

Pap Smear Images Classification with Machine Learning

Subjects: **Engineering, Biomedical**

Contributor: Nur Ain Alias , Wan Azani Mustafa , Mohd Aminudin Jamlos , Hiam Alquran , Hafizul Fahri Hanafi , Shahrina Ismail , Khairul Shakir Ab Rahman

Cervical cancer is regularly diagnosed in women all over the world. This cancer is the seventh most frequent cancer globally and the fourth most prevalent cancer among women. Automated and higher accuracy of cervical cancer classification methods are needed for the early diagnosis of cancer. In addition, routine Pap smears could enhance clinical outcomes by facilitating the early diagnosis of cervical cancer. Liquid-based cytology (LBC)/Pap smears for advanced cervical screening is a highly effective precancerous cell detection technology based on cell image analysis, where cells are classed as normal or abnormal. Computer-aided systems in medical imaging have benefited greatly from extraordinary developments in artificial intelligence (AI) technology.

cervical cancer

cell classification

Pap Smear Images

1. Introduction

Cervical cancer is a major disease that seriously threatens women's health ^{[1][2]}. This cancer is also known as the second most commonly affected and killer type of cancer among women around the world ^[3]. It results from a chronic infection of the skin and mucosal cells in women's vaginal regions. The fact that this cancer does not manifest any signs when it first appears is the most alarming feature of it ^[4]. In a report by Elakkiya et al. ^[5], it has been mentioned that this type of cancer is curable with early detection in the early stage. Unfortunately, the mortality rate is getting higher among women around the world that have been affected by this cancer ^{[6][7][8]}. The traditional method of manual inspection, also known as a Pap smear examination, is inaccurate due to human error that may lead to a false patient diagnosis ^{[9][10]}. The technology for automated cervical cancer screening is indeed very significant for lessening the risk of cervical cancer. However, the existing approach using machine learning has drawbacks, including poor generalisation capacity in complicated situations, as well as low efficiency, accuracy, and generalisation ability ^[11]. Several studies have attempted to investigate the ability of machine learning to classify cervical cancer cells for the purpose of enhancing manual screening ^{[10][12][13]}. The most often used approach for predicting characteristics from a high-dimensional collection of cancer imaging data is the random forest approach ^{[14][15][16]}. However, if a large number of decision trees are utilised, the random forest approach might become too sluggish and ineffective for real-time predictions ^[10]. In addition, current classification approaches, such as deep learning (DL) or hand-crafted techniques, mostly rely on single detection structures and have high processing complexity and low accuracy ^[17].

2. Pap Smear Images Classification

The cancer disease’s significance has increased, as public health worries about the region’s development and success. Microscopic image-based analysis has been extensively used in pathological research and disease diagnosis. However, the misauthentication of cell lines due to pathologists’ errors has been identified as a severe issue. Therefore, a comparative evaluation of the proposed model was conducted to illustrate the utility of feature selection and class imbalance based on the classifier’s accuracy, sensitivity, precision, F-measure, and specificity.

Classification of Cells Based on Machine Learning Approach

There are numerous approaches that have been done by previous researchers in the area of cell classification. The methodology and results of the approaches are summarised to make it easier to compare the findings of the studies. **Table 1** illustrates the summary of the prevailing works related to the classification of cervical cancer cells using machine learning.

Table 1. Summary of the prevailing works.

Title	Class and Database	Methodology	Results
CVM-Cervix: A hybrid cervical Pap smear image classification framework using CNN, visual transformer, and multilayer perceptron [18]	Database: CRIC dataset, SIPaKMeD dataset and combination of CRIC and SIPaKMeD Class: CRIC-6 class, SIPaKMeD 5 Class	Framework CVM-Cervix based on deep learning. Type: Machine learning—Neural network	Effective and potential of the proposed CVM-Cervix proved.
A Comparative Analysis of Deep Learning Models for Automated Cross-Preparation Diagnosis of Multi-Cell Liquid Pap Smear Images [19]	Database: Thin Prep Pap dataset Class: 4 Bethesda class	An ensemble novel convolutional neural network (CNN) and a CNN with autoencoder (AE). Type: Machine learning—Neural network	All models’ accuracy >90%. The individual transfer model had high variability in performance, while CNN and AE CNN did not. ResNet101 accuracy is 92.65%.
Multi-class nucleus detection and classification using a deep convolutional neural network with enhanced high dimensional	Database: Herlev dataset, SIPaKMeD dataset, CRIC dataset Class: Herlev-7	Segmentation—hybrid system that incorporates two binary image patches obtained by a 19-layered convolutional neural network (ConvNet) model with an enhanced deep	Accuracy 99.12% Specificity 99.45% Sensitivity 99.25% Execution time 99.6248 s The proposed model is more effective compared to existing approaches.

Title	Class and Database	Methodology	Results
dissimilarity translation model on cervical cells [20]	class, SIPaKMeD-5 Class, CRIC-6 class	high dimensional dissimilarity translation (HDDT). A Pre trained Resnet-50 model. T-distribution stochastic neighbour embedding (t- SNE) for down-sampled. Classification using a multi-class weighted kernel extreme learning machine (WKELM) classifier via a sparse multicanonical correlation (SMCCA) method. Type: Machine learning— Neural network	
Hybrid Loss-Constrained Lightweight Convolutional Neural Networks for Cervical Cell Classification [21]	Database: Herlev dataset Class: Herlev-7 class	Classification using hybrid loss function with label smoothing to improve the distinguishing power of lightweight convolutional neural networks (CNN) Type: Machine learning— Neural network	ShufflenetV2 results:- Accuracy 96.18% Precision 96.30% Recall 96.23% Specificity 99.08% GhostnetV2 results:- Accuracy 96.39% Precision 96.42% Recall 96.39% Specificity 99.09%
Detection of cervical cells based on improved SSD network [22]	Database: Herlev dataset Class: Herlev-7 class	Integration of Single Shot MultiBox Detector with the positive and negative features to address the problem of insufficient sensitivity for small objects. Type: Machine learning— Neural network	Accuracy 90.80% Mean average precision (mAP) is 81.53%, which is 7.54% and 4.92% higher than YOLO and classical SSD.
Detection of cervical cancer cells in a complex situation based on improved YOLOv3 network [1]	Database: Herlev dataset Class: Herlev-7 class	Detection using the YOLO algorithm. Feature extraction generalization by adding the dense block and S3Pool algorithm on the basis of the feature extraction network DarkNet-53.	Mean average precision (mAP) is 78.87%, which is 7.54% and 4.92% higher than YOLO (You Only Look Once) and classical SSD

Title	Class and Database	Methodology	Results
		Clustering algorithm of improved algorithm k-means++.	
		Type: Machine learning—Neural network	
Pap smear-based cervical cancer detection using residual neural networks deep learning architecture [23]	Database: Mendeley LBC SIPaKMeD dataset Class: Mendeley LBC SIPaKMeD—4 class	Data augmentation module of DTWCT module and convolutional neural networks (CNN). Classification using ResNet 18, which defines four classes of sources for Pap smear cell images. Type: Machine learning—Neural network	Average Pap smear detection index (PDI) is 99%.
Cervical cell multi-classification algorithm using global context information and attention mechanism [3]	Database: SIPaKMeD dataset Class: SIPaKMeD—5 class	Convolutional neural network (L-PCNN) that integrates global context information and attention mechanism. Improved ResNet-50 backbone network for feature extraction. Type: Machine learning—Neural network	Accuracy 98.89%. Sensitivity 99.9%. Specificity 99.8%. F-measure 99.89%.
DeepCyto: a hybrid framework for cervical cancer classification using deep feature fusion of cytology images [24]	Database: Herlev dataset, SIPaKMeD dataset, LBC dataset Class: Herlev-7 class, SIPaKMeD-5 class, LBC-4 class	Novel classification using DeepCyto. Principal component analysis and machine learning ensemble for classification of Pap smear images. Artificial neural network with feature fusion vectors as an input for classification. Type: Machine learning—Neural network	DeepCyto is a powerful tool for precise feature extraction and Pap smear image classification.
Classification of Cervical Cytology Overlapping Cell Images with Transfer Learning Architectures [25]	Database: Cervix93 cervical cytology image	Transfer learning using deep learning convolutional neural network.	Accuracy 99.03%. Kappa coefficient showing perfect agreement. AlexNet proved a successful

Title	Class and Database	Methodology	Results
	Class: 3 class	Cutting edge pretrained networks: AlexNet, ImageNet, and Places 365. Type: Machine learning—Neural network	assistive tool for cervical cancer detection.
Optimal deep convolution neural network for cervical cancer diagnosis model [8]	Database: Herlev dataset Class: Herlev-7 class	Detection using intelligent deep convolutional neural network. Classification (IDCNN-CDC) model using biomedical Pap smear images. Noise removal using Gaussian Filter. Segmentation using the Tsallis entropy technique with the dragonfly optimization. Deep learned feature using SqueezeNet. Classification using weighted extreme learning machine (ELM). Type: Machine learning—Neural network	Higher performance of the proposed technique in terms of sensitivity, specificity, accuracy, and F-Score.
Modified metaheuristics with stacked sparse denoising autoencoder model for cervical cancer classification [9].	Database: Herlev dataset Class: Herlev-7 class	Novel Modified Firefly Optimization Algorithm with Deep Learning-enabled cervical cancer classification (MFFOA-DL3) model. Noise removal using Bilateral Filtering (BF)-based. Segmentation technique of Kapur's entropy-based image to define affected area. Generate feature vectors using EfficientNet. Classification of the cell using MFFOA with Stacked Sparse Denoising Autoencoder (SSDA) model.	The findings of a comprehensive comparison investigation revealed that the MFFOA-DL3 model outperformed other recent approaches.

Title	Class and Database	Methodology	Results
		Type: Machine learning— Neural network	
Imaging based cervical cancer diagnostics using small object detection—generative adversarial networks [5]	Database: Herlev dataset, Colposcopy images, Clinical references Class: not applicable	An effective hybrid deep learning technique using Small-Object Detection-Generative Adversarial Networks (SOD-GAN) with Fine-tuned Stacked Autoencoder (F-SAE). Generation and discrimination of the cervical cell using Region-based Convolutional Neural Network (RCNN).	The proposed method identifies and classifies cervical premalignant and malignant diseases based on deep characteristics without the necessity for initial classification and segmentation.
		Type: Machine learning— Neural network	
Cervical cancer diagnosis based on modified uniform local ternary patterns and feed forward multilayer network optimized by genetic algorithm [26]	Database: Herlev dataset Class: Herlev-7 class	Segmentation of the image using a thresholding approach. Feature extraction by applying a texture descriptor titled modified uniform local ternary patterns (MULTP). Classification of the cell using an optimized multilayer feed-forward neural network.	MULTP, the proposed texture descriptor, is a generic operator that may be used to characterise texture features of images in numerous computer vision issues. In addition, the suggested optimization approach may be utilised to increase performance in deep networks.
		Type: Machine learning— Neural network	
Early cervical cancer diagnosis using Sooty tern-optimized CNN-LSTM classifier [11]	Database: Herlev dataset Class: Herlev-7 class	Augmentation process of image enhancement, image flipping, and image rotating to reduce the number of parameters necessary. Segmentation of the cancer-affected regions with the help of kernel weighted fuzzy local information c-means clustering (KWFLICM) model. Classification using the	Accuracy 99.80%. Specificity 99%. Sensitivity 98.83%. F-Score 97.8. Improvement of 28.5% better than Random Forest and 19.46% better than ensemble classifier.

Title	Class and Database	Methodology	Results
		Sooty Tern Optimization (STO) algorithm with CNN-based long short-term memory classifier (CNN-LSTM). Type: Machine learning—Neural network	
Hybrid Model for Detection of Cervical Cancer Using Causal Analysis and Machine Learning Techniques [10]	Database: Private Class: not applicable	Boruta analysis and SVM method for an efficient feature selection and prediction of the model for the cervical cell dataset. Type: Machine learning—Linear model	Boruta analysis shows a better performance approach compared to the existing techniques available.
Cervical Cancer Classification Using Combined Machine Learning and Deep Learning Approach [27]	Database: Herlev dataset Class: Herlev-2 class	Feature extraction using ResNet-101. Classification using Support vector Machine (SVM). Type: Machine learning—Linear model	Accuracy 97.30%.
Auxiliary classification of cervical cells based on the multi-domain hybrid deep learning framework [17]	Database: Herlev dataset, SIPaKMeD dataset, BJTU dataset Class: Herlev-2&7 class, SIPaKMeD-5 Class, BJTU-7 class	Deep features extraction using deep Convolutional Neural Network of pretrained Visual Geometry Group-19 (VGG-19). Hand-crafted images undergo the process of feature selection, clustering, and dimensionality reduction. Classification using a Support Vector Machine (SVM) classifier. Type: Machine learning-Linear model	Accuracy 98.70%. Sensitivity 98.20%. Specificity 98.90%. The suggested novel screening methodology is promising for early cervical cancer detection, with multi-domain and hybrid characteristics proving realistic in clinical practise.
An Evaluation of Computational Learning-based Methods for the Segmentation of Nuclei in Cervical Cancer Cells	Database: Z-Stack cellular microscopy proliferation images provided	Machine learning architecture of Random Forest, Ada Boost, and MLP algorithm.	All machine learning architectures gave outstanding nuclei segmentation in cervical cancer cells but did not solve

Title	Class and Database	Methodology	Results
from Microscopic Images [28]	by the HCS Pharma Class: not applicable	Type: Machine learning—Nonlinear model	the overlapping nuclei and Z-stack segmentation problems.
Prognosis of Cervical Cancer Disease by Applying Machine Learning Techniques [4]	Database: Dataset of 858 cervical cancer patients with 36 risk factors and one outcome variable Class: not applicable	Analysis of the different supervised machine learning techniques. The classification algorithm used Artificial Neural Network, Bayesian Network, SVM, Random Tree, Logistic Tree and XG-Boost Tree. Selection algorithm for feature selection: relief rank, wrapper method, and LASSO regression. Type: Machine learning—Nonlinear model	Maximum accuracy achieved using XG-Boost with complete features 94.94%. This approach offers much potential for clinical use and cervical cancer cell detection.
Is the aspect ratio of cells important in deep learning? A robust comparison of deep learning methods for multi-scale cytopathology cell image classification: From convolutional neural networks to visual transformers [29]	Database-SIPaKMeD dataset Class: not applicable	Twenty-two deep learning models were used to classify the cervical cancer cells into two categories of standard and scaled datasets. Type: Machine learning—Nonlinear model	Deep learning models are robust to changes in the aspect ratio of cervical cells in cervical cytopathological images.
A Fast Hybrid Classification Algorithm with Feature Reduction for Medical Images [30]	Database: Herlev dataset Class: Herlev-7 class	A novel fast hybrid fuzzy classification algorithm with feature reduction for medical images. Integration of quantum-based grasshopper computing algorithm (QGH) with a fuzzy clustering technique for feature extraction. The second integration of the fusion technique utilises QGH with the fuzzy c-means algorithm to determine the best features.	Established the importance of the feature selection on the accuracy of the proposed classifier
[11]			in the area es for the , such as hers. For on (STO)

algorithm with a CNN-based long short-term memory classifier (CNN-LSTM) and achieved better performance results compared to other literature reviews. The results achieved in the study have shown that the accuracy is 99.80%, specificity is 99%, sensitivity is 98.83%, and F-score is 97.8. Their findings show an improvement of

Title	Class and Database	Methodology	Results	
		Type: Machine learning— Nonlinear model		stent with
				specificity,
				evolutional
				improved
				usion was
				al images.
				ation and
				necessity
				ethods of
				sification,
Cervical Cancer Classification from Pap Smear Images Using Modified Fuzzy C Means, PCA, and KNN [7]	Database: Herlev dataset Class: Herlev-7 class	Geometrical and feature extraction using a novel approach of modified fuzzy c-means. [5] Augmentation of the images using Principal Component Analysis (PCA) to maintain the uncorrelated features and thus reduce the algorithm processing time. Classification of the Pap smear image into normal and abnormal cells using K Nearest Neighbour (KNN).	Minimum accuracy 94.15%. Maximum accuracy 96.28%. Average accuracy 94.86%. Sensitivity 97.96%. Specificity 83.65%. F1-Score 96.87%. Precision 96.31%.	
		Type: Machine learning— Nonlinear model		lex
				1.
				ol. 2021,
A Semi-supervised Deep Learning Method for Cervical Cell Classification [31]	Database: Herlev dataset, SIPaKMeD dataset Class: Herlev-7 class, SIPaKMeD-5 Class	A novel manual features and voting mechanism to achieve data expansion in semi-supervised learning. Clarity function to filter out higher-quality images, annotating a small amount of the high-quality images, and voting mechanism for balancing and training data.	Accuracy 91.94%.	Cell
		Type: Machine learning— Classifier		gnosis
				omput.

5. Elakkiya, R.; Teja, K.S.S.; Jegatha Deborah, L.; Bisogni, C.; Medaglia, C. Imaging based cervical cancer diagnostics using small object detection—Generative adversarial networks. *Multimed. Tools Appl.* 2022, 81, 191–207.
6. Lia, M.; Horn, L.C.; Sodeikat, P.; Höckel, M.; Aktas, B.; Wolf, B. The diagnostic value of core needle biopsy in cervical cancer: A retrospective analysis. *PLoS ONE* 2022, 17, e0262257.
7. Lavanya Devi, N.; Thirumurugan, P. Cervical Cancer Classification from Pap Smear Images Using Modified Fuzzy C Means, PCA, and KNN. *IETE J. Res.* 2021, 68, 1591–1598.
8. Waly, M.I.; Sikkandar, M.Y.; Aboamer, M.A.; Kadry, S.; Thinnukool, O. Optimal deep convolution neural network for cervical cancer diagnosis model. *Comput. Mater. Contin.* 2022, 70, 3297–3309.

9. Vaiyapuri, T.; Alaskar, H.; Syed, L.; Aljohani, E.; Alkhayyat, A.; Shankar, K.; Kumar, S. Modified metaheuristics with stacked sparse denoising autoencoder model for cervical cancer classification. *Comput. Electr. Eng.* 2022, 103, 108292.
10. Lilhore, U.K.; Poongodi, M.; Kaur, A.; Simaiya, S.; Algarni, A.D.; Elmannai, H.; Vijayakumar, V.; Tunze, G.B.; Hamdi, M. Hybrid Model for Detection of Cervical Cancer Using Causal Analysis and Machine Learning Techniques. *Comput. Math. Methods Med.* 2022, 2022, 4688327.
11. Chitra, B.; Kumar, S. Early cervical cancer diagnosis using Sooty tern-optimized CNN-LSTM classifier. *Int. J. Imaging Syst. Technol.* 2022, 32, 1846–1860.
12. Kaur, H.; Sharma, R.; Kaur, L. Automated Cervical Cancer Image Analysis using Deep Learning Techniques from Pap-Smear Images: A Literature Review. In *Proceedings of the 2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, Noida, India, 3–4 September 2021; pp. 1–7.
13. Diniz, D.N.; Rezende, M.T.; Bianchi, A.G.C.; Carneiro, C.M.; Luz, E.J.S.; Moreira, G.J.P.; Ushizima, D.M.; de Medeiros, F.N.S.; Souza, M.J.F. A deep learning ensemble method to assist cytopathologists in pap test image classification. *J. Imaging* 2021, 7, 111.
14. Jahan, S.; Islam, M.D.S.; Islam, L.; Rashme, T.Y.; Prova, A.A.; Paul, B.K.; Islam, M.D.M.; Mosharof, M.K. Automated invasive cervical cancer disease detection at early stage through suitable machine learning model. *SN Appl. Sci.* 2021, 3, 806.
15. Diniz, D.N.; Rezende, M.T.; Bianchi, A.G.C.; Carneiro, C.M.; Ushizima, D.M.; de Medeiros, F.N.S.; Souza, M.J.F. A hierarchical feature-based methodology to perform cervical cancer classification. *Appl. Sci.* 2021, 11, 4091.
16. Ali, M.M.; Ahmed, K.; Bui, F.M.; Paul, B.K.; Ibrahim, S.M.; Quinn, J.M.W.; Moni, M.A. Machine learning-based statistical analysis for early stage detection of cervical cancer. *Comput. Biol. Med.* 2021, 139, 104985.
17. Zhang, C.; Jia, D.; Li, Z.; Wu, N. Auxiliary classification of cervical cells based on multi-domain hybrid deep learning framework. *Biomed. Signal Process. Control* 2022, 77, 103739.
18. Liu, W.; Li, C.; Xu, N.; Jiang, T.; Rahaman, M.M.; Sun, H.; Wu, X.; Hu, W.; Chen, H.; Sun, C.; et al. CVM-Cervix: A hybrid cervical Pap-smear image classification framework using CNN, visual transformer and multilayer perceptron. *Pattern Recognit.* 2022, 130, 108829.
19. Karasu Benyes, Y.; Welch, E.C.; Singhal, A.; Ou, J.; Tripathi, A. A Comparative Analysis of Deep Learning Models for Automated Cross-Preparation Diagnosis of Multi-Cell Liquid Pap Smear Images. *Diagnostics* 2022, 12, 1838.
20. Karri, M.; Annavarapu, C.S.R.; Mallik, S.; Zhao, Z.; Acharya, U.R. Multi-class nucleus detection and classification using deep convolutional neural network with enhanced high dimensional dissimilarity translation model on cervical cells. *Biocybern. Biomed. Eng.* 2022, 42, 797–814.

21. Chen, W.; Shen, W.; Gao, L.; Li, X. Hybrid Loss-Constrained Lightweight Convolutional Neural Networks for Cervical Cell Classification. *Sensors* 2022, 22, 3272.
22. Jia, D.; Zhou, J.; Zhang, C. Detection of cervical cells based on improved SSD network. *Multimed. Tools Appl.* 2022, 81, 13371–13387.
23. Sellamuthu Palanisamy, V.; Athiappan, R.K.; Nagalingam, T. Pap smear based cervical cancer detection using residual neural networks deep learning architecture. *Concurr. Comput. Pract. Exp.* 2022, 34, e6608.
24. Shinde, S.; Kalbhor, M.; Wajire, P. DeepCyto: A hybrid framework for cervical cancer classification by using deep feature fusion of cytology images. *Math. Biosci. Eng.* 2022, 19, 6415–6434.
25. Mulmule, P.V.; Kanphade, R.D. Classification of Cervical Cytology Overlapping Cell Images with Transfer Learning Architectures. *Biomed. Pharmacol. J.* 2022, 15, 277–284.
26. Fekri-Ershad, S.; Ramakrishnan, S. Cervical cancer diagnosis based on modified uniform local ternary patterns and feed forward multilayer network optimized by genetic algorithm. *Comput. Biol. Med.* 2022, 144, 105392.
27. 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.
28. Maylaa, T.; Windal, F.; Benhabiles, H.; Maubon, G.; Maubon, N.; Vandenhaute, E.; Collard, D. An Evaluation of Computational Learning-based Methods for the Segmentation of Nuclei in Cervical Cancer Cells from Microscopic Images. *Curr. Comput. Aided Drug Des.* 2022, 18, 81–94.
29. Liu, W.; Li, C.; Rahaman, M.M.; Jiang, T.; Sun, H.; Wu, X.; Hu, W.; Chen, H.; Sun, C.; Yao, Y.; et al. Is the aspect ratio of cells important in deep learning? A robust comparison of deep learning methods for multi-scale cytopathology cell image classification: From convolutional neural networks to visual transformers. *Comput. Biol. Med.* 2022, 141.
30. Mahmoud, H.A.H.; Alarfaj, A.A.; Hafez, A.M. A Fast Hybrid Classification Algorithm with Feature Reduction for Medical Images. *Appl. Bionics Biomech.* 2022, 2022, 1367366.
31. Zhao, S.; He, Y.; Qin, J.; Wang, Z.; Urzua, U. A Semi-supervised Deep Learning Method for Cervical Cell Classification. *Anal. Cell. Pathol.* 2022, 2022, 4376178.

Retrieved from <https://www.encyclopedia.pub/entry/history/show/84890>