

Prediction of Cervical Cancer

Subjects: **Computer Science**, **Artificial Intelligence**

Contributor: Hanen Karamti , Raed Alharthi , Amira Al Anizi , Reemah M. Alhebshi , Ala' Abdulmajid Eshmawi , Shtwai Alsubai , Muhammad Umer

Cervical cancer ranks among the top causes of death among females in developing countries. The most important procedures that should be followed to guarantee the minimizing of cervical cancer's aftereffects are early identification and treatment under the finest medical guidance.

cervical cancer

cancer prediction

1. Introduction

Cervical cancer is a form of cancer that arises in the cells of the cervix, the lower region of the uterus that connects to the vagina. Typically, cervical cancer is initiated by an infection resulting from human papillomavirus (HPV), a sexually transmitted infection. HPV is a prevalent virus capable of inducing abnormal alterations in cervical cells, which, if left untreated, can potentially progress into cancer ^[1].

Cervical cancer is ranked as the third leading cause of death for women, following breast cancer ^[2] and lung cancer. Unfortunately, it is commonly believed that cervical cancer remains incurable in advanced stages. However, significant progress has been made recently to improve the detection rate of the disease by using imaging techniques. Based on statistics provided by the World Health Organization (WHO), cervical cancer ranks as the fourth most prevalent cancer worldwide. In 2018 alone, around 570,000 new cases were documented, constituting 7.5% of all female cancer-related fatalities ^[3]. Out of the reported 311,000 annual deaths attributed to cervical cancer, approximately 85% occur in countries with lower- and middle-income economies. Timely detection of cervical cancer plays a crucial role in preserving lives. Women with HIV face a six-fold higher risk of developing cervical cancer compared to those without HIV, and it is estimated that 5% of all cervical cancer cases are associated with HIV. Several factors contribute to the effectiveness of screening, including access to equipment, consistent screening tests, adequate supervision, and the identification and treatment of detected lesions ^[4].

Cervical cancer can be categorized into two main types: squamous cell cancer, which accounts for 70–80% of cases, and adenocarcinoma, which originates from glandular cells responsible for producing cervical canal mucus. Although squamous cell carcinoma is more common, the occurrence of adenocarcinoma has been on the rise in recent years, now accounting for 10 to 15% of uterine cancers. Detecting adenocarcinoma through screening presents greater challenges as it develops in the cervical canal rather than the cervix itself. However, the treatment approaches for both types of cancer are similar ^{[5][6]}. The primary cause of cervical cancer is human papillomavirus (HPV), particularly high-risk types. Several risk factors can increase the likelihood of developing cervical cancer in

women infected with HPV. These factors include smoking, early sexual activity, multiple sexual partners, genital herpes infection, a weakened immune system, lower socioeconomic status, poor genital hygiene, and a higher number of childbirths [7][8]. Symptoms of cervical cancer can vary depending on the tumor's size and the stage of the disease. However, the challenge lies in the pre-cancerous stage, as it often lacks noticeable symptoms and is typically detected incidentally during routine annual check-ups. In advanced stages, approximately 90% of cases present clear symptoms, with irregular vaginal bleeding being the primary symptom associated with cervical cancer.

The process of cervical cancer screening often involves a gynecological examination, which can be painful [9][10] and uncomfortable [11] for patients. The discomfort experienced during the examination can result in delays or avoidance, which hinders early diagnosis. Additionally, inadequate public health policies in developing nations contribute to low rates of cervical cancer screening. As a result, the mortality rate in these countries is 18 times higher [12], with approximately nine out of 10 deaths related to cervical cancer transpiring in low-income countries [13]. Considering that early-stage cervical cancer has relatively high survival rates, reaching up to 90% over a 5-year period [14], it is imperative to improve cervical cancer screening rates. However, screening rates differ between countries, with higher rates observed in developed countries [15] and alarmingly low rates in developing nations.

A range of preventive measures are implemented to combat cervical cancer; however, relying solely on screening tests is insufficient. The timely detection of cervical cancer in its early stages is vital for preventing deaths caused by invasive cervical cancer. Presently, computer vision, machine learning (ML) [16], artificial intelligence (AI) [17], and deep learning (DL) techniques [18] are extensively utilized in disease detection [19]. ML models, in particular, have garnered considerable interest due to their ability to swiftly identify specific diseases [20]. By employing various preprocessing techniques such as data cleaning, dimensionality reduction, and feature selection on the disease dataset, ML algorithms can be applied to achieve precise and accurate results. These analyzed outcomes can aid medical professionals in swiftly diagnosing diseases and providing optimal treatments to patients.

2. Prediction of Cervical Cancer

Machine learning (ML) [21] is an extraordinary tool that finds application in numerous domains, extending to the identification and diagnosis of diseases in diverse animal and plant species. In recent years, numerous ML models have been developed and utilized to enhance research efforts and expedite progress in specific areas of interest.

Machine and deep learning models are used for different types of medical diagnoses like breast cancer [22], Lung cancer [23], endoscopy [24], and many others. CT images are the most accurate dataset for image-based medical diagnosis [25][26][27]. Some other research works make use of deep learning models for cross-domain work like image-captioning [28][29], drowsy driver detection [30], and neural stem differentiation [31]. CNN applications are also extended to mirror detection with visual chirality cue [32]. In a research study conducted by Kalbhor et al. [33], the discrete cosine transform (DCT) and discrete wavelet transform (DWT) were employed to extract features. To effectively reduce the dimensionality of these features, the fractional coefficient approach was utilized. The reduced features were then utilized as input for seven machine learning classifiers to differentiate between various

subgroups of cervical cancer. The study achieved an accuracy of 81.11%. Devi and Thirumurugan [34] conducted another study where they utilized the C-means clustering algorithm to segment cervical cells. Texture features, including the Gray-Level Co-occurrence Matrix (GLCM) and geometrical descriptors, were extracted from these cells. To reduce the dimensionality of the extracted features, principal component analysis (PCA) was employed. Subsequently, the K-nearest neighbors (KNN) algorithm was utilized to classify the cervical cells, resulting in an accuracy of 94.86%.

In their study, Alquran et al. [35] focused on the classification of cervical cancer using the Harvel dataset. They combined deep learning (DL) with a cascading support vector machine (SVM) classifier to achieve accurate results. By integrating these techniques, they successfully classified cervical cancer into seven distinct categories with an impressive accuracy of up to 92%. In their research, Kalbhor et al. [36] introduced an innovative hybrid technique that combined deep learning architectures, machine learning classifiers, and a fuzzy min–max neural network. Their approach focused on the feature extraction and classification of Pap smear images. The researchers utilized pre-trained deep learning models, including AlexNet, ResNet-18, ResNet-50, and GoogleNet. The experimental evaluation was conducted using benchmark datasets, namely, Herlev and Sipakmed. Notably, the highest classification accuracy of 95.33% was achieved by fine-tuning the ResNet-50 architecture, followed by AlexNet, on the Sipakmed dataset.

Tanimu et al. [37] conducted a study focusing on the identification of risk factors associated with cervical cancer using the decision tree (DT) classification algorithm. They utilized recursive feature elimination (RFE) and least absolute shrinkage and selection operator (LASSO) feature selection techniques to identify the most important attributes for predicting cervical cancer. The dataset used in the study had missing values and exhibited a high level of imbalance. To address these challenges, the researchers employed a combination of under and oversampling techniques called SMOTETomek. The results demonstrated that the combination of DT, RFE, and SMOTETomek achieved an impressive accuracy score of 98.72%. Quinlan et al. [38] conducted a comparative analysis to assess different machine learning models for cervical cancer classification. The dataset used in their study exhibited class imbalance, requiring a solution to address this issue. To mitigate the class imbalance problem, the researchers employed the resampling technique called SMOTE-Tomek in combination with a tuned Random Forest algorithm. The results demonstrated that the Random Forest classifier with SMOTE-Tomek achieved a remarkable accuracy score of 99.69%.

Gowri and Saranya [39] proposed a machine learning framework for accurate cervical cancer prediction. Their approach involved the utilization of DBSCAN and SMOTE-Tomek to identify outliers in the dataset. Two prediction scenarios were conducted: DBSCAN + SMOTE-Tomek + RF and DBSCAN + SMOTE + RF. The research findings demonstrated that the DBSCAN + SMOTE + RF approach achieved an impressive accuracy rate of 99%. Abdoh et al. [40] proposed a cervical cancer classification system that utilized the Random Forest (RF) classification technique along with the synthetic minority oversampling technique (SMOTE) and two feature reduction methods: recursive feature elimination and principal component analysis (PCA). The experiment utilized a dataset containing 30 features. The study investigated the impact of varying the number of features and found that using SMOTE with RF and all 30 features resulted in an impressive accuracy of 97.6%.

Ijaz et al. [41] proposed a data-driven system for the early prediction of cervical cancer. Their approach incorporated outlier detection and the SMOTE oversampling method. The classification task was performed using the random forest algorithm in combination with Density-Based Spatial Clustering of Applications with Noise (DBSCAN). Results of their study showed that the DBSCAN + SMOTE-Tomek + RF approach achieved an impressive accuracy score of 97.72% when applied to a dataset with 10 features. Jahan et al. [42] presented an automated system for the detection of invasive cervical cancer. Their research focused on comparing the performance of eight different classification algorithms in identifying the disease. The study involved selecting various top feature sets from the dataset and employed a combination of feature selection techniques, including Chi-square, SelectBest, and Random Forest, to handle missing values. Notably, the MLP algorithm achieved an impressive accuracy of 98.10% when applied to the top 30 features. Mudawi and Alazeb [43] introduced a comprehensive research system consisting of four phases for the prediction of cervical cancer. Their study involved utilizing various machine learning models such as logistic regression (LR), random forests (RF), decision trees (DT), k-nearest neighbors (KNN), Gradient Boosting Classifier (GBC) Adaptive Boosting, support vector machines (SVM), and XGBoost (XGB). The findings revealed that SVM achieved an impressive accuracy score of 99% in the prediction task.

Through an extensive literature survey, it has been observed that various existing approaches have demonstrated favorable performance in predicting cervical cancer across different datasets. Nevertheless, researchers have utilized various optimization techniques to improve performance metrics such as accuracy, precision, and recall.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 2018, 68, 394–424.
2. Umer, M.; Naveed, M.; Alrowais, F.; Ishaq, A.; Hejaili, A.A.; Alsubai, S.; Eshmawi, A.; Mohamed, A.; Ashraf, I. Breast Cancer Detection Using Convolved Features and Ensemble Machine Learning Algorithm. *Cancers* 2022, 14, 6015.
3. Dong, N.; Zhao, L.; Wu, C.H.; Chang, J.F. Inception v3 based cervical cell classification combined with artificially extracted features. *Appl. Soft Comput.* 2020, 93, 106311.
4. Zhang, T.; Luo, Y.m.; Li, P.; Liu, P.Z.; Du, Y.Z.; Sun, P.; Dong, B.; Xue, H. Cervical precancerous lesions classification using pre-trained densely connected convolutional networks with colposcopy images. *Biomed. Signal Process. Control* 2020, 55, 101566.
5. Seoud, M.; Tjalma, W.A.; Ronsse, V. Cervical adenocarcinoma: Moving towards better prevention. *Vaccine* 2011, 29, 9148–9158.
6. Gien, L.T.; Beauchemin, M.C.; Thomas, G. Adenocarcinoma: A unique cervical cancer. *Gynecol. Oncol.* 2010, 116, 140–146.

7. Villa, L.L. Human papillomaviruses and cervical cancer. *Adv. Cancer Res.* 1997, 71, 321–341.
8. Schiffman, M.; Castle, P.E.; Jeronimo, J.; Rodriguez, A.C.; Wacholder, S. Human papillomavirus and cervical cancer. *Lancet* 2007, 370, 890–907.
9. Baser, E.; Togrul, C.; Ozgu, E.; Esercan, A.; Caglar, M.; Gungor, T. Effect of pre-procedural state-trait anxiety on pain perception and discomfort in women undergoing colposcopy for cervical cytological abnormalities. *Asian Pac. J. Cancer Prev.* 2013, 14, 4053–4056.
10. Wong, G.; Li, R.; Wong, T.; Fan, S. The effect of topical lignocaine gel in pain relief for colposcopic assessment and biopsy: Is it useful? *BJOG Int. J. Obstet. Gynaecol.* 2008, 115, 1057–1060.
11. Michail, G.; Androutsopoulos, G.; Panas, P.; Valasoulis, G.; Papadimitriou, I.; Poulas, K.; Adonakis, G. Effects of Orally Administered Preliminary Analgesic Therapy in Diagnostic Colposcopy Patients: A Prospective Questionnaire Study. *Open Med. J.* 2021, 8, 1–7.
12. Zhang, S.; Xu, H.; Zhang, L.; Qiao, Y. Cervical cancer: Epidemiology, risk factors and screening. *Chin. J. Cancer Res.* 2020, 32, 720.
13. Bedell, S.L.; Goldstein, L.S.; Goldstein, A.R.; Goldstein, A.T. Cervical cancer screening: Past, present, and future. *Sex. Med. Rev.* 2020, 8, 28–37.
14. Guimarães, Y.M.; Godoy, L.R.; Longatto-Filho, A.; Reis, R.d. Management of early-stage cervical cancer: A literature review. *Cancers* 2022, 14, 575.
15. Maver, P.; Poljak, M. Primary HPV-based cervical cancer screening in Europe: Implementation status, challenges, and future plans. *Clin. Microbiol. Infect.* 2020, 26, 579–583.
16. Lu, J.; Song, E.; Ghoneim, A.; Alrashoud, M. Machine learning for assisting cervical cancer diagnosis: An ensemble approach. *Future Gener. Comput. Syst.* 2020, 106, 199–205.
17. Bi, W.L.; Hosny, A.; Schabath, M.B.; Giger, M.L.; Birkbak, N.J.; Mehrtash, A.; Allison, T.; Arnaout, O.; Abbosh, C.; Dunn, I.F.; et al. Artificial intelligence in cancer imaging: Clinical challenges and applications. *CA Cancer J. Clin.* 2019, 69, 127–157.
18. Devnath, L.; Luo, S.; Summons, P.; Wang, D. Performance comparison of deep learning models for black lung detection on chest X-ray radiographs. In *Proceedings of the 3rd International Conference on Software Engineering and Information Management*, Sydney, NSW, Australia, 12–15 January 2020; pp. 150–154.
19. Arzhaeva, Y.; Wang, D.; Devnath, L.; Amirgholipour, S.; McBean, R.; Hillhouse, J.; Luo, S.; Meredith, D.; Newbiggin, K.; Yates, D. Development of Automated Diagnostic Tools for Pneumoconiosis Detection from Chest X-ray Radiographs. The Final Report Prepared for Coal Services Health and Safety Trust 2019. Available online: <https://www.coalservices.com.au/wp-content/uploads/2017/11/Project-No.-20647-Final-Report.pdf> (accessed on 15 July 2023).

20. Das, K.; Cockerell, C.J.; Patil, A.; Pietkiewicz, P.; Giulini, M.; Grabbe, S.; Goldust, M. Machine learning and its application in skin cancer. *Int. J. Environ. Res. Public Health* 2021, 18, 13409.
21. Song, X.; Tong, W.; Lei, C.; Huang, J.; Fan, X.; Zhai, G.; Zhou, H. A clinical decision model based on machine learning for ptosis. *Bmc Ophthalmol.* 2021, 21, 169.
22. Feng, H.; Yang, B.; Wang, J.; Liu, M.; Yin, L.; Zheng, W.; Yin, Z.; Liu, C. Identifying malignant breast ultrasound images using ViT-patch. *Appl. Sci.* 2023, 13, 3489.
23. Zhuang, Y.; Chen, S.; Jiang, N.; Hu, H. An Effective WSSNet-Based Similarity Retrieval Method of Large Lung CT Image Databases. *KSII Trans. Internet Inf. Syst.* 2022, 16, 2359–2376.
24. Liu, Y.; Tian, J.; Hu, R.; Yang, B.; Liu, S.; Yin, L.; Zheng, W. Improved feature point pair purification algorithm based on SIFT during endoscope image stitching. *Front. Neurorobotics* 2022, 16, 840594.
25. Lu, S.; Yang, B.; Xiao, Y.; Liu, S.; Liu, M.; Yin, L.; Zheng, W. Iterative reconstruction of low-dose CT based on differential sparse. *Biomed. Signal Process. Control* 2023, 79, 104204.
26. Zhuang, Y.; Jiang, N.; Xu, Y. Progressive distributed and parallel similarity retrieval of large CT image sequences in mobile telemedicine networks. *Wirel. Commun. Mob. Comput.* 2022, 2022, 6458350.
27. Liu, H.; Xu, Y.; Chen, F. Sketch2Photo: Synthesizing photo-realistic images from sketches via global contexts. *Eng. Appl. Artif. Intell.* 2023, 117, 105608.
28. Yang, S.; Li, Q.; Li, W.; Li, X.; Liu, A. Dual-level representation enhancement on characteristic and context for image-text retrieval. *IEEE Trans. Circuits Syst. Video Technol.* 2022, 32, 8037–8050.
29. Liu, A.; Zhai, Y.; Xu, N.; Nie, W.; Li, W.; Zhang, Y. Region-aware image captioning via interaction learning. *IEEE Trans. Circuits Syst. Video Technol.* 2021, 32, 3685–3696.
30. Wang, F.; Wang, H.; Zhou, X.; Fu, R. A driving fatigue feature detection method based on multifractal theory. *IEEE Sens. J.* 2022, 22, 19046–19059.
31. Zhu, Y.; Huang, R.; Wu, Z.; Song, S.; Cheng, L.; Zhu, R. Deep learning-based predictive identification of neural stem cell differentiation. *Nat. Commun.* 2021, 12, 2614.
32. Tan, X.; Lin, J.; Xu, K.; Chen, P.; Ma, L.; Lau, R. Mirror detection with the visual chirality cue. *IEEE Trans. Pattern Anal. Mach. Intell.* 2022, 45, 3492–3504.
33. Kalbhor, M.; Shinde, S.V.; Jude, H. Cervical cancer diagnosis based on cytology pap smear image classification using fractional coefficient and machine learning classifiers. *TELKOMNIKA (Telecommun. Comput. Electron. Control)* 2022, 20, 1091–1102.
34. Lavanya Devi, N.; Thirumurugan, P. Cervical cancer classification from pap smear images using modified fuzzy C means, PCA, and KNN. *IETE J. Res.* 2022, 68, 1591–1598.

35. Alquran, H.; Mustafa, W.A.; Qasmieh, I.A.; Yacob, Y.M.; Alsalatie, M.; Al-Issa, Y.; Alqudah, A.M. Cervical cancer classification using combined machine learning and deep learning approach. *Comput. Mater. Contin* 2022, 72, 5117–5134.
36. Kalbhor, M.; Shinde, S.; Popescu, D.E.; Hemanth, D.J. Hybridization of Deep Learning Pre-Trained Models with Machine Learning Classifiers and Fuzzy Min–Max Neural Network for Cervical Cancer Diagnosis. *Diagnostics* 2023, 13, 1363.
37. Tanimu, J.J.; Hamada, M.; Hassan, M.; Kakudi, H.; Abiodun, J.O. A machine learning method for classification of cervical cancer. *Electronics* 2022, 11, 463.
38. Quinlan, S.; Afli, H.; O'Reilly, R. A Comparative Analysis of Classification Techniques for Cervical Cancer Utilising at Risk Factors and Screening Test Results. In *Proceedings of the AICS 2019 Artificial Intelligence and Cognitive Science (AICS) Proceedings of The 28th Irish Conference on Artificial Intelligence and Cognitive Science, Dublin, Ireland, 7–8 December 2019*; pp. 400–411.
39. Gowri, K.; Saranya, M. Cervical Cancer Prediction using Outlier deduction and Over sampling methods. *Int. J. Innov. Res. Eng.* 2022, 3, 186–190.
40. Abdoh, S.F.; Rizka, M.A.; Maghraby, F.A. Cervical cancer diagnosis using random forest classifier with SMOTE and feature reduction techniques. *IEEE Access* 2018, 6, 59475–59485.
41. Ijaz, M.F.; Attique, M.; Son, Y. Data-driven cervical cancer prediction model with outlier detection and over-sampling methods. *Sensors* 2020, 20, 2809.
42. Jahan, S.; Islam, M.S.; Islam, L.; Rashme, T.Y.; Prova, A.A.; Paul, B.K.; Islam, M.M.; Mosharof, M.K. Automated invasive cervical cancer disease detection at early stage through suitable machine learning model. *SN Appl. Sci.* 2021, 3, 806.
43. Al Mudawi, N.; Alazeb, A. A model for predicting cervical cancer using machine learning algorithms. *Sensors* 2022, 22, 4132.

Retrieved from <https://encyclopedia.pub/entry/history/show/113425>