Mansoura Engineering Journal

Volume 48 | Issue 3 Article 3

2023

Deep Learning-Based Technique for The Perception of The Cervical Cancer

Aya Haraz
Biomedical Engineering Program, Faculty of Engineering, Mansoura University, Egypt

Hossam El-Din Moustafa
Department of Electronics and Communications Department, Faculty of Engineering, Mansoura University, Egypt

Abeer Twakol Khaleel
Department of Electronics and Communications Department, Faculty of Engineering, Mansoura University, Egypt

Ahmed H. Eltanboly
Department of Mathematics and Engineering Physics, Faculty of Engineering, Mansoura University, Egypt, eng_ahmed_hazem@mans.edu.eg

ORIGINAL RESEARCH

Deep Learning-based Technique for the Perception of the Cervical Cancer

Aya Haraz [a], Hossam ED. Moustafa [b], Abeer T. Khaleel [b], Ahmed H. Eltanboly [c,d,*]

[a] Biomedical Engineering Program, Faculty of Engineering, Mansoura University, Mansoura, Egypt
[b] Department of Electronics and Communications Department, Faculty of Engineering, Mansoura University, Mansoura, Egypt
[c] Department of Mathematics and Engineering Physics, Faculty of Engineering, Mansoura University, Mansoura, Egypt
[d] Department of Mathematics, Faculty of Basic Sciences, Galala University, New Galala City, Egypt

Abstract

In third-world countries, cervical cancer is the most prevalent and leading cause of death. It is affected by a variety of factors, including smoking, poor nutritional status, immunological inadequacy, and prolonged use of contraception. The Pap smear test, which is intended to prevent cervical cancer, finds preneoplastic changes in cervical epithelial cells. This study framework classified cervical cancer cells from Pap smears into five specified cell types using machine learning-based classification algorithms. The SIPaKMeD database is used in this investigation. This public dataset, which was manually cropped from 966 cluster cell images taken from Pap smear slides, has 4045 isolated Pap smear cells. Depending on their cellular form and structure, experts categorize the cells into five distinct types which are superficial-intermediate, Parabasal, Koilocytotic, Dyskeratotic, and Metaplastic cells. Introducing a pipeline to improve algorithm accuracy and easy implementation by using a specified feature extractor and conducting a suitable preprocessing pipeline is the main contribution. The study reached the conclusion that machine learning could improve Pap smear screening classification results. We concluded that the support vector machine (SVM) is the most suitable algorithm for this application. The SVM has the highest accuracy of 0.968, the Neural Network (NN) at 0.958, and the (K-Nearest Neighbor) KNN at 0.941. These results support the proposed framework as a reliable classification diagnostic tool.

Keywords: Deep features, Feature engineering, Principal component analysis (PCA), Neural network, Pap smear (PS)

1. Introduction

CANCER is the body's abnormal cells growing uncontrollably. Uncontrolled growth and division of cancer cells allow them to invade healthy tissues and organs and subsequently spread throughout the body (Win et al., 2020). One condition that places women's health at hazard and leads to infertility is cervical cancer, even though it is challenging to identify any early warning symptoms. It is the leading cause of cancer mortality in 42 countries, according to the World Health Organization's (WHO) most recent estimates, and it is the fourth most prevalent cancer among women worldwide (Rezende et al., 2021; Sung et al., 2021). Furthermore, Pap smear screening is disapproved of socially, mainly in low-income areas. However, automated Pap smear cell detection and identification at an early stage of cell growth is vital for accurate disease diagnosis and early detection (Khamparia et al., 2020). The Pap smear test detects preneoplastic alterations in cervical epithelial cells. Additionally, such early detection could prevent cervical cancer (Diniz et al., 2021). Cervical screening programs (Pap smear) in wealthy countries have lowered death rates due to their widespread adoption and increased public knowledge of the disease. Cervical cancer is one of the worst illnesses, although it may be treated if caught early. Human papillomavirus (HPV) infection is the main risk factor for cervical cancer (Nithya and Ilango, 2019). Over 150 distinct viruses are members of the

Received 26 December 2022; revised 31 March 2023; accepted 7 April 2023.
Available online 26 June 2023

* Corresponding author at. Department of Mathematics and Engineering Physics, Faculty of Engineering, Mansoura University, Mansoura, Egypt.
E-mail address: Eng_Ahmed_Hazem@mans.edu.eg (A.H. Eltanboly).

https://doi.org/10.58491/2735-4202.3043

HPV family of viruses. Some of them cause lesions, or papilloma, which are a type of growth. The blood or internal organs like the heart or lungs are immune to HPV infection; however, it can affect skin surface cells, as well as lines the genitals, anus, mouth, and throat. HPV may spread from one person to another via sexual intercourse (‘Cervical screening, 2022; Asadi et al., 2020). Colposcopy and Pap smear are the two techniques to capture the cells and identify the individual. Colposcopy is not recommended because the Pap smear test is painless, affordable, and delivers a variety of results (Arora et al., 2021). A Pap smear is a test to aid in avoiding the presence of cervical cancer; it is not a test for cancer (‘The Pap Smear, 2022; ‘Cancer Sourcebook for Women, 2022). Furthermore, the Pap smear test is used to reduce infertility. During the Pap test, cells from the vagina and cervix's surface are collected using a piece of cotton and a brush (Singh and Goyal, 2023). Women who are classified as infertile are more likely to have epithelial cell abnormalities than fertile women are. This indicates that the cells lining the vagina or cervix have changes that might be cancer or precancerous cells (AO et al., 2013). A Pap smear test is strongly indicated for women who are being assessed for infertility problems as part of their normal treatment. Doctors in the hospital have trouble recognizing the cancer cell since the nucleus of the cell might occasionally be a little challenging to view with the naked eye. The size of the nucleus in a normal cell is less than that in an aberrant cell. When categorizing cervical cancer, the abnormal nucleus is bigger, and often the size cannot be exactly detected by eyesight alone (Mustafa et al., 2020). This inspired us to develop a powerful methodology for accurately identifying the Pap smear cell types. In our intended investigation, we used Pap smear cervical cancer dataset images that seem to be HPV-infected. It includes 4049 isolated cropped cell images separated into five classes. Normal cells are separated into two main types (superficial-intermediate, and parabasal), abnormal but not malignant cells are split into two main categories (Koilocytes and Dyskeratotic), and benign cells have one category (metaplastic cells). Each image in the dataset describes the cytoplasm and nucleus of the cells. There are 2048 abstract features calculated automatically from each region of interest in each image, describing the embedded characteristics of images (more illustration in the methodology section). The main part of implementing the proposed system is that can process RGB images in the transformation of painting characteristics into numerical descriptors. In solving this task, automated painter recognition focuses on the measurable elements in a painting that are represented with a set of global image features. The set of computed image descriptors includes statistical features that describe the intensity of a grayscale image, features based on color, textural features, etc. and it is calculated automatically and obtained using deep learning. Cell size, cell region, nucleus area, cytoplasm region, and cell shape are the most significant general features that were derived from an infected Pap smear. The features of cell shape like rotating the cytoplasm and the nucleus, cell topology either cell distribution and position, nucleus, cytoplasm color densities, and last cell texture of mononuclear or multinucleated cells (Ware et al., 2018). Finally, we present evaluation results based on feature and image-based classification systems, together with metrics for the discriminative power of each classifier. To identify the five different types of cells, we primarily concentrate on the proposed work that provides improved classification accuracy.

1.1. Type A: Superficial intermediate cells

This kind of cell covers most of the population and is considered normal. It has a flat circular, oval, or polygonal form. The cytoplasm is colored pink, and the nucleus is prominent in size and form with thickening and a rise in the condensation of the pigmenting material as shown in Fig. 1a.

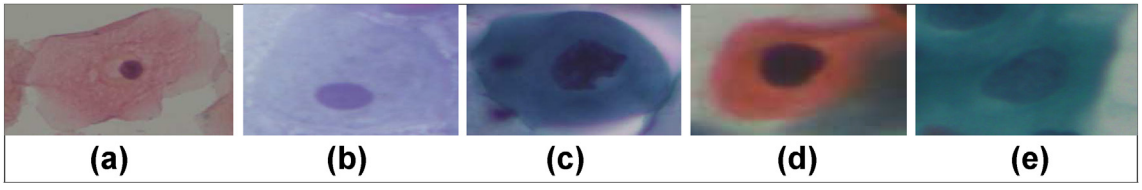

Fig. 1. Optical microscope Pap smear images for different types of cells: (a) Superficial cells, (b) Parabasal cells, (c) Koilocytotic cells, (d) Dyskeratotic cells, and (e) Metaplastic cells.

1.2. Type B: Parabasal cell

Occasionally regarded as normal cells since cancer cells are immature and the smallest epithelial cells have cyanophilic cytoplasm and big vascular nuclei as shown in Fig. 1b.

1.3. Type C: Koilocytotic cells

Consist of abnormal cells with mature cancer cells that are bluish and weakly stained, as well as cells that have strongly stained cytoplasm and a large cavity around the nucleus. Nuclear membrane irregularity and diploid or multinucleated cells are present as shown in Fig. 1c.

1.4. Type D: Dyskeratotic cells

It is also regarded as an abnormal cell since the cytoplasm is bright and the nucleus is semi-koilocytotic, as shown in Fig. 1d.

1.5. Type E: Metaplastic cells

These cells might be small or huge, but their edges are clearly visible. Within each cell is a large vacuole. The cytoplasm will be deeper in color and the central portion will be light brown. The cytoplasm has a round shape as shown in Fig. 1. E is regarded as benign cells, the area of transformation from benign to cancerous cells.

1.6. Related work

We looked at the latest approaches that focus on the early diagnosis and prediction of cervical cancer and make a comparison between them from datasets, the classification algorithm, the results, and the publication info perspectives as shown in Table 1. From this comparison, we concluded that there are no existing methods used such as our proposed method which is differentiated in easy implementation and provide high accuracy.

2. Material and method

Using Pap smear images which are publicly available and contain 4049 isolated cell images, a framework for the early diagnosis of cervical cancer was developed. As shown in Fig. 2. We built the Computer Aided Diagnosis (CAD) system using data mining software to complete six essential stages. Importing the images is the focus of the first phase. The second stage begins with integrating images into the embedded painting feature extractor. Focusing on removing sparse features from the image is the third stage. The fourth stage is where the preprocessing pipeline starts to improve the outcomes, containing three sequential steps.

Table 1. A comparison of the relevant research from the perspectives of datasets, classification methods, outcomes, and publishing information.

Paper	Dataset	Classifier	Results	Year
Machine Learning Approach for Prediction of Cervical Cancer (Priyanka, 2021)	Images are gathered from the 'Herlev University Hospital,' which has 20 unique features and 7 classes.	Convolution Neural Networks (CNN)	Classification Accuracy (CA) 74.04%	2021
SIPaKMeD: A New Dataset for Feature and Image-Based Classification of Normal and Pathological Cervical Cells in Pap Smear Images (Plissiti et al., 2018)	SIPaKMeD's image dataset includes 4049 cell images.	SVM, multi-layer perceptron (MLP), and CNN	MLP for Nuclei features 78.81 ± 1.83 CA SVM for Nuclei features 83.45 ± 1.53 CA MLP for Cytoplasm features 88.54 ± 5.60 CA SVM for Cytoplasm features 91.68 ± 0.98 CA CNN for (RGB) color features 95.35 ± 0.42 CA SVM for Deep (convolutional) features 93.35 ± 0.62 CASVM for Deep (fully connected) features 94.44 ± 1.21 CA	2018
Classification of Clinical Dataset cervical Cancer using KNN (Sharma et al., 2016)	Isolated cells gathered from Fortis Hospital Mohali, Punjab (India)	KNN	82.9% CA	2016
Comparison of Feature selection methods for diagnosis of cervical cancer using the SVM classifier (Ashok and Aruna, 2016)	Rajah Muthiah Medical College collected 150 images of Pap smear testing.	SVM	CA of 98.5%, the sensitivity of 98%, and specificity of 97.5%	2016

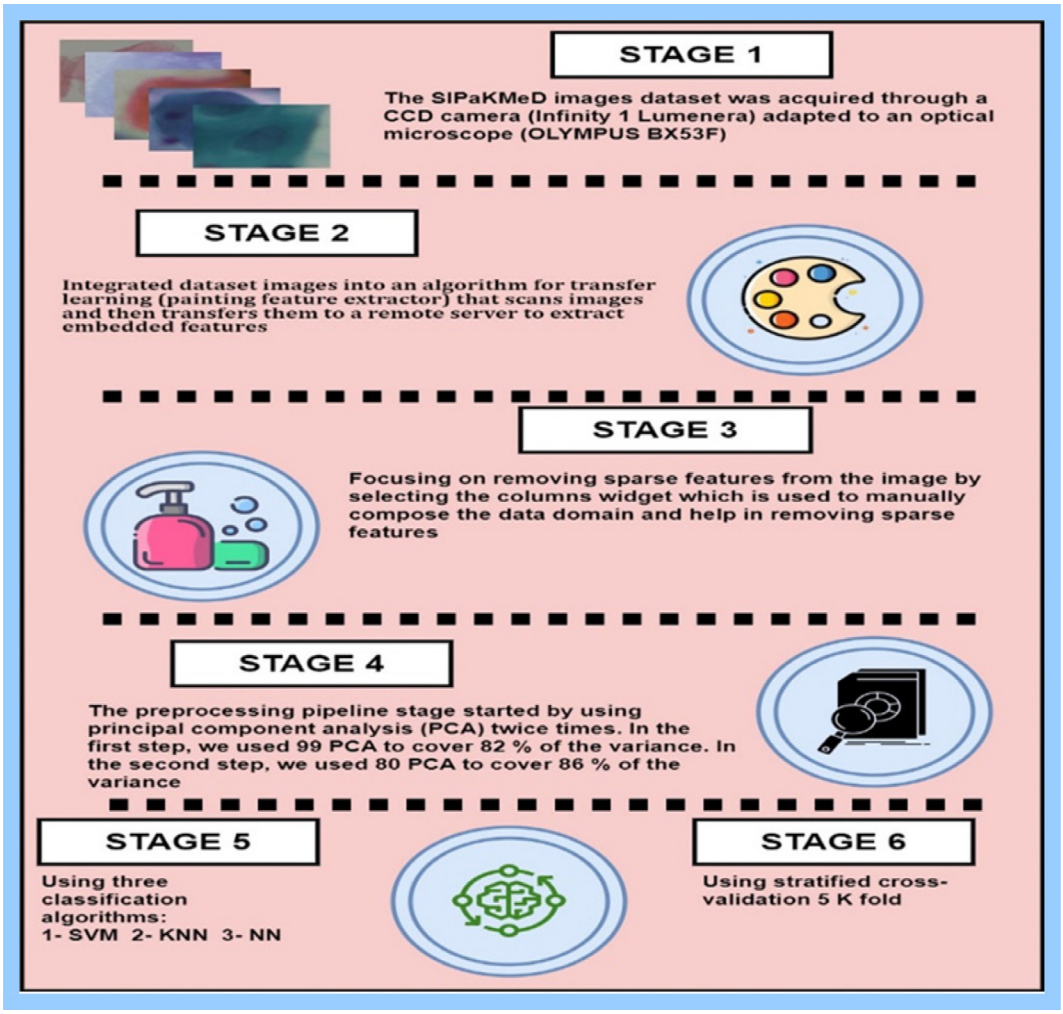

Fig. 2. The proposed method Flow chart demonstrates each step in Computer Aided Diagnosis (CAD) system Procedure.

Three distinct algorithms were utilized in the fifth stage to aid in training and learning. The outcomes of the diagnostic were assessed in the last phase.

In the first stage, we first prepared all five types, which contain about 800 images for each categorical cervix class, they have been configured and embedded into the software. In the second stage, we integrated these images into an algorithm for transfer learning that scan images and then transfers them to a remote server. Each image is given a feature vector using a deep learning method (Kalantidis et al., 2016; Babenko and Lempitsky, 2015). An enhanced data table with extra columns (image descriptors) was made available. We used the painting feature extractor, which is trained to predict painters from artwork images (Pap smear images) that read and upload them to a remote server to evaluate them locally (‘Orange Data Mining, 2022; Hariharan et al., 2015; Sfikas et al., 2016). The painting feature extractor is Google's deep neural network for image identification trained on 79,433 artworks by 1584 different painters. It uses the penultimate layer as an activation function for the network embedder, essentially encodes images as vectors, and produces 2048 abstract features from each image. The third stage starts by selecting the columns widget, which is used to manually compose the data domain and help in removing sparse features and metadata produced from the feature extraction stage. The fourth stage is the preprocessing pipeline stage which contains two essential steps started using principal component analysis (PCA), which transforms correlated data into a collection of linearly uncorrelated features via an orthogonal transformation by compressing the selected features to 99 principal components with standard deviation normalization which covers 82% of the variance in features as shown in Fig. 3 likewise, we repeated this step by taking 80 principal components with standard deviation normalization

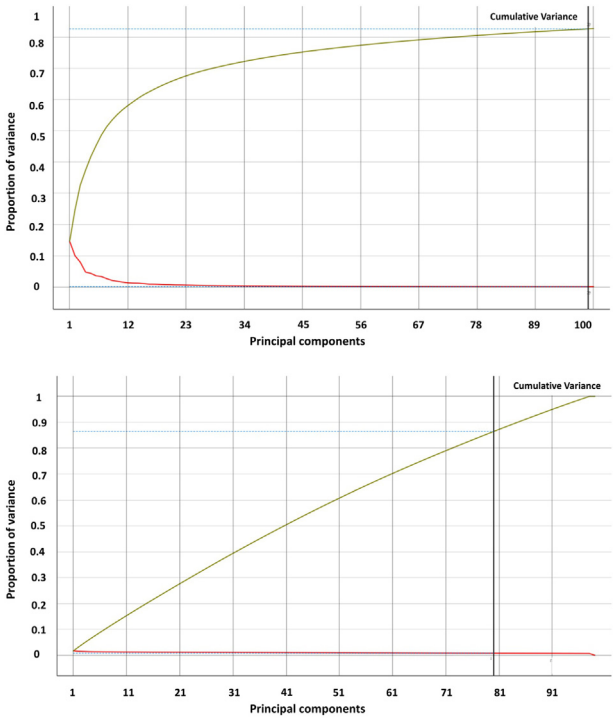

Fig. 3. The sequence steps of principle component analysis (PCA) vs. the proportion of variance.

to cover 86% of the variance. This step is very important to help in shortening the training time for the whole process training time. As an example, the SVM consumes about 41 s of training time. Conversely, by adding the last step in that stage the time is reduced to 27 s.

In the fifth stage, we employed powerful classifiers to begin training the algorithms. First, we used SVM (Ashok and Aruna, 2016). It is a well-established Supervised Learning method for solving classification problems. We used the radial basis function kernel, (RBF) kernel, with a hyperparameter cost (c) = 7, regression loss epsilon = 0.50, and iteration limit = 2000. Furthermore, we utilized the K-NN (‘ K-Nearest Neighbor(KNN, 2022), which used the Euclidean metric with the number of neighbors = 8. Finally, used a NN. It is a group of algorithms that uses a technique inspired by the way the human brain functions to look for hidden correlations in a set of data. The output criteria didn't need to be revised because the network produced the greatest possible outcome (‘ Neural Network Definition, 2022). We utilized the hyperparameter which consists of the Adam solver, 100 neurons in the hidden layer, ReLu as an activation function, optimized the regularization to reach 0.0001, and the number of iterations = 200. In the final stage, we used stratified cross-validation (K fold) to evaluate the results by taking 5 K folds. We are the first group to propose, such a CAD system, and to the best of our knowledge, we were unable to identify any work in the literature that tries to improve classification accuracy while minimizing training time using an easy implementation method such as our work.

3. Experimental results

The SIPaKMeD images dataset was collected using an optical microscope-compatible CCD camera (Infinity 1 Lumenera) (OLYMPUS BX53F) (Plissiti et al., 2018; ‘Single Cell Conventional Pap Smear, 2016). We employ three machine-learning algorithms (SVM, KNN, and NN) across the 5 K folds to evaluate and validate the suggested methodology. Additionally, we provide a confusion matrix for each of the used classifiers, as shown in Tables 3–5. The confusion matrix of SVM, which is the best technique for this kind of issue, revealed that 796 instances of the Dyskeratotic type were correctly classified, as opposed to 766 and 785 instances for the KNN and NN algorithms. Furthermore, the SVM confusion matrix demonstrated that 769 occurrences of the Koilocytotic type were correctly classified, as opposed to 716 and 756 instances for the KNN and NN algorithms, respectively. Additionally, the SVM confusion matrix demonstrated that 758 instances of the Metaplastic type were correctly classified, as opposed to 738 and 746 instances for the KNN and NN algorithms, respectively. Moreover, the SVM confusion matrix emphasized that 777 instances of the Parabasal type were correctly classified, as opposed to 776 and 777 instances for the KNN and NN methods, respectively. Finally, the SVM confusion matrix revealed that 820 instances of the Superficial-intermediate type were correctly classified, as opposed to 814 instances for the KNN and 815 instances for the NN algorithms, respectively. As shown in Table 2 when we evaluated the performance of the suggested method using the area under the curve (AUC),

Table 2. A comparison between the CA, and the AUC info viewpoints for the proposed method.

Classifier model	CA	AUC	F1	Precision	Recall	Log loss	Specificity
SVM	0.968	0.998	0.968	0.968	0.968	0.103	0.992
Neural Network	0.958	0.997	0.958	0.958	0.958	0.141	0.989
KNN	0.941	0.993	0.941	0.941	0.941	0.340	0.985

Table 3. KNN confusion matrix.

Cell type	Dyskeratotic	Koilocytotic	Metaplastic	Parabasal	Superficial-intermediate
Dyskeratotic	766	34	6	7	0
Koilocytotic	44	716	45	8	12
Metaplastic	16	26	738	7	6
Parabasal	2	2	4	776	3
Superficial-intermediate	0	3	8	6	814

Table 4. Neural Network confusion matrix.

Cell type	Dyskeratotic	Koilocytotic	Metaplastic	Parabasal	Superficial-intermediate
Dyskeratotic	785	23	3	2	0
Koilocytotic	25	756	35	3	6
Metaplastic	7	27	746	8	5
Parabasal	2	0	4	777	4
Superficial-intermediate	0	6	7	3	815

classification accuracy (CA), precision, recall, F1-score, log loss, and Specificity equations (1)–(5), we found that the SVM classifier performed best overall compared to other classifiers. The SVM emphasized that it's the best suitable algorithm regarding to evaluated matrix in comparison with KNN's CA (0.941) and NN's CA (0.958). SVM has a high CA of 0.968. We also use other parameter to show the performance of our system. According to the results, the SVM has F1 score (0.968), recall (0.968), log loss (0.103; the lower the better), specificity (0.992), and AUC of 0.998. Contrarily, NN has an AUC of 0.997, a precision of 0.958, an F1 score of 0.958, recall of 0.958, log loss of 0.141, and specificity of 0.989. Finally, KNN has AUC of 0.993, and F1 score of 0.941, recall of 0.941, log loss of 0.340, and specificity of 0.985. The SVM algorithm significantly outperformed NN and KNN in terms of performance.

$$\textbf{Precision} = \frac{TP}{TP + FP} \tag{1}$$

$$\textbf{Specificity} = \frac{TN}{TN + FP} \tag{2}$$

$$\textbf{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

$$\textbf{Recall} = \frac{TP}{TP + FN} \tag{4}$$

Table 6. A comparison between the existing method and the proposed method classification accuracy (CA) info perspectives.

Proposed Method	CA
SVM	0.968
Neural Network	0.958
KNN	0.941
Existing Method	
MLP for Nuclei features	78.81 ± 1.83
SVM for Nuclei features	83.45 ± 1.53
MLP for Cytoplasm features	88.54 ± 5.60
SVM for Cytoplasm features	91.68 ± 0.98
CNN for (RGB) color features	95.35 ± 0.42
SVM for Deep (convolutional) features	93.35 ± 0.62
SVM for Deep (fully connected) features	94.44 ± 1.21

$$\textbf{F1 Score} = \frac{TP}{[TP + 0.5(FP + FN)]} \tag{5}$$

where FP, FN, TN, and TP denote the counts of false positive, false negative, true negative and true positive, respectively.

$$\log \text{loss} = -\frac{1}{N}\sum_{i=1}^{N} \times \sum_{j=1}^{M} y_{ij} \log\left(p_{ij}\right), y_{ij} = \begin{cases} 1 \; if \; Observation \in Class \; j \\ 0 \; elsewhere \end{cases} \tag{6}$$

Table 5. SVM confusion matrix.

Cell type	Dyskeratotic	Koilocytotic	Metaplastic	Parabasal	Superficial-intermediate
Dyskeratotic	796	15	2	0	0
Koilocytotic	18	769	30	0	8
Metaplastic	3	26	758	4	2
Parabasal	2	1	3	777	4
Superficial-intermediate	0	4	4	3	820

Where **N** denotes the number of rows in the test set, while **M** is the number of fault delivery classes, and p_{ij} is the predicted probability that the observation belongs to class j.

4. Discussion

In this framework, we used a robust preprocessing pipeline and selected the best feature extraction and dimension reduction algorithms that focus on reducing dimensionality and getting the best focusing for important features and help shorten training time while improving data compatibility with a learning model class. Furthermore, we used a strong classifier and optimized hyperparameter to help improve results very well. In opposition to that, the existing method (Plissiti et al., 2018), which we compared with us, utilized the SIPaKMeD dataset to compare the performance of the classification algorithms in recognizing the various cell types. They employed three separate patterns such as cell features, image features, and deep features (convolutional and fully connected layer) scheme—to extract the characteristics offered for each pattern of cells in the SIPaKMeD dataset while using a 5K-fold cross-validation. Furthermore, they utilized MLP, SVM, and CNN to train the model and classify images. Consequently, while comparing the CA of our proposed work with the existing work, we found that utilizing SVM was much more accurate in the proposed method than SVM's CA in the existing method, which provides CA = 0.968 as shown in Table 6.

4.1. Conclusion

In conclusion, image cell clusters that were obtained from Pap smear slides are present in the publicly accessible SIPaKMeD cell image dataset (‘Single Cell Conventional Pap Smear, 2016). We provide a preprocessing pipeline and feature engineering pipeline that handle dataset using specified powerful feature extractor, such as a painting extractor, and used strongly optimized classifiers during the training phase to get results that are more precise. This procedure produced encouraging findings, particularly regarding the fact that early detection of this kind of illness by Pap smear testing is essential in lowering death and infertility rates globally. It was crucial to creating a new schematic approach to improve CAD system results because the Pap smear reveals the presence of dysplastic cells or adenoma cells and classifying these types of cells greatly aids in an early solution to any problem in the female reproductive system, not just for early detection of cervical cancer.

Author contribution/author credit statement

Eng. Aya Haraz: Study the design of the work, Data collection and tools, Data analysis and interpretation, Methodology, Software. Dr. Ahmed Hazem Eltanboly: Supervision, Drafting the article, Critical revision of the article, Investigation, Project administration, Resources, Final approval of the version to be published, Visualization. Prof. Hossam El-Din Moustafa: Funding acquisition, Supervision, Critical revision of the article, Project administration, Resources, Final approval of the version to be published. Assoc. Prof. Abeer Twakol Khaleel: Supervision, Critical revision of the article, Project administration, Final approval of the version to be published.

Conflict of interest

None declared.

References

Ao, Almobarak, Elhoweris, M.H., Nour, H.M., Ahmed, M.A.A.M., Omer, A.F.A., Ahmed, M.H., 2013. Frequency and patterns of abnormal Pap smears in Sudanese women with infertility: what are the perspectives? J Cytol Indian Acad Cytologists 30, 100. https://doi.org/10.4103/0970-9371.112651.

Arora, A., Tripathi, A., Bhan, A., 2021. Classification of cervical cancer detection using machine learning algorithms. In: 2021 6th International Conference on Inventive Computation Technologies (ICICT), pp. 827–835.

Asadi, F., Salehnasab, C., Ajori, L., 2020. Supervised algorithms of machine learning for the prediction of cervical cancer. J Biomed Phys Eng 10, 513.

Ashok, B., Aruna, P., 2016. Comparison of Feature selection methods for diagnosis of cervical cancer using SVM classifier. Int. J. Eng. Res. Afr. 6, 94–99.

Babenko, A., Lempitsky, V., 2015. Aggregating local deep features for image retrieval. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 1269–1277.

Cancer Sourcebook for women: basic consumer health information about leading … – book google. https://books.google.com.eg/books?id=Fds6AQAAIAAJ&q=cancer+source+book+for+women&dq=cancer+source+book+for+women&hl=ar&sa=X&ved=2ahUKEwiP2-_Uv536AhUCixoKHbc-AQ8Q6AF6BAgIEAI. (Accessed 18 September 2022) accessed.

‘Cervical screening - NHS. https://www.nhs.uk/conditions/cervical-screening/. (Accessed 18 September 2022).

Diniz, D.N., Rezende, M.T., Bianchi, A.G.C., Carneiro, C.M., Luz, E.J.S., Moreira, G.J.P., et al., 2021 Jul 9. A deep learning ensemble method to assist cytopathologists in pap test image classification. J. Imaging 7 (7), 111.

Hariharan, B., Arbeláez, P., Girshick, R., Malik, J., 2015. Hypercolumns for object segmentation and fine-grained localization. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 447–456.

Kalantidis, Y., Mellina, C., Osindero, S., 2016. Cross-dimensional Weighting for Aggregated Deep Convolutional Features. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in

Bioinformatics), vol. 9913. https://doi.org/10.1007/978-3-319-46604-0_48. LNCS:685–701.
Khamparia, A., Gupta, D., Vhc de Albuquerque, 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. J. Supercomput. 76, 8590–8608.
'K-Nearest Neighbor(KNN. Algorithm for Machine Learning - Javatpoint. https://www.javatpoint.com/k-nearest-neighbor-algorithm-for-machine-learning. (Accessed 18 September 2022).
Mustafa, W.A., Halim, A., Ma, Jamlos, SZS Idrus, 2020. A Review: pap smear analysis based on image processing approach. J. Phys. Conf. 1529, 022080.
'Neural Network Definition. https://www.investopedia.com/terms/n/neuralnetwork.asp. (Accessed 18 September 2022).
Nithya, B., Ilango, V., 2019. Evaluation of machine learning based optimized feature selection approaches and classification methods for cervical cancer prediction. SN Appl. Sci. 1, 1–16.
Orange data mining - image embedding. https://orangedatamining.com/widget-catalog/image-analytics/imageembedding/. (Accessed 18 September 2022).
Plissiti, M.E., Dimitrakopoulos, P., Sfikas, G., Nikou, C., Krikoni, O., Charchanti, A., 2018. SIPAKMED: a new dataset for feature and image based classification of normal and pathological cervical cells in Pap smear images. In: 2018 25th IEEE International Conference on Image Processing (ICIP), pp. 3144–3148.
Priyanka, B.J., 2021. Machine learning approach for prediction of cervical cancer. Turkish J Comp Mathe Education (TURCOMAT) 12, 3050–3058.
Rezende, M.T., Silva, R., Bernardo, F.D., Tobias, A.H., Oliveira, P.H., Machado, T.M., et al., 2021 Jun 10. Cric searchable image database as a public platform for conventional pap smear cytology data. Sci Data 8 (1), 151.
Sfikas, G., Retsinas, G., Gatos, B., 2016. Zoning aggregated hypercolumns for keyword spotting. In: 2016 15th International Conference on Frontiers in Handwriting Recognition (ICFHR), pp. 283–288.
Sharma, M., Singh, S.K., Agrawal, P., Madaan, V., 2016. Classification of clinical dataset of cervical cancer using KNN. Indian J Sci Technol 9, 1–5.
Singh, S.K., Goyal, A., 2023. Performance analysis of machine learning algorithms for cervical cancer detection. In: Performance analysis of machine learning algorithms for cervical cancer detection. In Research Anthology on Medical Informatics in Breast and Cervical Cancer. IGI Global, pp. 347–370.
Single cell conventional pap smear images | kaggle. https://www.kaggle.com/datasets/mohaliy2016/papsinglecell. (Accessed 18 September 2022) accessed.
Sung, H., et al., 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA A Cancer J. Clin. 71, 209–249. https://doi.org/10.3322/CAAC.21660.
'The pap smear - mathilde E boon, albert J H suurmeijer - book Google. https://books.google.com.eg/books?id=umcl4R3sp2AC&pg=PR15&dq=The+Pap+smear+book&hl=ar&sa=X&ved=2ahUKEwi91e-Gv536AhWPZMAKHYReCTcQ6AF6BAgEEAI#v=onepage&q=The%20Pap%20smear%20book&f=false. (Accessed 18 September 2022).
William, W., Ware, A., Basaza-Ejiri, A.H., Obungoloch, J., 2018. A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images. Comput. Methods Progr. Biomed. 164, 15–22.
Win, K.P., Kitjaidure, Y., Hamamoto, K., Myo Aung, T., 2020. Computer-assisted screening for cervical cancer using digital image processing of pap smear images. Appl. Sci. 10, 1800.