International Journal of Multidisciplinary Engineering in Current Research - IJMEC

Volume 10, Issue 5, May-2025, http://ijmec.com/, ISSN: 2456-4265

Cervical Cancer Prediction Using Image Dataset

Pericharla Rajeswari Reethika
PG scholar, Department of MCA, CDNR collage, Bhimavaram, Andhra Pradesh.
K.Suparna
(Assistant Professor), Master of Computer Applications, DNR collage, Bhimavaram, Andhra Pradesh.

Abstract:

Cervical cancer is a leading cancer in the female population. This disease is considered dangerous as its slow and unpredicted growth. The prevention of such cancer can be mostly achieved by screening its transformation zones. The cervical pre-cancerous zones can be considered as three types: type 1, type 2, and type 3. Screening and analyzing these three stages can be crucial for preventing their transformation into cancer. Hence, it is essentially important to have an automated and intelligent system that can grade the cervical pre-cancerous colposcopy images into one of the three types. This can help in providing the right treatment and prevent cancer transformation. In this paper, we develop a residual learning-based model (ResNet-50) to be trained for classifying the type of a colposcopy cervical image into type 1, type 2, and type 3. Experimentally, the model was fine-tuned and evaluated on a public dataset of colposcopy cervical images and achieved promising results in cervical cancer screening of accuracy of 77% and F1-score of 79%.

Keywords: Cervical cancer, colposcopy, ResNet-50, Residual learning.

INTRODUCTION

1.1 Aim of study

The aim of this study is to develop a highly accurate and reliable deep learning-based model for predicting cervical cancer from medical images, utilizing the ResNet50 architecture. This model is intended to assist healthcare professionals in the early detection and diagnosis of cervical cancer, ultimately improving patient outcomes through timely intervention and treatment.

1.2 Objective:

- Select a comprehensive and balanced dataset of cervical images, including both normal and abnormal cases.
- Utilize the ResNet50 architecture, pretrained on the ImageNet dataset, for feature extraction and classification.
- Evaluate the model's performance using a variety of metrics, including accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC.

1.3 Introduction

Their study found that ResNet-50 could detect cervical cancer through cervicography images more accurately than these three conventional machine learning models. A deep learning pipeline has been proposed to detect cervix regions and classify cervical tumors (Fig. 8) [24].

Following the breast cancer, cervical cancer is the most common deadly cancer for women in America [1]. For now, such cancer is still considered uncurable in case it reaches its later stages. Moreover, it is found by the statistics of the World Health Organization (WHO) that this type of cancer is the fourth most prevalent cancer globally, with a reporting rate of 5,70,000 new cases in 2018, accounting for 7.5% of all women cancer deaths [1]. Currently, there are about 85% of the women who face death every year due to cervical cancer in the developing care where professional medical care is not well available [2].

This toll of deaths can be reduced or stopped by granting the developing world countries the access to better technology and healthcare systems in this field. This includes the availability of regular screening tests, which can help in providing an effective treatment for the cervical precancer stage. On the other hand, cervical cancer can be hard to be detected or diagnosed in its

earlier stages as it has no clear and visible signs in its early stage [3]. This makes the regular screening and checkup very significant in order to prevent such disease as this earlier identification can increase the 5-years survival rate of this cancer to 66% [4]. The current methods available for screening cervical cancer are mainly the following: HPV testing, PAP smear testing, colposcopy, and biopsy [5].

The most used screening technique for the diagnosis and treatment of cervical cancer is the PAP smear. However, this technique has several drawbacks: it needs big number of microscopic examinations to for the diagnosis of cancer and noncancer patients, in addition, it is time consuming and requires trained professionals. Nevertheless, PAP smear and HPV testing are considered very expensive treatment and they provide low sensitivity in detecting cervical cancer [4].

Thus, the development of computer aid systems that can detect the precancerous stages of cervical cancer are in need. Such systems that rely only on images can be used as tools to assist medical experts in diagnosing the early stages of cervical cancer and whether or there is a probability of malignant cancer transformations. Deep learning has shown itself as a powerful method in analyzing medical images [6]. Starting from classification to segmentation, deep learning methods have made a sharp progress in that field, making identification and detection of diseases easier and sooner.

Residual learning [7] is a deep learning method that uses identity mapping and skip connections to improve the performance of a very deep network designed to perform complex classification tasks such as cervical precancer screening. In this paper, we investigate the capability of a residual learning-based network, denoted ResNet50, in grading the cervical precancer images into 3 stages or types: Type 1, Type 2, and Type 3. For training and testing, we used the popular cervical screening dataset [8] that consists of sufficient number of colposcopy images of the three different types.

Cervical cancer is one of the most prevalent cancers among women worldwide, significantly impacting global public health. Early detection is paramount in reducing mortality rates and improving the prognosis of patients diagnosed with this disease. Traditional diagnostic methods, such as Pap smears and colposcopy, rely heavily on the manual examination of cervical images by trained medical professionals. However, this process can be subjective, time-consuming, and prone to human error, underscoring the necessity for automated and accurate diagnostic tools.

Recent advancements in machine learning and artificial intelligence, particularly in deep learning, have revolutionized the field of medical image analysis. Deep learning algorithms, specifically convolutional neural networks (CNNs), have demonstrated remarkable success in various image classification tasks, offering a promising solution for the automated detection of cervical cancer. These algorithms can learn and extract intricate features from large datasets, enabling them to classify images with high accuracy.

This study focuses on utilizing the ResNet50 algorithm, a well-established deep residual network, to predict cervical cancer from medical images. ResNet50 is known for its deep architecture and the introduction of residual connections, which help mitigate the vanishing gradient problem and allow for the training of much deeper networks. This capability is crucial for capturing complex patterns in medical images, leading to more reliable and accurate predictions.

LITERATURE SURVEY

1. **M. Abdulsamad, E.A. Alshareef and F. Ebrahim, "Cervical Cancer Screening Using Residual Learning", pp. 2022-12, 2022.**

Cervical cancer is a leading cancer in the female population. This disease is considered dangerous as its slow and unpredicted growth. The prevention of such cancer can be mostly achieved by screening its transformation zones. The cervical pre-cancerous zones can be considered as three types: type 1, type 2, and type 3. Screening and analyzing these three stages can be crucial for preventing

their transformation into cancer. Hence, it is essentially important to have an automated and intelligent system that can grade the cervical pre-cancerous colposcopy images into one of the three types. This can help in providing the right treatment and prevent cancer transformation. In this paper, we develop a residual learning-based model (ResNet-50) to be trained for classifying the type of a colposcopy cervical image into type 1, type 2, and type 3. Experimentally, the model was fine-tuned and evaluated on a public dataset of colposcopy cervical images and achieved promising results in cervical cancer screening of accuracy of 77% and F1-score of 79%.

2. **S. Adhikary, S. Seth, S. Das, T.K. Naskar, A. Barui and S.P. Maity, "Feature assisted cervical cancer screening through dic cell images",** ***Biocybernetics and Biomedical Engineering*****, vol. 41, no. 3, pp. 1162-1181, 2021.**

The mortality rate of cervical cancer is increasing alarmingly. Conventional cytological methods are not always efficient to diagnose cancer at an early stage. Several label-free, quantitative screening approaches are emerging rapidly for fast and accurate detection of cervical cancer. Differential interference contrast (DIC) imaging is one of such label-free methods for the detection of cellular abnormality. The combination of DIC imaging and prediction algorithm enables the development of an efficient computer-aided diagnosis (CAD) system for cervical cancer detection at an early stage. In the present study, the DIC dataset is categorized into 2-classes (abnormal and normal) and 3-classes (normal, precancer, and squamous cell carcinoma). After segmentation of the cells using the modified valley-based Otsu thresholding method, three classifiers, namely support vector machine (SVM), multilayer perceptron (MLP), and k-nearest neighbour (k-NN) are applied. Further, to improve the classification performances, principal component analysis (PCA) is applied for feature selection. The experimental results reveal that the SVM classifier has the greatest accuracy of 0.97 (2-class classification) and 0.90 (3-class classification)

3. **N. Al Mudawi and A. Alazeb, "A model for predicting cervical cancer using machine learning algorithms",** ***Sensors*****, vol. 22, no. 11, pp. 4132, 2022**

A growing number of individuals and organizations are turning to machine learning (ML) and deep learning (DL) to analyze massive amounts of data and produce actionable insights. Predicting the early stages of serious illnesses using ML-based schemes, including cancer, kidney failure, and heart attacks, is becoming increasingly common in medical practice. Cervical cancer is one of the most frequent diseases among women, and early diagnosis could be a possible solution for preventing this cancer. Thus, this study presents an astute way to predict cervical cancer with ML algorithms. Research dataset, data pre-processing, predictive model selection (PMS), and pseudo-code are the four phases of the proposed research technique. The PMS section reports experiments with a range of classic machine learning methods, including decision tree (DT), logistic regression (LR), support vector machine (SVM), K-nearest neighbors algorithm (KNN), adaptive boosting, gradient boosting, random forest, and XGBoost. In terms of cervical cancer prediction, the highest classification score of 100% is achieved with random forest (RF), decision tree (DT), adaptive boosting, and gradient boosting algorithms. In contrast, 99% accuracy has been found with SVM. The computational complexity of classic machine learning techniques is computed to assess the efficacy of the models. In addition, 132 Saudi Arabian volunteers were polled as part of this study to learn their thoughts about computer-assisted cervical cancer prediction, to focus attention on the human papillomavirus (HPV).

4. **S. Fekri-Ershad and S. Ramakrishnan, "Cervical cancer diagnosis based on modified uniform local ternary patterns and feed forward multilayer network optimized by genetic algorithm",** ***Computers in Biology and Medicine*****, vol. 144, pp. 105392, 2022.**

Cervical cancer is one of the most common types of cancer for women. Early and accurate diagnosis can save the patient's life. Pap smear testing is nowadays commonly used to diagnose cervical cancer. The type, structure and size of the cervical cells in pap smears images are major factors which are used by specialist doctors to diagnosis abnormality. Various image processing-based approaches have been proposed to acquire pap smear images and diagnose cervical cancer in pap

smears images. Accuracy is usually the primary objective in evaluating the performance of these systems. In this paper, a two-stage method for pap smear image classification is presented. The aim of the first stage is to extract texture information of the cytoplasm and nucleolus jointly. For this purpose, the pap smear image is first segmented using the appropriate threshold. Then, a texture descriptor is proposed titled modified uniform local ternary patterns (MULTP), to describe the local textural features. Secondly, an optimized multi-layer feed-forward neural network is used to classify the pap smear images. The proposed deep neural network is optimized using genetic algorithm in terms of number of hidden layers and hidden nodes. In this respect, an innovative chromosome representation and cross-over process is proposed to handle these parameters. The performance of the proposed method is evaluated on the Herlev database and compared with many other efficient methods in this scope under the same validation conditions. The results show that the detection accuracy of the proposed method is higher than the compared methods. Insensitivity to image rotation is one of the major advantages of the proposed method. Results show that the proposed method has the capability to be used in online problems because of low run time. The proposed texture descriptor, MULTP is a general operator which can be used in many computer vision problems to describe texture properties of image. Also, the proposed optimization algorithm can be used in deep-networks to improve performance.

5. **W. Liu, C. Li, N. Xu, T. Jiang, M.M. Rahaman, H. Sun, et al., "CVM-Cervix: A hybrid cervical Pap-smear image classification framework using CNN visual transformer and multilayer perceptron", *Pattern Recognition*, vol. 130, pp. 108829, 2022.**

Cervical cancer is the seventh most common cancer among all the cancers worldwide and the fourth most common cancer among women. Cervical cytopathology image classification is an important method to diagnose cervical cancer. Manual screening of cytopathology images is time-consuming and error-prone. The emergence of the automatic computer-aided diagnosis system solves this problem. This paper proposes a framework called CVM-Cervix based on deep learning to perform cervical cell classification tasks. It can analyze pap slides quickly and accurately. CVM-Cervix first proposes a Convolutional Neural Network module and a Visual Transformer module for local and global feature extraction respectively, then a Multilayer Perceptron module is designed to fuse the local and global features for the final classification. Experimental results show the effectiveness and potential of the proposed CVM-Cervix in the field of cervical Pap smear image classification. In addition, according to the practical needs of clinical work, we perform a lightweight post-processing to compress the model.

PROPOSED METHOD

3.1 Proposed Work

1. **Select Dataset:**
   - **Description:** The first step involves selecting a comprehensive and diverse dataset of cervical images. The dataset should include images of both normal and abnormal cervical conditions to ensure the model can learn to distinguish between the two.
   - **Sources:** Publicly available medical image repositories, hospital records, or specialized datasets such as the Herlev dataset or ISBI challenge datasets.
2. **Preprocess:**
   - **Data Cleaning:** Remove any noisy, corrupted, or irrelevant images from the dataset to ensure high-quality input data.
   - **Image Resizing:** Standardize the size of all images to match the input size required by ResNet50 (e.g., 224x224 pixels) for consistency.
   - **Normalization:** Normalize pixel values to a common scale, typically between 0 and 1, to facilitate better training convergence.

- **Data Augmentation:** Apply various augmentation techniques such as rotation, flipping, scaling, and cropping to artificially expand the dataset. This helps in improving the robustness and generalization capability of the model by simulating real-world variations.

3. **Apply ResNet50:**
   - **Model Initialization:** Initialize the ResNet50 model, pretrained on ImageNet, to leverage transfer learning for better feature extraction.
   - **Fine-Tuning:** Replace the final fully connected layer of ResNet50 with a new layer suited to the specific task of cervical cancer prediction (binary classification: normal vs. abnormal).
   - **Training:** Train the model on the preprocessed dataset, utilizing techniques such as dropout and regularization to prevent overfitting. Employ various optimizers (e.g., Adam, SGD) and tune learning rates for optimal performance.
   - **Validation:** Use a separate validation set to monitor the model's performance during training and to adjust hyperparameters accordingly.
4. **Performance Analysis:**
   - **Metrics:** Evaluate the trained model using a range of performance metrics including accuracy, precision, recall, F1-score, confusion matrix, and ROC-AUC.
   - **Cross-Validation:** Perform cross-validation to ensure the model's robustness and generalizability across different subsets of the dataset.
   - **Comparison:** Compare the performance of ResNet50 with other baseline models and previously reported results in the literature to benchmark its effectiveness.
5. **Prediction:**
   - **Inference:** Use the trained ResNet50 model to predict the classification of new, unseen cervical images. The model outputs the probability of each image belonging to the normal or abnormal class.
   - **Post-Processing:** Apply thresholding to the predicted probabilities to make final binary classification decisions.
   - **Deployment:** Develop a user-friendly interface or integrate the model into existing medical diagnostic systems to assist healthcare professionals in real-time cervical cancer prediction.

This proposed method outlines a systematic approach to leveraging the ResNet50 deep learning algorithm for predicting cervical cancer from medical images.

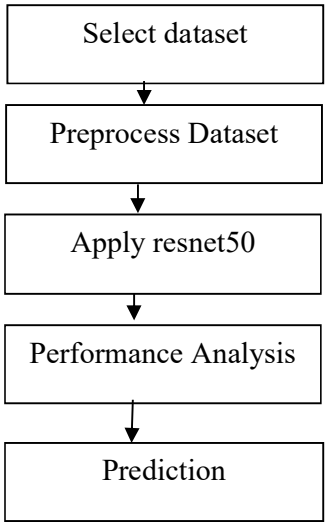
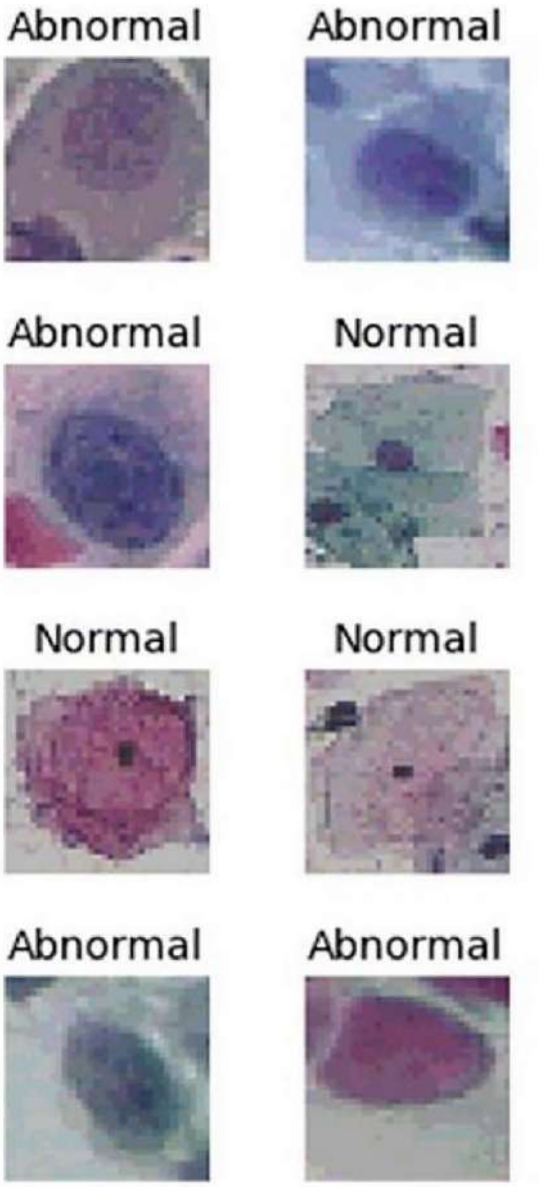


Fig. 3.1 Flowchart of proposed methodology

RESULT



CONCLUSION

In this paper, a transfer learning approach was applied to transfer the knowledge of ResNet-50 model trained on ImageNet, to perform a new classification task which is: cervical pre-cancerous colposcopy image types. The network was trained and tested on a large-scale dataset comprised of 8215 image of the three different types of cervical colposcopy images: type 1, type 2, and type 3. Experimentally, the model has shown a good performance in diagnosing the types of images, despite of the complexity of images as they are all quite similar and contain some artefacts such as blood and others. It is concluded that ResNet-50 is a well-designed and deep architecture of sufficient complexity, could achieve significantly higher classification accuracy when distinguishing between these three types of cervical pre-cancerous colposcopy images which helps in better screening. Furthermore, it is noticed that this model learned features and activation maps of every type, as it was shown in the Grad-Cam visualization which demonstrates that mid and high-level features are learned effectively by the model. Overall, it can be stated that such deep model can be used in real-life application and a helping tool for cervical pre-cancer stages screening in addition to their visualization.

REFERENCES

2 M. Abdulsamad, E.A. Alshareef and F. Ebrahim, "Cervical Cancer Screening Using Residual Learning", pp. 2022-12, 2022.

3 S. Adhikary, S. Seth, S. Das, T.K. Naskar, A. Barui and S.P. Maity, "Feature assisted cervical cancer screening through dic cell images", *Biocybernetics and Biomedical Engineering*, vol. 41, no. 3, pp. 1162-1181, 2021.

4 N. Al Mudawi and A. Alazeb, "A model for predicting cervical cancer using machine learning algorithms", *Sensors*, vol. 22, no. 11, pp. 4132, 2022

5 S. Fekri-Ershad and S. Ramakrishnan, "Cervical cancer diagnosis based on modified uniform local ternary patterns and feed forward multilayer network optimized by genetic algorithm", *Computers in Biology and Medicine*, vol. 144, pp. 105392, 2022.

6 W. Liu, C. Li, N. Xu, T. Jiang, M.M. Rahaman, H. Sun, et al., "CVM-Cervix: A hybrid cervical Pap-smear image classification framework using CNN visual transformer and

multilayer perceptron", *Pattern Recognition*, vol. 130, pp. 108829, 2022.

7 B.A. Mohammed, E.M. Senan, Z.G. Al-Mekhlafi, M. Alazmi, A.M. Alayba, A.A. Alanazi, et al., "Hybrid Techniques for Diagnosis with WSIs for Early Detection of Cervical Cancer Based on Fusion Features", *Applied Sciences*, vol. 12, no. 17, pp. 8836, 2022.

8 H. Alquran, M. Alsalatie, W.A. Mustafa, R. Al Abdi and A.R. Ismail, "Cervical Net: A Novel Cervical Cancer Classification Using Feature Fusion", *Bioengineering 2022*, vol. 9, pp. 578, 2022.

9 J. Kim, C.M. Park, S.Y. Kim and A. Cho, "Convolutional neural network-based classification of cervical intraepithelial neoplasias using colposcopic image segmentation for acetowhite epithelium", *Scientific Reports*, vol. 12, no. 1, pp. 17228, 2022.

10 W. Chen, W. Shen, L. Gao and X. Li, "Hybrid Loss-Constrained Lightweight Convolutional Neural Networks for Cervical Cell Classification", *Sensors*, vol. 22, no. 9, pp. 3272, 2022.

11 J. Guo, J. Wang and C. Xia, "An improved classification method for cervical cancer based on ResNet", *2023 IEEE 12th Data Driven Control and Learning Systems Conference (DDCLS)*, pp. 15501555, 2023, May.

12 W. Chen, X. Li, L. Gao and W. Shen, "Improving computer-aided cervical cells classification using transfer learning based snapshot ensemble", *Applied Sciences*, vol. 10, no. 20, pp. 7292, 2020.

13 M.M. Rahaman, C. Li, X. Wu, Y. Yao, Z. Hu, T. Jiang, et al., "A survey for cervical cytopathology image analysis using deep learning", *IEEE Access*, vol. 8, pp. 61687-61710, 2020.

14 Ghoneim, G. Muhammad and M.S. Hossain, "Cervical cancer classification using convolutional neural networks and extreme learning machines", Future Generation Computer Systems, vol. 102, pp. 643-649, 2020. Tripathi, A. Arora and A. Bhan, "Classification of cervical cancer using Deep Learning Algorithm", *2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*, pp. 1210-1218, 2021, May.

15 M.M. Ali, K. Ahmed, F.M. Bui, B.K. Paul, S.M. Ibrahim, J.M. Quinn, et al., "Machine learning-based statistical analysis for early stage detection of cervical cancer", *Computers in biology and medicine*, vol. 139, pp. 104985, 2021.

16 S. Jahan, M.S. Islam, L. Islam, T.Y. Rashme, A.A. Prova, B.K. Paul, et al., "Automated invasive cervical cancer disease detection at early stage through suitable machine learning model", *SN Applied Sciences*, vol. 3, pp. 1-17, 2021.

17 M. Glučina, A. Lorencin, N. Anđelić and I. Lorencin, "Cervical cancer diagnostics using machine learning algorithms and class balancing techniques", *Applied Sciences*, vol. 13, no. 2, pp. 1061, 2023.

18 M. Rahimi, A. Akbari, F. Asadi and H. Emami, "Cervical cancer survival prediction by machine learning algorithms: a systematic", 2023.

19 U.K. Lilhore, M. Poongodi, A. Kaur, S. Simaiya, A.D. Algarni, H. Elmannai, et al., "Hybrid model for detection of cervical cancer using causal analysis and machine learning techniques", *Computational and Mathematical Methods in Medicine 2022*, 2022.

20 E.L. Silva, A.F. Sampaio, L.F. Teixeira and M.J.M. Vasconcelos, "Cervical Cancer Detection and Classification in Cytology Images Using a Hybrid Approach", *Advances in Visual Computing: 16th International Symposium ISVC 2021 Virtual Event October 4 - 6 2021 Proceedings Part II*, pp. 299-312, 2021.

21 V. Chandran, M.G. Sumithra, A. Karthick, T. George, M. Deivakani, B. Elakkiya, et al., "Diagnosis of cervical cancer based on ensemble deep learning network using colposcopy images", *BioMed Research International 2021*, 2021.

22 D. Arifianto and A.S. Agoes, "Cervical cancer image classification using CNN transfer learning", *2nd International Seminar of Science and Applied Technology (ISSAT 2021)*, pp. 145-149, 2021, November.

23 Y. Park, Y.J. Kim, W. Ju, K. Nam, S. Kim and K.G. Kim, "Classification of cervical cancer

using deep learning and machine learning approach", 2021.

24 Y. Park, Y.J. Kim, W. Ju, K. Nam, S. Kim and K.G. Kim, "Classification of cervical cancer using deep learning and learning approach", 2021.