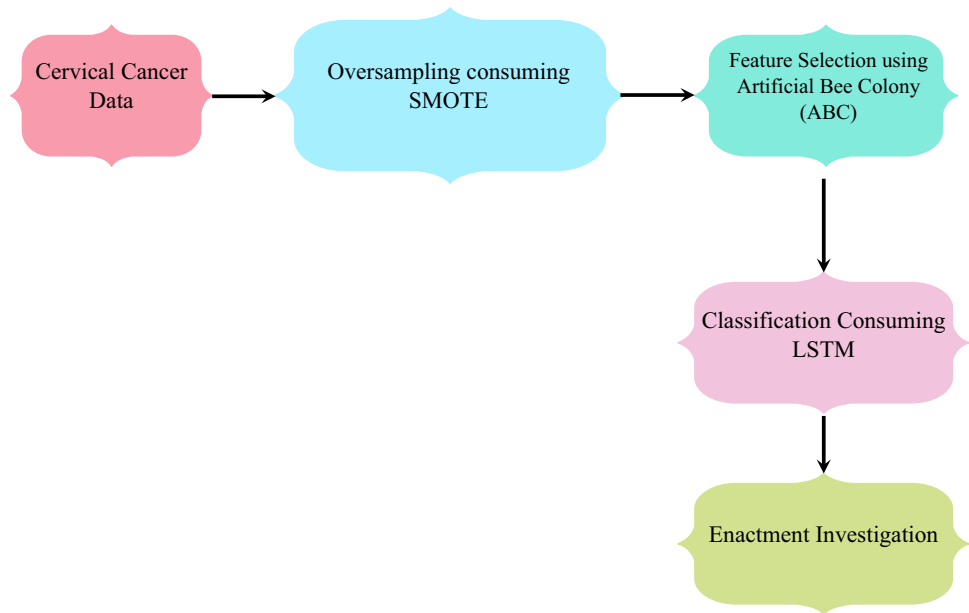
Figure 2 depicts the process diagram for this study. In order to lower the classifier miss rate, the suggested work uses a bio-inspired technique dubbed the ABC algorithm for topography selection. The honey bees' foraging behaviour in the ABC algorithm is used as a topography selection approach in the suggested work. In this strategy, answers are referred to as potential honey bee food sources. Bees can be dispersed over large areas, allowing them to make more efficient use of available food. This ABC technique comprises three important components—food sources, employed foragers, and unemployed foragers—in order to create a minimum model to pick the foragers. Recruitment to the rich food source results in positive feedback, while abandonment of a food source results in negative feedback; both of these behaviours are related to the self-organization of bees.

The proposed study uses an ABC algorithm to identify and prioritise data points that contribute to understanding the root causes of cervical cancer. This algorithm takes into account three different types of bees—the observer bee, the worker bee, and the scout bee—in order to function properly. There are always as many worker bees as there are observers

Table 1 The existing work done review and tabular analysis

S. no.	Citations	Research work	Findings
1	[6]	They used logistic regression for the analysis of Pap smear images on the dataset that was described with 133 topographies for 1728 patients. The CART cell level approach was used in the model which improved the sensitivity of the model and pAUC curve	CART, L1 regularized logistic regression, and elastic net regression were utilized as classifiers. To evaluate model accuracy, training and validation datasets were combined and subjected to fivefold cross-validation. Due to fewer cervical cancer cases, estimating sensitivity and positive predictive value was challenging. Model specificity was 96%, with a negative predictive value of 97%
2	[10]	They implemented three learning approaches Support Vector Machine, C5.0, and extreme learning machine to predict recurrent cervical cancer by identifying the important risks for the disease	Data from a Tumor repository included records of 168 patients with 12 attributes. 118 records trained the model, 50 tested it. Identifying risk variables for recurrent cervical cancer, the C5.0 algorithm was pivotal. Independent factors included pathologic stage, invasive tumors, pathologic T, Cell Type, and RT target summary. Given the limited patient count, medical diagnosis interventions were vital
3	[14]	They carried out research on Pap smear images by implementing the SVM algorithm and Random Forest method on the dataset with 38 attributes on 75 samples and grouped into 7 classes	Pap smear results are cancer-positive or negative. Cell topographies in Pap smear images further confirmed findings. RF approach achieved 80.10% classification accuracy
4	[17]	They proposed research work on Pap smear images with 20 Topographies selected using the Relief method and Random Forest as the classifier. Out of the selected topographies, 13 topographies showed contribution to the findings	Naive Bayes compared experimental outcomes, while RF achieved superior classification (94.44% accuracy). Cervical cell categorization, excluding cytoplasm, used various dimensional parameters. Automating Pap smear predictions benefits from including cytoplasm
5	[19]	Their proposal combines genetic algorithms with rough set theory and the ID3 algorithm. Genetic algorithm tunes network weights and structure simultaneously	CNCI institution contributed a dataset of 221 patients, each with 21 Boolean attributes. The proposed model outperformed standard multilayer perception in handling complex inputs. Despite 81.5% accuracy, it excelled with low-dimensional data featuring few characteristics
6	[22]	Authors proposed a bottom-up approach using SVM with 149 cell images from Pap smear images to automatically detect cancer cells	This study demonstrated an increase in cell-level classification accuracy of 98%, however it only applied to the small image dataset

Fig. 2 The proposed algorithm flow diagram



in any one colony. Every worker bee finds its way to the flowering plants and back to the hive, where it dances for the benefit of the other bees in attendance. A bystander bee chooses the food source by watching the dance of a worker bee. A worker bee becomes a scout bee when it leaves its food source and begins searching for a new one.

Proposed method algorithm:

Input: Topographies count in the cervical cancer data.

Output: Optimum topographies.

1. Reset the topographies count, $i = 1 \dots SP$
2. Reset topography locus
3. Calculate suitability a of specific topographies
4. Established cycle to 1
5. Recurrence
6. FOR every engaged bee
7. Harvest new designated topographies
8. Compute the appropriateness function
9. Relate the avaricious assortment procedure
10. Compute the probability P_a for the topography
11. FOR apiece onlooker bee
12. Select a designated topography X_a dependent on P_a
13. Yield novel designated topographies
14. Compute the appropriateness function
15. Relate the avaricious assortment procedure
16. If there is an uninhibited resolution for emissary bees then
17. Substitute it with novel designated topographies
18. Remember the greatest designated topographies attained so far
19. $cycle = cycle + 1$
20. Until
21. $cycle = M.C.N.$

In this study, we employ LSTM, a type of RNN with additional functionality, to examine the parameters involved in cervical cancer categorization and their associated risks.

The fusion of deep learning and evolutionary computation enhances cervical cancer risk factor identification by leveraging the strengths of each approach. Deep learning excels in extracting intricate patterns from complex data, while evolutionary computation optimizes feature selection, improving model interpretability and performance. This synergy allows for more accurate and robust identification of risk factors, leading to enhanced diagnostic capabilities and better-informed medical decision-making in cervical cancer management.

This strategy is highly resistant to the issue of short memory. An input gate, an output gate, and a forget gate are its three gates. It is essentially a two-layer neural network with 100 nodes in each hidden layer. These cell units get activation signals from a wide variety of sources. The designed multipliers regulate the cell's activity. An LSTM gate blocks constant changes to the information stored in a network's memory cells. The input to LSTM in this investigation comes from the topographies chosen using the ABC module. The number of categories is displayed as a result. If the memory output passes through the gate, then cervical cancer has occurred.

Result Analysis and Discussion

The suggested and current methodologies undergo experimental investigation in the Java simulation environment. As a direct consequence of this, the dataset has certain minor inaccuracies, which are rectified by the modifications to

the techniques that have been explained. In this study, the accuracy, sensitivity, and specificity of the proposed long short-term memory with ABC (LSTM-ABC) approach are evaluated and contrasted with those of the existing method.

Accuracy

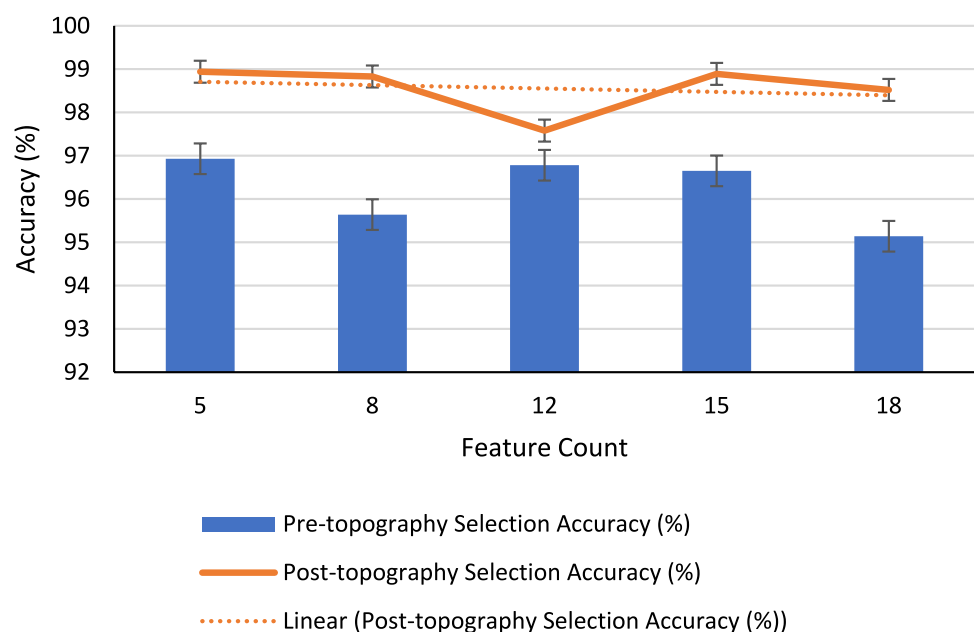
To better choose characteristics for the classifier, this study proposes using the ABC algorithm for optimal topography selection. The experimental results demonstrate that the suggested system selects optimal topographies with a higher degree of accuracy using the ABC methodology. It is obvious that the classifier's accuracy will suffer greatly if a topography selection strategy is not used. The topography selection method is crucial for classification, as demonstrated by the experimental investigation. Accuracy values with and without topography selection are listed in Table 2.

When evaluating a model, accuracy is defined as the ratio of the observed classification parameters to the full set of parameters used for classification.

Table 2 Accuracy Assessment for pre and post topography selection

S. no.	Topography count	Pre-topography selection accuracy (%)	Post-topography selection accuracy (%)
1	5	96.93	98.94
2	8	95.64	98.83
3	12	96.78	97.58
4	15	96.65	98.89
5	18	95.14	98.52

Fig. 3 Accuracy assessment for pre and post topography selection



$$\text{Accuracy} = \frac{TN + TP}{TP + TN + FP + FN} \times 100 \quad (1)$$

In Fig. 3, we see a comparison of the model's classification accuracy before and after applying topography selection approaches.

Sensitivity

The formula is used to determine a test's sensitivity, which is its ability to correctly identify persons with the condition (true positive rate).

$$\text{Sensitivity} = \frac{TP}{TP + FN} \times 100 \quad (2)$$

Specificity

Specificity measures how well a test can distinguish healthy individuals from those who are merely suspicious (Table 3).

Table 3 Evaluation assessment of the proposed work with existing work done

S. no.	Assessment parameter	SVM-PCA	SVM-ABC	LSTM-ABC
1	Accuracy (%)	96.73	96.89	98.68
2	Sensitivity (%)	92.58	94.28	95.67
3	Specificity (%)	92.69	93.49	94.28

$$\text{Specificity} = \frac{TN}{TN + FP} \times 100 \tag{3}$$

Figure 4 provides a comparative analysis between the proposed LSTM with ABC and the current SVM-PCA method, focusing on specificity. To mitigate the class imbalance inherent in cervical cancer datasets, the study employs Synthetic Minority Over-sampling Technique (SMOTE). Furthermore, it harnesses the capabilities of long short-term memory (LSTM) networks for cervical cancer classification. LSTM's adeptness in capturing long-term dependencies and retaining previously acquired knowledge enhances its ability to predict the proportion of healthy patients reliably.

The suggested LSTM with ABC algorithm exhibits superior performance across sensitivity, accuracy, and specificity metrics when compared to traditional SVM-PCA and SVM-ABC approaches. This highlights its potential as an advanced tool for cervical cancer diagnosis and risk assessment. Notably, the study incorporates feature selection techniques, particularly in classes such as Cytology and Biopsy, to enhance model efficiency and interpretability. The One-Versus-All (OVA) technique is applied to both SVM and LSTM models, enabling multi-class classification.

Moreover, the study evaluates the impact of adaptive gradient boosting on model performance by comparing results obtained with and without its application. This analysis sheds light on the efficacy of different optimization strategies in enhancing the predictive capabilities of the models. Overall, the integration of LSTM with ABC, coupled with feature selection and advanced optimization techniques, presents a novel and promising approach to cervical cancer classification, offering improved diagnostic accuracy and contributing to the advancement of medical decision-making processes (Table 4).

Table 4 Performance measure for cytology class

Parameters	SVM-ABC		LSTM-ABC	
	Without gradient boosting	With gradient boosting	Without gradient boosting	With gradient boosting
Accuracy (%)	88.57	95.12	89.35	95.62
Recall (%)	91.54	98.67	89.88	98.14
Specificity (%)	80.24	89.53	86.43	94.28
Precision (%)	78.72	94.27	91.27	96.37

From Figs. 5 and 6 the proposed algorithm is compared with SVM-ABC algorithm with and without gradient boosting for the class Cytology. We can analyse that the proposed algorithm is outperformed in both the cases with the existing algorithm. All the parameters are improved in comparison to the state of the art algorithm. Similarly in the case of biopsy class the proposed algorithm outperformance is observed. Using a probability calculation, this method verifies the selected features that are relevant to cervical cancer prediction. They will then be instructed to begin regular screening procedures to reduce their risk of developing cervical cancer.

The qualities are employed to populate a correlation matrix, showcasing variables' correlation coefficients' magnitudes. This matrix serves as a valuable tool for summarizing the dataset and its inherent patterns effectively. Utilized in regression models with multiple independent variables, the correlation coefficient plays a crucial role. Figure 7 provides a visual representation of the correlation among the variables within the cervical cancer dataset,

Fig. 4 Evaluation assessment of the proposed work with existing work done

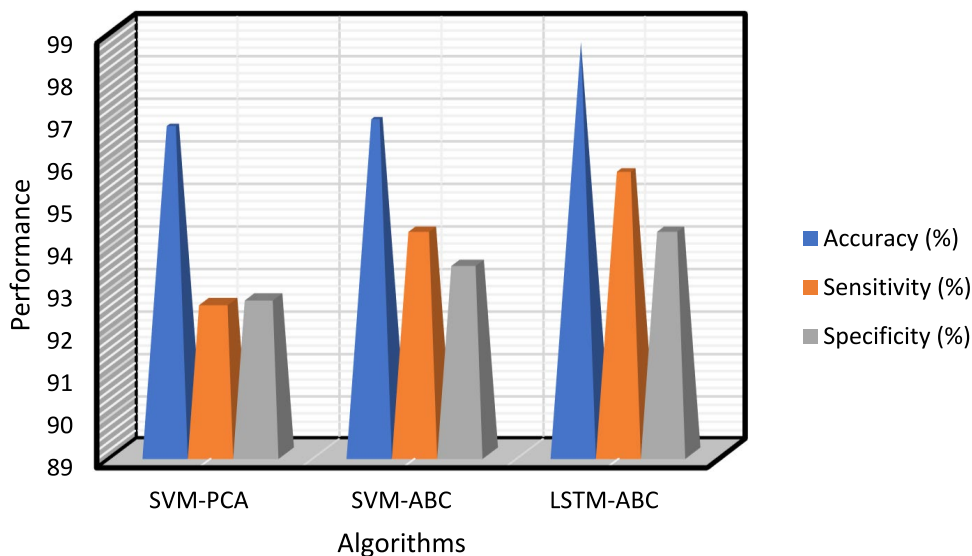


Fig. 5 Comparison of the proposed method in case of without gradient boosting

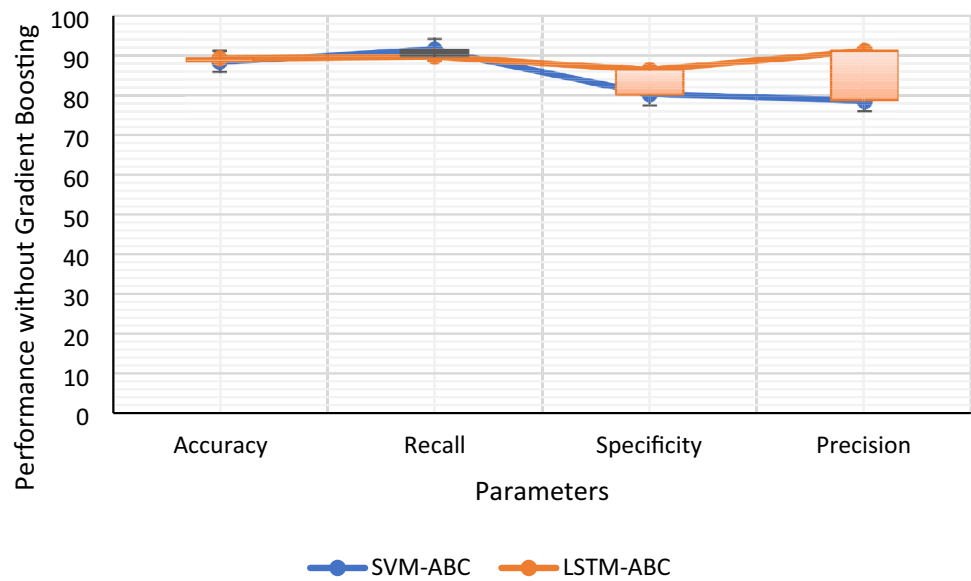
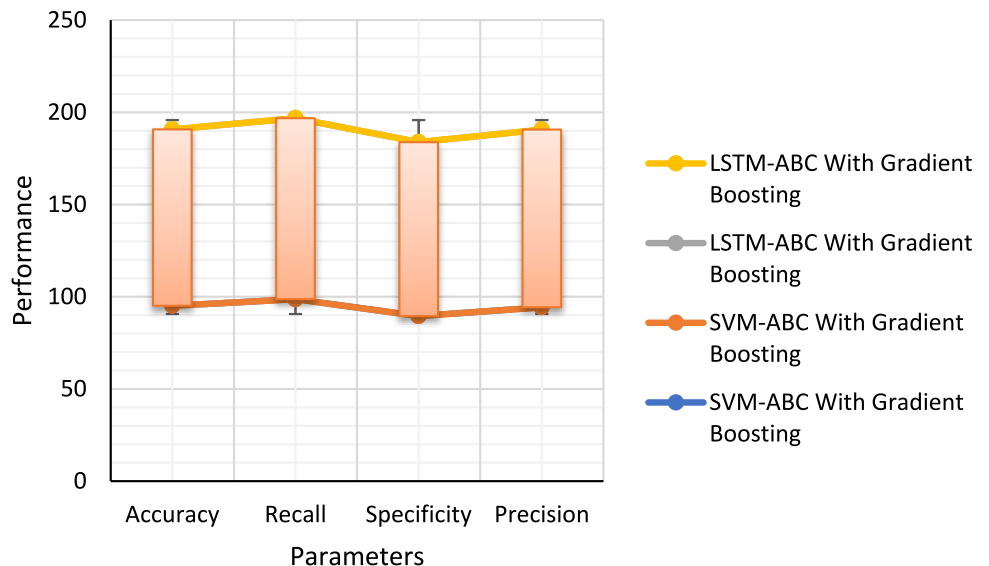


Fig. 6 Comparison of the proposed method in case of with gradient boosting



offering insights into their interrelationships and aiding in the interpretation of data-driven findings.

Figure 7 shows that the strongest association exists between the characteristic "smokes" and "smokes (years)," which makes sense given that smoking is an independent risk factor for cervical cancer. Number of STD diagnoses is also highly correlated. Because sexually transmitted diseases (STDs) are so prevalent and have been linked to an increased risk of cervical cancer.

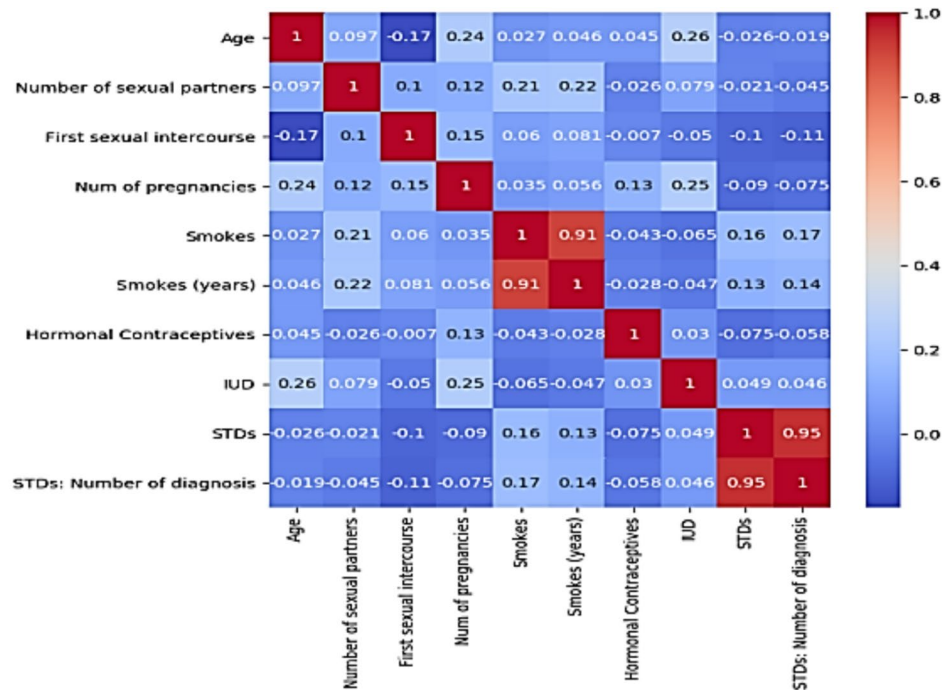
The study revealed cervical cancer's alarming rise, ranking as the fourth most prevalent malignancy among women. Additionally, it underscored the impact of aberrant cell alterations in the uterine lining and cervix on cervical cancer development, emphasizing the urgent need

for effective screening and diagnostic strategies to combat this growing public health concern.

Conclusion and Future Scope

Machine learning approaches can be found in this period of widespread digitalization and high speed computation, allowing for the creation of effective techniques and the acceleration of diagnostic operations. In order to accurately diagnose cervical cancer, the suggested system employs an algorithm dubbed the long short-term memory and ABC (LSTM-ABC) method. In order to improve the classifier's performance, we employ the SMOTE method

Fig. 7 Correlation matrixes of independent variables from cervical cancer dataset



e

to fix issues caused by an unbalanced data set and the ABC algorithm to choose topographies optimally. The topographies are used in a classification process utilising the long short-term memory (LSTM) method. The experimental results demonstrate that the suggested system produces higher accuracy, sensitivity, and specificity compared to the state-of-the-art models.

To further enhance classification accuracy, many different classification methods, including ensemble classifiers and convolution neural network (CNN), will be applied in the future. In the future, this study's focus will expand to include predicting cancer stage and identifying risk factors associated with the disease. Women can be protected from sexually transmitted diseases (STDs) with vaccination against human papillomavirus (HPV) before the age of 15, as is recommended by the World Health Organisation.

Author contributions Ramu K: consumption and design of study, acquisition of the data, Arun Ananthanarayanan: analysis and interpretation of the data, P. Joel Josephson: drafting, formalization and editing, N. R. Rejin Paul: review and investigation, Praveen Tumuluru: conceptualization, Ch. Divya and Sanjay Kumar Suman: investigation and analysis.

Funding Not applicable.

Data availability Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Declarations

Conflict of interest First author and second author declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with animals performed by any of the authors.

References

1. Khamparia A, Gupta D, Rodrigues JJ, de Albuquerque VHC. DCAVN: cervical cancer prediction and classification using deep convolutional and variational autoencoder network. *Multim Tools Appl.* 2021;80:30399–415. <https://doi.org/10.1007/s11042-020-09607-w>.
2. Abbas S, Jalil Z, Javed AR, Batool I, Khan MZ, Noorwali A, et al. BCD-WERT: a novel approach for breast cancer detection using whale optimization-based efficient features and extremely randomized tree algorithm. *PeerJ Comput Sci.* 2021;7: e390. <https://doi.org/10.7717/peerj-cs.390>.
3. Ayoub A, Mahboob K, Javed AR, Rizwan M, Gadekallu TR, Abidi MH, et al. Classification and categorization of COVID-19 outbreak in Pakistan. *Comput Mater Continua.* 2021;69:1253–69. <https://doi.org/10.32604/cmc.2021.015655>.
4. Iwendi C, Bashir AK, Peshkar A, Sujatha R, Chatterjee JM, Pasupleti S, et al. COVID-19 patient health prediction using boosted random forest algorithm. *Front Public Health.* 2020;8:357. <https://doi.org/10.3389/fpubh.2020.00357>.
5. Rehman SU, Javed AR, Khan MU, Nazar Awan M, Farukh A, Hussien A. Personalised-comfort: a personalized thermal comfort

- model to predict thermal sensation votes for smart building residents. *Enterprise Inform Syst.* 2020. <https://doi.org/10.1080/17517575.2020.1852316>.
6. Yamal JM, Guillaud M, Atkinson EN, Follen M, MacAulay C, Cantor SB, Cox DD. Prediction using hierarchical data: applications for automated detection of cervical cancer. *Stat Anal Data Min.* 2015;8(2):65–74.
 7. Javed AR, Faheem R, Asim M, Baker T, Beg MO. A smartphone sensors-based personalized human activity recognition system for sustainable smart cities. *Sustain Cities Soc.* 2021;71: 102970. <https://doi.org/10.1016/j.scs.2021.102970>.
 8. Roy V. Breast cancer classification with multi-fusion technique and correlation analysis. *Fus Pract Appl.* 2023;9(2):48–61.
 9. Iwendi C, Mahboob K, Khalid Z, Javed AR, Rizwan M, Ghosh U. Classification of COVID-19 individuals using adaptive neuro-fuzzy inference system. *Multim Syst.* 2021. <https://doi.org/10.1007/s00530-021-00774-w>.
 10. Tseng C-J, Lu C-J, Chang C-C, Chen G-D. Application of machine learning to predict the recurrence-proneness for cervical cancer. *Neural Comput Appl.* 2014;24(6):1311–6.
 11. Javed AR, Sarwar MU, Beg MO, Asim M, Baker T, Tawfik H. A collaborative healthcare framework for shared healthcare plan with ambient intelligence. *Hum Centric Comput Inform Sci.* 2020;10:1–21. <https://doi.org/10.1186/s13673-020-00245-7>.
 12. Javed AR, Fahad LG, Farhan AA, Abbas S, Srivastava G, Parizi RM, et al. Automated cognitive health assessment in smart homes using machine learning. *Sustain Cities Soc.* 2021;65: 102572. <https://doi.org/10.1016/j.scs.2020.102572>.
 13. Roy V, Shukla S, Shukla PK, Rawat P. Gaussian elimination-based novel canonical correlation analysis method for EEG motion artifact removal. *J Healthc Eng.* 2017;2017;Article ID 9674712. <https://doi.org/10.1155/2017/9674712>.
 14. Kurniawati, YE, Permanasari AE, Fauziati S. Comparative study on data mining classification methods for cervical cancer prediction using pap smear results. In: 2016 1st international conference on biomedical engineering (IBIOMED). 2016. pp. 1–5.
 15. Bruni LAG, Serrano B, Mena M, Gómez D, Muñoz J, Bosch FX, de San José S. Human papillomavirus and related diseases report in world; summary report; ICO/IARC Information Centre on HPV and Cancer (HPV Information Centre), Institut Català d'Oncologia: Barcelona, Spain, 2023.
 16. Yuan C, Yao Y, Cheng B, Cheng Y, Li Y, Li Y, Liu X, Cheng X, Xie X, Wu J, et al. The application of deep learning based diagnostic system to cervical squamous intraepithelial lesions recognition in colposcopy images. *Sci Rep.* 2020;10:11639.
 17. Sun G, Li S, Cao Y, Lang F. Cervical cancer diagnosis based on random forest. *Int J Perform Eng.* 13(4).
 18. Shukla S, Roy V, Prakash A. Wavelet based empirical approach to mitigate the effect of motion artifacts from EEG signal. In: 2020 IEEE 9th international conference on communication systems and network technologies (CSNT), Gwalior, India, 2020. pp. 323–6. <https://doi.org/10.1109/CSNT48778.2020.9115761>.
 19. Mitra P, Mitra S, Pal SK. Staging of cervical cancer with soft computing. *IEEE Trans Biomed Eng.* 2000;47(7):934–40.
 20. Zing DSW, Han NKC, Lee ZYK, Tan KL. AI for medical imaging goes deep. *Nat Med.* 2018;24:539–40.
 21. Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell.* 2018;172:1122–31.
 22. Zhang J, Liu Y. Cervical cancer detection using SVM based topography screening. In: International conference on medical image computing and computer-assisted intervention. 2004. pp. 873–80.
 23. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer.* 2018;18:500–10.
 24. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, Mahajan V, Rao P, Warier P. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *Lancet.* 2018;392:2388–96.
 25. Ali A, et al. Detection of gene ontology clusters using biclustering algorithms. *SN Comput Sci.* 2023;4(3):217.
 26. Ali A, et al. DPEBic: detecting essential proteins in gene expressions using encoding and biclustering algorithm. *J Ambient Intell Humaniz Comput.* 2021:1–8.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.