



HAL
open science

A Survey on Cervical Cancer Detection and Classification Using Deep Learning

K. Hemalatha, V. Vetrivelvi

► **To cite this version:**

K. Hemalatha, V. Vetrivelvi. A Survey on Cervical Cancer Detection and Classification Using Deep Learning. 5th International Conference on Computational Intelligence in Data Science (ICCIDS), Mar 2022, Virtual, India. pp.18-29, 10.1007/978-3-031-16364-7_2 . hal-04381282

HAL Id: hal-04381282

<https://inria.hal.science/hal-04381282v1>

Submitted on 9 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

A survey on cervical cancer detection and classification using deep learning

Hemalatha K¹ and V Vetrivelvi²

Department of Computer Science and Engineering(DCSE), College of Engineering
Guindy, Anna University,Chennai-600025, India

hemshema5@gmail.com

Department of Computer Science and Engineering(DCSE), College of Engineering
Guindy, Anna University,Chennai-600025, India

kalvivetri@gmail.com

Abstract. Cervical cancer is one of the main cause of cancer death, impacting 570,000 people globally. Cervical cancer is caused by the Human Papillomavirus (HPV), which causes abnormal cell growth in the cervical region. Periodic HPV testing in woman has helped to minimize the death rate in developed countries. However, due to a shortage of affordable medical facilities, developing countries are still striving to deliver low-cost solutions. The most commonly used screening test for early detection of abnormal cells and cancer is the Pap smear test. This paper explores existing deep learning model and recent research works done using publicly accessible Intel and Mobile-ODT Kaggle dataset for cervix detection and classification and ensures that these automated technologies helps pathologist in providing fast and cost-effective results.

Keywords: Cervical Cancer · Cervix Detection · Pap-Smear · Image Classification · Convolutional neural networks.

1 Introduction

Cervical cancer is the second most prevalent carcinoma in women and one of the most common gynaecological malignancies, although the cure rate is about 100 percent if detected and treated early. Cervical cancer is a carcinoma that affects a woman's cervix. The cervical is the uterus's neck-shaped canal at the bottom. The pap test is a procedure that involves extracting a tissue from the cervix and examining it under a microscope. The pap test, on the other hand, is inaccurate, showing false negative rates ranges from 6% to 55% [1]. The Human papilloma virus test is a DNA test which identifies cervical cancer by relating it to a particular HPV type [1]. This test is usually not recommended because it has a significant false positive rate. Furthermore, such exams are quite expensive. Due to a shortage of affordable medical facilities, developing countries are still striving to deliver low-cost solutions. Cervical cancer accounts for 70% to 90% in several developing nations and are influenced by various factors like, including birth control pills, smoking, and exposure to diethylstilbestrol (DES) are the main

causes of cervical cancer [4]. The endocervix and the ectocervix are two distinct areas of the cervix, with the endocervix closer to the female internal reproductive organ and the ectocervix closest to the vagina. The cervical transformation zone is where cervical cancer generally occurs. The endocervix and ectocervix cells meet in the transformation zone of the cervix [3]. There are three main cervix



Fig. 1. Images of the three different cervix types

types. Type1, Type2 and Type3 (a) Cervix type 1 the zone of transformation it is completely ectocervical and it is completely visible (b) Cervix type 2, which is the second type of cervix. The zone of transformation has a cervix that is endocervical and it is a component completely visible (c) Cervix type 3, the zone of transformation has a cervix that is endocervical it is a component not completely visible and is the most common type of cervix. Because of the type of treatment varies depending on the type of cervix, so it's crucial to know which one you have. Figure 1 shows the three forms of cervix types[3]. Deep learning models are used to identify the cervical types more appropriately.

Deep learning is supposed to assist pathologists in making more precise diagnoses by providing a precise analysis of cancerous lesions, as well as allowing for a faster clinical workflow. In this study, a survey is conducted on how prominent deep learning architectures are being employed in cervical cancer detection and classification. Also, to ensure that these automated technologies assist health-care providers for fast and cost-effective results.

2 Related Work

Cancer is the world's most dangerous disease. Lung cancer, cervical cancer, breast cancer, ovarian cancer, sarcoma cancer, brain cancer, carcinoma cancer, and more types of cancer exist. Among several cancers, cervical cancer is the fourth most frequent in terms of death rate. If correct treatment/diagnosis is not received in a regular basis, it can be fatal and result in death. Many researchers used deep learning algorithms such as GoogleNet [1], CervixNet [2], CapsNet [3], Inception V3, ResNet50 and VGG19 [5], ResNet53 and ResNet 101 [6], AlexNet and SqueezeNet [7], MobileNet and VGG16 [9], RetinaNet [10], Densenet169 [13] and many more to predict cervical cancer.

Table 1. Comparison of Existing Algorithms

Authors	Year	Model	Description
J. Payette et al.[6]	2017	ResNet	To develop an image classifier for different types of cervixes.
O. E. Aina et al.[7]	2019	AlexNet and SqueezeNet	To use a smaller CNN architecture to automate the classification of cervix types and deploy it on a mobile device.
X. Q. Zhang et al.[4]	2019	CapsNet	To develop a deep learning model for classifying images of cervical lesions in order to improve diagnostic efficiency and accuracy.
R. Gorantla et al.[2]	2019	CervixNet	Segmenting the Region of Interest (RoI) and then classify the RoI to improve cervigram contrast.
Z. Alyafeai et al.[1]	2020	GoogleNet	To create a completely automated cervix detection and cervical image classification pipeline using cervigram images.
P. Guo et al.[10]	2020	RetinaNet and Deep SVDD	In a smartphone-acquired cervical image dataset, a unique ensemble deep learning method is presented for identifying cervix images and non-cervix images.
S. Dhawan et al.[5]	2021	InceptionV3, VGG and ResNet50	To create a cervical cancer prediction model that uses deep learning and transfer learning approaches to recognise and classify cervix image.
Ankur Manna et al.[13]	2021	Inception V3, Xception and DenseNet169	Three Convolutional Neural (CNN) architectures were used to create an ensemble-based classification model.

A new fully automated deep learning pipeline for detecting and classifying cervical cancer in the cervical region was presented by Zaid alyafeai et al. [1]. With an accuracy rate of 0.68 on the intersection of union (IoU) measure, the pipeline starts with a detection method that locates the cervix section 1000 times faster than current date-driven methods. Cervical cancers may be classified using two compact convolutional neural network (CNN) versions. Cervical cancers are classified using self-extracted features learned by suggested CNN models with Area under curve(AUC) values of 0.82 and a 20-fold increase in the long it takes to recognise each cervix image, CNN-based classifiers outperformed previously created feature-based classifiers. Finally, cervigram images are used to train and assess the proposed deep learning pipeline. Sanjeev dhawan et al.,[5] outlines an effort to build a cervical cancer prediction model that uses deep learning and transfer learning techniques to recognize and categories cervix images into one of three categories (Type 1/Type 2/Type 3). ConvNet, which will classify the cervix images, is created using the three models mentioned above: InceptionV3, ResNet50, and VGG19. The results of the experiment show that the Inception v3 model outperforms Vgg19 and ResNet50 on the cervical cancer dataset, with an accuracy of 96.1 %.

Qing et al.,[4] proposed a method were images of cervical lesions were diagnosed using a deep-learning algorithm. The research was subdivided into two halves. The lesions in the cervical images were segmented. Finally, the training set accuracy of the model was 99 percent, while the testing set accuracy was 80.1 percent, resulting in better results than earlier classification approaches and enabling for efficient prediction and classification of huge amount of data. Jack payettee et al., [6] experimented with a number of convolutional designs before settling on residual neural networks with batch normalisation and dropout. For each class, the loss was estimated using the multi-class method loss with a logarithmic scale. The dataset was made available by Kaggle which has 1481 training images, 512 test images. The author also used 4633 additional images for training and employed a variety of data sources due to the short size of the dataset strategies for augmentation.

Rohan et al., [2] suggested an innovative CervixNet methodology that involves image enhancement on cervigrams, segmentation was done to identify the Region of Interest (RoI) for selecting the appropriate treatment. A novel Hierarchical Convolutional Mixture of Experts (HCME) technique was presented for the classification problem. In the field of biomedical imaging, smaller datasets represent an inherent issue, HCME was capable of solving the problem of over fitting. The work was carried out on publicly accessible Intel Mobile-ODT Kaggle datasets, the suggested methodology outperformed all current methodologies, with an accuracy of 96.77 percent. The goal of Liming et al. [12] was to develop a deep learning-based visual evaluation algorithm which could automatically diagnose cervical cancer and precancer. Multiple cervical screening approaches and histopathologic confirmation of precancers were done on a population-based longitudinal cohort of 9406 women ages 18–94 years in Guanacaste, Costa Rica for 7 years (1993–2000). Tumors that was present even for up to 18 years were

discovered. Using computerized cervical images for screening recorded with a fixed-focus camera, the deep learning-based technique was trained and validated. The image prediction score (0–1) could be categorized to balance sensitivity and specificity for precancer or cancer detection. On a two-sided basis, all statistical tests were performed. The findings support the idea of using modern digital cameras to do automated visual evaluations of cervical images.

Summary of the existing literature work is shown in Table 1. The author’s name appears in column 1. The year is listed in column 2, while the author’s model is listed in column 3. Column 4 displays the detailed description. The following section provides the materials and methods used for cervix detection and classification.

3 Materials and Methods

The research’s purpose is to use deep learning technology to identify and classify objects in cervical images. Cervical cells are difficult to classify since the cytoplasm and nucleus are so distinct. In recent years, modern artificial intelligence techniques such as machine learning and deep learning have become increasingly popular in the medical health industry, and have a remarkable success with it. Further in this section we discuss the dataset used and about the object detection and classification techniques in previous research.

3.1 Data Collection

The dataset used for this research was provided by Kaggle as part of the competition [3]. Table 2 provides a breakdown of the dataset. Three types of cervix are reported in this data set are all classified normal (not malignant). It is difficult for healthcare providers to identify the transformation zones, Although the transformation zones are not always evident, some patient needs further testing while others do not. As a result, using an algorithm to make a choice will considerably increase the efficiency and effectiveness of pap smear tests.

Table 2. Intel Mobile ODT Kaggle Dataset [3]

Dataset	Type 1	Type 2	Type 3
Train	250	781	450
Additional	1189	1558	1886
Test	512 (unlabeled)		

3.2 Object Detection Techniques

CNNs and other deep learning-based techniques have dominated recent advances in object detection [1-11]. Faster R-CNN is a good example of a CNN-based

object detection system that works well. (a) Two-stage object detectors, which includes a region proposal network followed by a classification and bounding box regression network, and (b) single-shot networks, which do everything in single network; examples of this method include YOLO [1], single shot detector, and RetinaNet [10]. Regardless, for feature extraction, each object recognition framework can use a number of base network designs, such as VGG [5], ResNet [5,6,11], GoogleNet [1], CervixNet [2], and CapsNet [4]. More training data is necessary as the number of parameters in architecture increases. Deep learning algorithms are frequently combined with transfer learning [5], data augmentation, and various optimization procedures to improve accuracy in specific applications.

3.3 Image Classification Techniques

In the field of computer vision and image processing, image classification has been a subject of great interest. The Convolutional neural network (ConvNet) is the frequently used deep learning network for classification of image [1-11]. The ConvNet has three layers, similar to a neural network: input layer, hidden layer, and output layer. It is a deeper neural network because it has hundreds of hidden layers. Raw image pixel values are given into the input layer, while neurons are given into the output layer according to the number of output classes. A cervigram image is taken as the input in the cervix classification problem, and the output is the possibility that the image belongs to Type I, II, or III. The final layer of the convolution layer is a fully connected layer with the SoftMax activation layer. Because the number of images in the dataset is insufficient to yield significant results, several models, including VGG [5], ResNet [5,6,11], GoogleNet [1], CervixNet [2], and CapsNet [4], rely on weights from pre-trained models to produce notable results. The deep learning models utilised for cervix detection and classification are described in detail in the next section.

4 Models Used

4.1 ResNet

The Residual Network architecture, developed by Microsoft Research in 2015. This architecture introduces the Residual Network concept to address the vanishing/exploding gradient issue.

In this network, it uses a technique known as "skip connections", which allows you to skip the first few layers of training and go straight to the final layer. The advantage of introducing skip connections is that regularisation will bypass any layer that reduces architecture performance. As a result, vanishing/exploding gradients are no longer an issue when training very deep neural networks. Residual blocks have three convolutional layers: a 1x1 layer with fewer filters, a 3x3 layer with the same number of filters, and a 1x1 layer with the same number of filters as the original input size to up sample and allow for the shortcut connection [6].

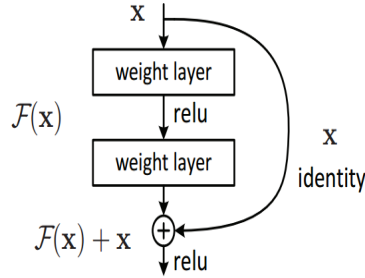


Fig. 2. Residual Learning Building Block[6]

4.2 AlexNet and SqueezeNet

For the purpose of deploying the trained model, SqueezeNet, a compact network with reduced memory computation requirements and equivalent accuracy to AlexNet, was proposed[7]. AlexNet won the ILSVRC-2012 competition with a top-five testing error rate of 15.3 %. The design is made up of five convolutional layers and three fully connected layers. AlexNet is a prominent object-detection network that could have a wide range of applications in the field of computer vision. In the future, AlexNet may be used for image processing more than CNNs. The accuracy of AlexNet and SqueezeNet are nearly same. SqueezeNet, on the other hand, is 500 times smaller and three times faster. Eight Fire modules follow the first convolution layer in the SqueezeNet before the final convolution layer is applied. Instead of using 3 by 3 filters, SqueezeNet uses 1 by 1, downsampling on convolution layers, a fire module, and a smaller number of input channels (3 by 3 filters). This results in smaller activation maps for convolution layers during late down sampling. The low memory and processing requirements make it ideal for embedded systems because it only needs roughly 5MB of storage space. In addition for mobile app on iOS, it has the highest energy efficiency and throughput, while its processing rate on Android is even more faster.

4.3 Inception V3

In the process of image classification Inception v3 model is widely used. Inception model have been proven to be more computationally efficient than VGGNet, both in terms of the number of parameters generated and the cost [5]. The Inception v3 network model uses a convolution kernel splitting method to turn large volume integrals into compact convolutions. For example, a 3 x 3 convolution is split into 3 x 1 and 1 x 3 convolutions. The splitting strategy reduces the number of parameters, allowing the network training process to be accelerated while the spatial feature is retrieved more quickly.

4.4 GoogleNet

The ILSVRC 2014 competition was won by GoogleNet from Google. It had a 6.67 percent top-five error rate. The Inception Network was a key development in the study of Neural Networks, especially CNNs. GoogleNet is the term given to the first iteration of the Inception network. The GoogleNet image classification model has been enhanced to create the Cervix detection module [1]. It comprises of 24 convolutional layers followed by two fully connected layers, as seen in Fig. 3. GoogleNet is faster at training than VGG model.

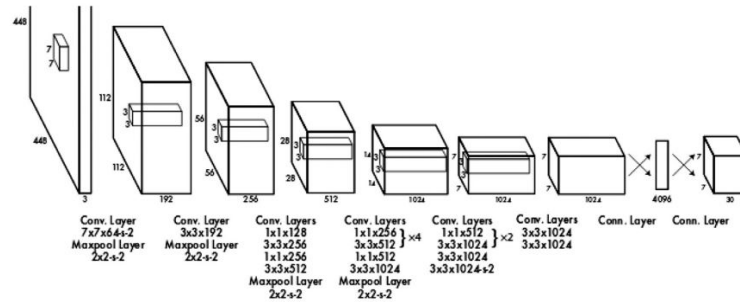


Fig. 3. Architecture of GoogleNet [1]

4.5 CapsNet

The capsule in the CapsNet neural network includes the entity properties and features, including position, deformation, velocity, reflectivity, colour, structure, and other feature space. This spatial information determines the consistency of orientation and size between features. Capsules are designed to record the probability of features and their variants rather than recording features of individual variants. As a result, the capsule's aim isn't just to detect features, but also to train the network to identify variants, because the same capsule can identify the same object class from multiple perspectives.

The CapsNet model [4] was used to classify cervical images. The cervical image training dataset was first fed into the capsNet classification model. After that, the image was put through the Primary Caps layer, which resulted in 16-D capsules. This layer generated 1600 vectors in total, with an output of 10 X 10 and 16 X 16. A 16 dimension was assigned to each vector. The DigiCaps layer was then applied. The output geometry of the DigiCaps layer was 32 X 32, indicating that there were three classes of 32-D capsule units. Then squashing non-linear function was used to transform this vector into the final output vector. Finally, 32 x 32 images was reconstructed using the completely connected layer FC (Fully Connected Layer).

4.6 CervixNet

CervixNet can handle a variety of noisy images. The CervixNet method [2] uses a flat field correction technique, imaging intensity modification methodology, Laplacian filter, and various morphological operations to automatically enhance the input image. The RoI is then retrieved from the improved image using Mask R-CNN segmentation. Cervix types are classified using a two-level hierarchical design based on the Hierarchical Convolutional Mixture of Expert (HCME) algorithm.

4.7 RetinaNet and Deep SVDD

Deep SVDD and RetinaNet, has been found to achieve state-of-the-art multi-class object detection performance on public datasets like the Intel & Mobile ODT kaggle dataset and COCO. Also implementing Deep SVDD, a method that involves training a neural network to extract and encode common features from training data. Deep SVDD [10] is a deep learning architecture-based one-class classification algorithm. The deep architecture's purpose is to extract features from input images and determine common aspects that may be used to represent the goal category. The architecture has three convolutional modules with $32 \times 5 \times 5$ filters, $64 \times 5 \times 5$ filters, and $128 \times 5 \times 5$ filters, as well as a 256-unit dense layer. RetinaNet [10] is said to perform better when dealing with class imbalance and focusing on difficult, misclassified samples. To assist in achieving higher level of training convergence.

5 Discussion

Existing research uses deep learning models such as ResNet, AlexNet, SqueezeNet, CapsNet, CervixNet, GoogleNet, RetinaNet, Inception V3 and others to predict cervical cancer. The applied models are compared on the advantages, disadvantages and various evaluation metrics as shown in table 3. The performance of the detection and classification modules are compared to that of existing state-of-the-art models.

Ensemble of three models InceptionV3, DenseNet 169 and Xception provide complementary information, than the individual performance of each model and overfitting is more likely to occur in dense networks as it has high number of connections, which reduces computational efficiency and parameter efficiency thus the ensemble models achieves an accuracy of 96.36%

AlexNet and ResNet have around 60M parameters in common, however their top-5 accuracy differs upto 10%. However, training a Residual network necessitates a large number of computations (about ten times that of AlexNet), requiring more train time and energy. The vanishing gradient problem is a difficulty that ResNet has to deal with it. AlexNet use dropout layers to solve the over-fitting problem. In a comparison of AlexNet with ResNet on a cervix

Table 3. Result Analysis of Applied Techniques

Model	Advantages	Disadvantages	Results
ResNet	In classification of images, the ResNet model outperformed humans and training a deeper network normally takes more time.	Training a deeper network normally takes more time.	Accuracy: 58%
AlexNet and SqueezeNet	It just requires a small amount of computing RAM and can be used on mobile devices.	The AlexNet model took around an hour to train, whereas the squeezeNet took about 22 Minutes.	AlexNet Accuracy: 62.6% SqueezeNet Accuracy: 63.3%
CapsNet	Detects specific cervical type more accurately based on images.	Capsule Network performs poorly on complex data and exhibits many of the same issues as CNNs initially.	Training Accuracy: 99%, Test Accuracy: 80.1%
CervixNet	Robust model to diverse noisy images and imaging acquisition conditions.	Not suitable for all types of image classification	Accuracy:96.77%, Precision:96.69%, Specificity:98.36%, Sensitivity:96.82%, F1 Score: 0.97
GoogleNet	The GoogleNet model has less parameters as well as more robust.	Google Net is theoretically limitless.	Accuracy: 82%
RetinaNet and Deep SVDD	On publicly available Datasets RetinaNet has been reported to reach state-of-the-art performance for multi-class Object detection.	The duration of RetinaNet training varies between 10 and 35 hours.	Accuracy: 91.6%
InceptionV3, VGG and ResNet50	In comparison to its peers, Inception v3 has the lowest error rate.	Inception V3 is made up of a series of complicated inception modules arranged in a hierarchy.	Accuracy: 96.1%
InceptionV3 and Xception, DenseNet	Ensemble of three models provide complementary information, than the individual performance of each model	Overfitting is more likely to occur in dense networks.	Accuracy: 96.36%

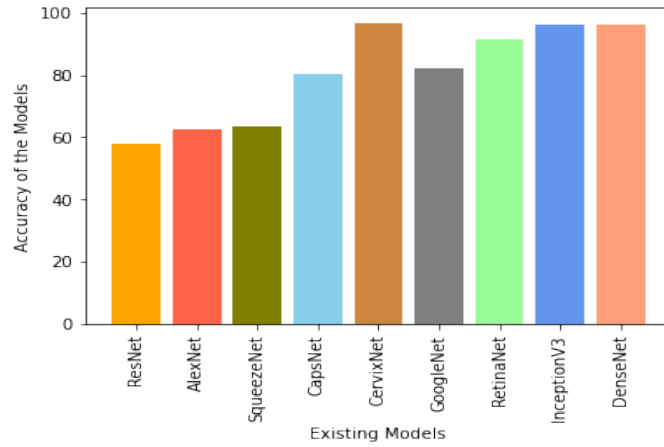


Fig. 4. Comparison of the Existing Models

image dataset, AlexNet had a 62.6 % accuracy. When comparing GoogleNet to AlexNet, GoogleNet has about 4 million parameters and identifies in 33 seconds per epoch, but AlexNet takes 54 seconds and is less accurate. The key distinction between GoogleNet is that it combines inception models, whereas the Inception network looks deep in terms of the number of layers and expands. The VGG model was created primarily to lower the overall number of parameters with fewer trainable parameters, learning is faster and over-fitting is less likely to occur. When compared to all previous models, DenseNet has roughly 12.8 million parameters, eases the vanishing-gradient problem, and considerably reduces the amount of parameters with very much quicker computation time. Figure 4 shows the comparison of existing models with their accuracies.

6 Conclusion

The study presented in this paper shows how various deep learning methods were developed and used to solve the problem of cervix detection and classification. The experiment's algorithm is built on a convolutional neural network approaches, the model's effectiveness by fine-tuning the value of several parameters resulting in progressive improvement. After reviewing of all the applied models, it was discovered that CervixNet, DenseNet and Inception V3 provides 96.77%, 96.36% and 96.1 % better prediction than all other models. The survey shows that a relatively modest design can achieve outcomes comparable to a non-human expert, automated technologies helps pathologist in providing fast and cost-effective results. We will continue to optimise and change the structure of the various network models addressed in this survey as future research to achieve better accuracy.

References

1. Alyafeai, Z. Ghouti, L. A fully-automated deep learning pipeline for cervical cancer classification. *Expert Syst. Appl.* 141, (2020).
2. Gorantla, R., Singh, R. K., Pandey, R. Jain, M. Cervical Cancer Diagnosis using CervixNet - A Deep Learning Approach. *Proc. - 2019 IEEE 19th Int. Conf. Bioinforma. Bioeng. BIBE 2019* 397–404 (2019) doi:10.1109/BIBE.2019.00078.
3. Samperna, R., Boonstra, L., Van Rijthoven, M., Scholten, V. Tran, L. CERVICAL CANCER SCREENING (KAGGLE COMPETITION).
4. Zhang, X. Q. Zhao, S. G. Cervical image classification based on image segmentation preprocessing and a CapsNet network model. *Int. J. Imaging Syst. Technol.* 29, 19–28 (2019).
5. Dhawan, S., Singh, K. Arora, M. Cervix Image Classification for Prognosis of Cervical Cancer using Deep Neural Network with Transfer Learning. *EAI Endorsed Trans. Pervasive Heal. Technol.* 169183 (2018) doi:10.4108/eai.12-4-2021.169183.
6. Payette, J., Rachleff, J. Van De Graaf, C. Intel and MobileODT Cervical Cancer Screening Kaggle Competition: Cervix Type Classification Using Deep Learning and Image Classification. <https://www.kaggle.com/c/>.
7. Aina, O. E., Adeshina, S. A. Aibinu, A. M. Classification of cervix types using Convolution Neural Network (CNN). *2019 15th Int. Conf. Electron. Comput. Comput. ICECCO 2019 2019–2022* (2019) doi:10.1109/ICECCO48375.2019.9043206.
8. Guo, P., Xue, Z., Rodney Long, L. Antani, S. Cross-dataset evaluation of deep learning networks for uterine cervix segmentation. *Diagnostics* 10, (2020).
9. [9] H. Barros. Deep Learning for Cervical Cancer Diagnosis: Multimodal approach.[Online].Available:<https://www.mobileodt.com/products/eva-colpo/>.
10. Guo, P. et al. Ensemble deep learning for cervix image selection toward improving reliability in automated cervical precancer screening. *Diagnostics* 10, (2020).
11. Pfohl, S., Triebe, O. Marafino, B. Guiding the management of cervical cancer with convolutional neural networks.
12. Hu, L. et al. An Observational Study of Deep Learning and Automated Evaluation of Cervical Images for Cancer Screening. *J. Natl. Cancer Inst.* 111, 923–932 (2019).
13. Manna, A., Kundu, R., Kaplun, D., Sinitca, A. Sarkar, R. A fuzzy rank-based ensemble of CNN models for classification of cervical cytology. *Sci. Rep.* 11, 1–18 (2021).