Pertanika Proceedings 1 (5): 83 - 88 (2025)

PERTANIKA PROCEEDINGS

Journal homepage: http://www.pertanika.upm.edu.my/

A Systematic Review on Cervical Precancerous Classification Techniques and Diagnosis Using Deep Learning Models

P. Leela[1]*, K. Hari Krishna[1] and K. K. Baseer[2]

[1]*Department of Computer Science and Engineering, Mohan Babu University, Tirupati, Andhra Pradesh 517102, India*

[2]*Computer Science and Engineering, GITAM School of Technology, GITAM (Deemed to be University), Bengaluru, Karnataka 561203, India*

ABSTRACT

Many problems that are hard to tackle with conventional artificial intelligence (CAI) techniques can be handled using DL techniques. Deep learning (DL) is a widely employed technique in many different domains, including medicine, these days. With almost 700 deaths each day, cervical cancer (CRC) is one of the most common causes of death for women, after breast cancer. By 2040, this figure is projected to reach 400,000 each year. Nonetheless, cancer is entirely treatable if it is found in its early and precancerous stages. Decision support systems that use artificial intelligence (AI) to detect aberrant cervical cells were created since manual identification can be challenging, time-consuming, and prone to errors, and it requires skilled medical personnel. The review study comprises research that uses AI in imaging modalities like magnetic oscillation imaging, computed imaging, positron emission imaging scans, colposcopy, sociodemographic information and other risk variables, histological analysis, and Pap Smear tests. However, human error causes a high false rate in the manual screening method. Strict inclusion and exclusion criteria were adhered to, including academic publications, deep learning-based techniques, and search windows from the previous ten years. In this article, various examine cutting-edge methods that analyze cervical cytology and screening pictures using DL techniques are studied. It examines and talks about pertinent DL approaches, their structures, categorization schemes, and cervical cytology.

Keywords: Cervical cancer, computer aided diagnosis, conventional artificial intelligence, deep learning, machine learning

ARTICLE INFO

Article history:
Received: 29 August 2025
Published: 31 October 2025

DOI: https://doi.org/10.47836/pp.1.5.016

E-mail addresses:
leelapachappa@gmail.com (P. Leela)
khk396@gmail.com (K. Hari Krishna)
drkkbaseer@gmail.com (K. K. Baseer)
* Corresponding author

INTRODUCTION

The fourth most frequent malignancy in women, CrC (cervical cancer) is a serious global public health concern. Globally, CrC caused over 604,000 new cases and almost 342,000 fatalities in 2020 (World Health Organization, 2022). An essential but insufficient prerequisite for CrC is infection with a high-risk oncogenic human papillomavirus (HPV). S. F. Abdoh et al. (2018) Eighty percent of women who engage in sexual activity will eventually contract HPV,a highly prevalent STD. Cervical intraepithelial neoplasia (CIN), a type of premalignant cervical lesion, occurs beforeCrC. Since it usually takes ≥10 years to advance to CrC. The CrC is appropriate for early detection, screening, and prevention. Asymptomatic women with precancerous (PC) cervical disease or CrC can be detected early with routine screening for the condition. The purpose of CrC screening is to find women who have PC and treat them before CrC develops (Abdoh et al., 2018). This will lower the number of people who die from CrC and, ideally, eliminate most CrC-related deaths. Common methods for detecting CrC or PC include cytology, colposcopy/cervicography, sociodemographic traits, and histology; Squamous, endocervical, and inflammatory cells are typically visually assessed as part of the traditional cytology-based CrC screening process. Figure 1 provides the papers published from 2012 to 2024.

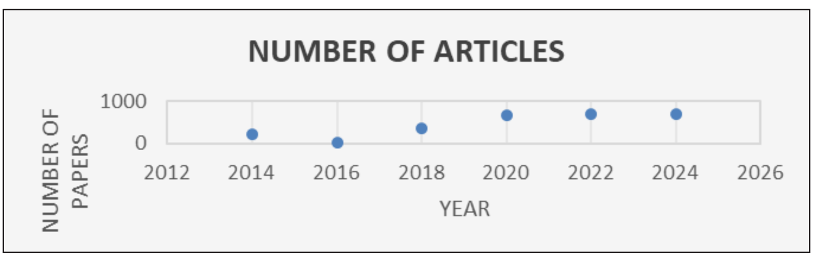

Figure 1. Number of articles released with respect to the year

Cervical Cancer Features

Li et al. (2020) used a dataset that was gathered at the Universitario de Caracas Hospital located in Caracas, and released on the California University in Irvine (UCI) repository. The dataset included 12 features for each of the 858 cases, including demographic data, behaviors, and past medical records. Due to privacy concerns, a few instances chose not to answer certain questions; therefore, the dataset has a large number of missing values (Rahaman et al., 2020). The features of the dataset, the overall number of written records, and the missing values for each characteristic are displayed in Table 1. Data post-processing is therefore essential to overcoming these constraints. Feature selection (FS), enhancement, feature fusion, standardization, normalization, regularization, and outlier removal are some of the data postprocessing methods. Classification refers to categorizing various conditions for a particular problem using machine learning (ML) and deep learning (DL) techniques. Traditional ML techniques such as support vector machine (SVM), random forest (RF), k-nearest neighbor (KNN), decision tree (DT), logistic regression (LR), linear discriminant analysis (LDA), and artificial neural network (ANN) are popular for decision-making algorithms. In contrast, the DL models yield improved performance with increasing data size as they produce an abstract representation of the raw information that is compressed over many layers of an ANN.

Table 1
Features, overall patients and missing data of the dataset

No	Characteristics	Dataset	Values
1	Group of Age category	952	0
2	Overall intimate partners	934	22
3	First sexual intercourse	962	6
4	Overall Pregnancy	954	45
6	Hormonal Contraceptives	852	98
7	IUD	832	105

LITERATURE REVIEW

In order to enable learning, DL consists of multiple layers of computer models for data processing that display input data at varying levels of abstraction. In order to solve real-world issues. Medical image processing can be greatly advanced using DL approaches in a number of clinical and scientific fields, including the detection and diagnosis of different types of cancer. The various forms of cancer, including colorectal, breast, bladder, and CC, are among the top causes of death worldwide. Asymptomatic women with precancerous (PC) cervical disease or CC can be detected early with routine screening for the condition. The purpose of CrC screening is to find women who have PC and treat them before CrC develops. Each screening and diagnosis method is briefly described in the section that follows.

CLASSIFICATION METHODS

Khamparia et al. (2020) provided a brief overview of earlier colposcopic image processing research. Next, a number of often-used feature fusion techniques were discussed. Lastly, we examine the relevant existing research and provide a quick introduction to graph neural networks (Khamparia et al., 2020). The various articles that summarize the use of AI for CRC detection and liquid-based cytology/PSC pictures. The SIPaK MeD dataset has five subclasses of cells: superficial–intermediate (SPIM), parabasal (PB), metaplastic (MP), dyskeratotic (DKT), and koilocytotic (KCT). These subclasses comprise normal, benign, and pathological cells. There are seven classes in the Herlev dataset: columnar, superficial squamous (SFS), intermediate squamous (IMS), and many class classification (MCC) problems are used to group cancer, light dysplastic (LDP), moderate dysplastic (MDP), and severe dysplastic (SDP) which is shown in Figure 2.

Khare et al. (2024) extract and categorize the DFCNN features, whereas Chitra and Kumar used a mutation-based atom search optimization approach to build a DenseNet model. The extracted DFCNN features were obtained using pretrained InceptionV3, Inception ResNet, and MobileNet models. Zhao et al. (2022) used a fuzzy distance-based ensemble to detect NC and ANC in the Herlev dataset and five cells in the SIPaK MeD dataset, which is shown in Table 2.

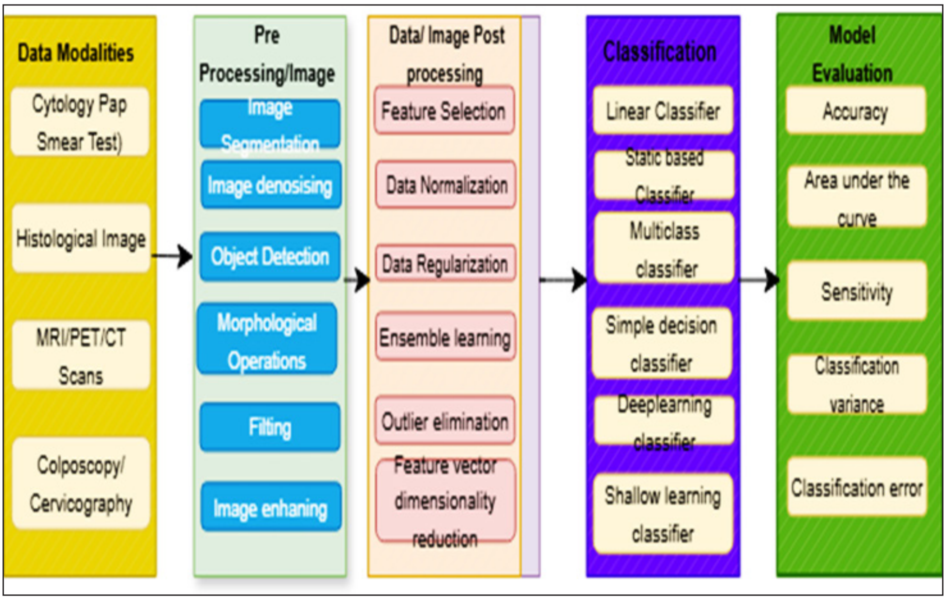

Figure 2. The process of pipeline for preprocessing, feature extraction and classification

Table 2
Cervical precancerous classification methods

Ref	**Concept**	**Figure**	**Name of the Dataset**	**Name of the Classes**	**User Access**	**Classifier**	**Validation**	**ACC**	**SPE**
Mustafa et al., 2022	-	978	MEDe SIOak	KcT, mP, DKTSFIM	Free	SVM	Holding out	99.20	-
Kaur et al., 2022	460	974	Mendeley	SCC, HSIL, LSIL, NILM	Free	-		99.46	99.56
Chitra & Kumar, 2021	-	927	Herlev	NC, LDP, SDP, Carcinoma	Free	CNN	Holding out	98.48	98.86

The extracted DFCNN features using pretrained InceptionV3, Inception ResNet, and MobileNet models. fuzzy distance-based ensemble to detect NC and ANC in the Herlev dataset and five cells in the SIPaK MeD dataset.

The normalized features from the ResNet-50 model resulted in the best performance in classifying seven and five-cell classes of Herlev and SIPaK MeD datasets using a fuzzy min–max neural network, respectively. The authors tested five class classifications on the SIPaK MeD pap-smear data set using 40 models based on CNN. Many considerations are made when utilizing deep learning which is shown in Figure 3.

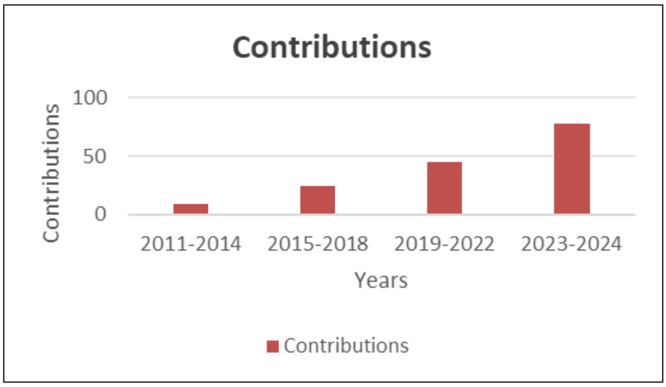

Figure 3. Bar charts representing the reviews

CONCLUSION AND FUTURE WORK

According to the WHO, cervical cancer can be prevented. A detailed investigation of an AI-assisted automated decision support system for CrC detection using several modalities is presented in this review paper. According to our review, automated feature extraction from imaging methods has been the method of choice for CrC. Better characterization of healthy and CrC cases is possible because of the deep characteristics taken from DL models. This study also discovered that clinicians oppose current automated systems because they lack model explainability and uncertainty quantification. Because deployment requires flexibility and data privacy, meta learning and federated learning shall also be investigated. Colposcopic biopsies can be time-consuming and uncomfortable. As previously stated, colposcopy is subjective and has a low sensitivity for accurately detecting high-grade illness. Studies have indicated that obtaining several more biopsies can enhance the effectiveness of colposcopy to improve prediction made by the glass-box-based AI-assisted system. Feature importance, feature visualization, and local/global explanations provide model explainability for clinical or engineering features.

ACKNOWLEDGEMENT

The authors sincerely thank Dr. K. Hari Krishna and Dr. K. K. Baseer for their support in this research project.

REFERENCES

Abdoh, S. F., Rizka, M. A., & Maghraby, F. A. (2018). Cervical cancer diagnosis using random forest classifier with SMOTE and feature reduction techniques. *IEEE Access*, *6*, 59475-59485. https://doi.org/10.1109/ACCESS.2018.2874063

Chitra, B., & Kumar, S. S. (2022). Recent advancement in cervical cancer diagnosis for automated screening: a detailed review. *Journal of Ambient Intelligence and Humanized Computing*, *13*(1), 251-269. https://doi.org/10.1007/s12652-021-02899-2

Kaur, B., Kumar, S., & Kaushik, B. K. (2022). Recent advancements in optical biosensors for cancer detection. *Biosensors and Bioelectronics*, *197*, Article 113805. https://doi.org/10.1016/j.bios.2021.113805.

Khamparia, A., Gupta, D., De Albuquerque, V. H. C., Sangaiah, A. K., & Jhaveri, R. H. (2020). Internet of health things-driven deep learning system for detection and classification of cervical cells using transfer learning. *Journal of Supercomputing*, *76*(11), 8590–8608. https://doi.org/10.1007/s11227-020-03159-4

Khare, S. K., Blanes-Vidal, V., Booth, B. B., Petersen, L. K., & Nadimi, E. S. (2024). A systematic review and research recommendations on artificial intelligence for automated cervical cancer detection. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *14*(6), Article e1550. https://doi.org/10.1002/widm.1550

Li, Y., Chen, J., Xue, P., Tang, C., Chang, J., Chu, C., Ma, K., Li, Q., Zheng, Y., & Qiao, Y. (2020). Computer-aided cervical cancer diagnosis using time-lapsed colposcopic images. *IEEE Transactions on Medical Imaging*, *39*(11), 3403-3415. https://doi.org/10.1109/TMI.2020.2994778

Mustafa, W. A., Alias, N. A., Jamlos, M. A., Ismail, S., & Alquran, H. (2022). A recent systematic review of cervical cancer diagnosis: detection and classification. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *28*(1), 81-96. https://doi.org/10.37934/araset.28.1.8196

Rahaman, M. M., Li, C., Wu, X., Yao, Y., Hu, Z., Jiang, T., Li, X., & Qi, S. (2020). A survey for cervical cytopathology image analysis using deep learning. *IEEE Access*, *8*, 61687-61710. https://doi.org/10.1109/ACCESS.2020.2983186

Zhao, S., He, Y., Qin, J., & Wang, Z. (2022). A semi-supervised deep learning method for cervical cell classification. *Analytical Cellular Pathology*, *2022*(1), Article 4376178. https://doi.org/10.1155/2022/4376178