Prediction of Cervical Cancer using Deep Learning

MSc Research Project
Data Analytics

Mohit Teotia
Student ID: x20144041

School of Computing
National College of Ireland

Supervisor: Dr. Catherine Mulwa

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Mohit Teotia
Student ID:	x20144041
Programme:	Data Analytics
Year:	2022
Module:	MSc Research Project
Supervisor:	Dr. Catherine Mulwa
Submission Due Date:	15/08/2022
Project Title:	Prediction of Cervical Cancer using Deep Learning
Word Count:	XXX
Page Count:	20

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Mohit Teotia
Date:	18th September 2022

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Prediction of Cervical Cancer using Deep Learning

Mohit Teotia
x20144041

Abstract

Nowadays, in medical fields, identifying and classifying the precancerous and cancerous phases of cervical cancer is a challenging task, which helps to analyse the various cytology slides. Segmenting nuclei and overlapping cells are frequent steps in the processing of cytology pictures. Deep learning is excelling in medical works for the detection and classification which is easing the tedious medical process and is still an open area for research. Significant results have been observed by utilizing the state-of-the-art architecture of deep learning algorithms therefore in this task, three different deep learning approaches are carried out which are Inception V3, Custom Model, and GAN. Here Inception V3 is based on a transfer learning approach while a custom model is made from scratch using convolutional blocks and GAN is a state-of-the-art model. After training these model on the cytology slide data all executed model is tested over the test set and their performance is evaluated on different metrics. After evaluation of the models, the metrics such as Accuracy, Precision, and recall has been determined. After analysing the results, it is observed that the GAN (Generative Adversarial Network) model has surpassed other models and therefore, can be utilized in the classification of Cytology slides into different classes in real-world application.

1 Introduction

The thriving lifestyle in today's time has introduced several human interaction behaviours whether it be physically or biologically. This at times maybe be advantageous for societal development or disadvantageous. In the medical field, over time multiple facets have evolved, developed, discovered and introduced. Modern times have led to introduce various diseases and viruses but it has brought various improvements into the medical field which led to the appropriate development of remedies. Nowadays, there are various disease and viruses which has been brought to light and is spreading rapidly. The utilization of field specialization, research, and development in collaboration with technological advancement can conveniently support the development of appropriate remedies against diseases and viruses. Researchers have been intensely developing various remedies and processes to hinder the growth of cancer. One such type of cancer is Cervical Cancer. This cancer is fatal and the fourth most frequently developed cancer in the world. Cervical Cancer is generally seen among women and mostly in developing countries (lower or middle-income groups). According to World Health Organization (WHO), about 600 thousand cases were estimated to be developed and fatalities rate more than 50 percent Brüggmann et al. (2022). Although this cancer is considered to be easily treatable only if detected earlier or in the pre-cancerous stage. Cervical Cancer is caused due to the

HPV (Human Papillomavirus) which positions and develops the pre-cancerous cell in the lining of the cervix which is the lower part of the uterus Zhang et al. (2020). There are three types of cervical cancer based on severity levels. HPV virus is generally transmitted to the cervix only through sexual contact. Furthermore, the researchers have stated that the presence of HIV may increase the chance of cervical cancer development in the body. Therefore to treat and eliminate the cancer, the early detection of the pre-cancerous cell which later transforms and replicate into a cancerous cell.

Detecting these cells with the traditional approach was a tedious task which led to the introduction of computer-aided detection. Although these computer-aided detection's were limited to human interventions. The pre-cancerous cell which is known as Papanicolaou is tested through manual screening. The researchers evaluate the chances of it being a Papanicolaou cell through the colour and the shape. The attained samples of the test are viewed in the microscopic environment. If detected, these are eliminated through appropriate treatments and approaches. It is noted that the pap test (Papanicolaou Test) can help reduce the mortality rate by 80%. Although, the approach of manual screening has certain fallacies such as lesser proficiency, time consuming process and also a tedious task. Therefore, an autonomous approach must be incorporated which can increase the proficiency of the detection and make it a less tiring process.

The implementation of machine learning and a deep learning approach for the detection of the cancerous counterpart will be advantageous. The screening process incorporated with the trained machine learning and deep learning model will increase the efficacy of the detection thereby increasing the capability, and reach and also decreasing the cost. The models shall be trained by inputting numerous sample images and setting certain attributes. In this paper, we aim to study and evaluate the implementation of various deep learning algorithms for the detection of cervical cancer possibility. Initially, we sourced the dataset from GitHub repositories which is the Cervix93 Cytology Dataset. These data will be pre-processed and then the features will be extracted from the dataset. It will contain images of different samples which may be infected or non-infected cells. Furthermore, the dataset contains the training set and testing set. Then the algorithms considered in our study are implemented, and the model is trained. After the training, the model is evaluated utilizing certain metrics such as Precision and recall score, accuracy score and validation loss. In the next section, we will be discussing the literature review related to this field of study by various researchers.

The result of this research is project is based on the limited data therefor it cannot be commercialized. It required more data to train so that accuracy of model can be improve then it can be commercialized.

1.1 Research Question

RQ: *"To what extent can we predict the cervical cancer using Deep learning techniques to support medical practitioner for saving a life ?"*
Cervical cancer using deep learning can be detected, in order to validate the predicted outcomes of deep learning model performance metrics such as Accuracy, Loss, Precision, Recall and F1-Score can be calculated.

Sub RQ: *"Which algorithm can correctly identify the Cervical Cancer grades (Negative, LSIL, HSIL)?"*
The algorithm with Highest Accuracy and PRF Score along with minimum loss will be considered as most optimal algorithm for Cervical Cancer Detection.

1.2 Research Objectives & Road map

Obj1: A critical review of literature on Cervical Cancer Detection.
Obj2(a): Apply data pre-processing, image augmentation techniques.
Obj2(b): Implement and evaluate the outcomes of Inception-V3 Model.
Obj2(c): Implement and evaluate the outcomes of Custom model.
Obj2(d): Implement and evaluate the outcomes of GAN architecture.
Obj2(e): Comparison the performance of implemented Models.
Obj2(f): Identify the Most optimal model for cervical cancer detection.

In the first section of the report (Chapter-1), the introduction of problem statement has been provided followed by our research aim and objectives, with research questions. In the next chapter (Chapter-2), we have reviewed the prior on cervical cancer detection problem. In Chapter-3, a framework of methodology has been explained as a solution to detect cervical cancer in an efficient manner. In the next chapter (Chapter-4), the architecture of each algorithm has been discussed. In chapter-5, the tools and technologies required for implementing the project has been discussed. Chapter-6 illustrates the evaluation of results with different experiments and algorithms and lastly, in chapter-7 the conclusion of our analysis is studied and discussed.

2 Literature Review of Cervical Cancer (2014-2022)

2.1 Introduction

In this literature review, a prior work studied by the authors and researchers will be discussed for cervical cancer detection, which can splitted into the following subsections i.e. (i) Critical review of Machine learning based approach for cervical cancer detection, (ii) Critiques of Ensemble Learning algorithm, (iii) Crituqes of Deep learning algorithm, (iv) Identified Gaps and Conclusion.

2.2 Critical Review of Machine Learning based approach for Cervical Cancer Detection

Different image analysis and machine learning approaches for the automated detection of cervical cancer through the Pap smear image test were studied by William et al. (2018). Here different research papers and articles have been studied and evaluated regarding the different algorithms. The general methods include the extraction and detection of the features of pap smear images to detect the chances of cervical cancer. This paper studied about thirty research papers from the last fifteen years. These papers have been sourced from online repositories such as Google Scholar, Scopus, IEEE, Science Direct and more. The automated screening mode includes five processes which are the acquisition of the image, pre-processing of the data, segmentation of the data, extraction of features and implementation of classification algorithms. Each paper studied was described with its merits and drawbacks in the paper. In a paper by Zhang and Liu (n.d.), a model for cervical cancer detection using SVM-based feature screening was proposed. The model utilized the pap smear image with corroboration of the bottom-up approach. Through these, there are about 4000 features extracted out of which the crucial and the related features are considered. This is done to make the model much more accurate and resource-efficient. During the evaluation of the model, the output showed a better performance than the conventional models. Furthermore, the paper implemented the pixel-level classification of the model which enhanced the detection of cervical cancer. During the evaluation by utilizing certain evaluation metrics, the model showed an accuracy of 98%. Although, the model could be enhanced in the future scope of work with various technological advancements and developments.

In another study, a machine learning-assisted cervical cancer detection approach was proposed by Mehmood et al. (2021). The proposed model analysis and evaluates the chances of cervical formation through automated machine learning algorithms. The novel approach proposed is named as CervDetect. For the implementation, the model has utilized the hybrid module by combing Random Forest (RF) and shallow neural networks for accurate output. To pre-process the data, the proposed model implements the Pearson correlation to both the input variables and output variables. Considering the feature selection approach, the model utilized the random forest technique. During the analysis of the performance of this model, different evaluation metrics have been such as PRF score, FPR, FNR, MSE error and more. On the other hand, Asadi proposed the supervised algorithms of Machine Learning for the prediction of cervical cancer Asadi et al. (2020). For the study, the authors considered different machine learning algorithms such as Quest, CT Tree, RBF-ANN, SVM, and MLP-ANN. Furthermore, the authors considered different parameters and features during the implementation of the algorithms. These parameters can be various patient's lifestyle pieces of information or the health pieces of information. After the implementation of these algorithms, these were evaluated based on different conditions. During the evaluation, it was observed that the Decision Tree outperformed the other algorithms implemented. Apart from the implementation, the paper also discussed various influencing parameters which effects the detection of cervical cancer. In addition to this, the paper also stated the future scope of work with a further set of improvements.

Similarly,the comparative approach to cervical cancer detection through machine learning algorithms has been studied by Shetty and Shah (2018). This paper review different papers which included different machine learning algorithms which are support vector machine, k-nearest neighbours, Random Forest tree (RFT), Classification and regression tree (CART), Artificial Neural Network, and more. The author here compared and evaluated the performance of each model throughout the study. Different parameters and features were considered for the analysis which were the sourcing of dataset, implementation of pre-processing approaches, implementation of machine learning algorithms, the comparative performance of the output, and more. Furthermore, the paper discussed various perspectives and suggestions that could be implemented. In addition to this, the future scope of work has been also stated in the paper. Michal proposed a novel approach to the prediction of cervical cancer identification by a photonic method combined with a machine learning algorithm Kruczkowski et al. (2022). The paper considered the internal and optical parameters while processing the data that make the model robust and efficient. The study model utilized various machine learning algorithms such as Random Forest, eXtreme Gradient Boosting, Naïve Bayes and Convolutional Neural Network. The primary aim of this study was to build an appropriate and accurate model for real-world application. The algorithms implemented here were evaluated using certain evaluation metrics. The dataset was acquired from a lab in Japan. The accuracy of these algorithms was compared by StratifiedKFold. In the final stance, it was observed that the performance of the algorithms was comparable. Although, the proposed model was self-sufficient for real-world application.

A novel and robust approach to automate invasive cervical cancer disease detection at an early stage has been proposed by implementing the appropriate machine learning algorithms Jahan et al. (2021). The mortality rate due to cervical cancer has been increasing only because of late detection of the development of the cancerous cell. Therefore, the authors of this paper develop a robust model to detect cancerous cells at the early stage. For the implementation of the model, this study considers eight ML algorithms which are Multi layer Perceptron (MLP), Random Forest (RF), K-nearest Neighbour (KNN), Decision Tree (DT), Logistic Regression (LR), Support vector classifier (SVC), Gradient Boosting and AdaBoost. Initially, after the acquisition of the dataset, certain processes are followed such as data pre-processing and feature extraction. In the final study, MLP outperformed all other models considered in the study. Furthermore, a certain future scope was also discussed in the paper. Singh studied the prediction of cervical cancer possibility through the implementation of machine learning methods Raghavendran (2015). For the prediction, the paper considered different machine learning classifiers. Furthermore, the paper considered the dataset from the UCI repository. The merit of this model is that it is being integrated with the IoT healthcare system. In the implementation process of the model, the data is initially pre-processed and the crucial features of the data are extracted. These are then implemented and validated for further utilization of the model. The study considered six different machine learning classifiers which are Naïve Bayes, Functions-based logistic SMO, Lazy-based LWL, Meta-based iterative classifier optimizer, rules-based decision table and trees-based decision stump. Although this model lacked certain integration of the features. Naveen and Sumana also studied the approach for early detection of cervical cancer by implementing machine learning algorithms Mugad and Sumana (2021). For the detection of cervical cancer, there are certain parameters to enhance the model which are age, number of sexual partners, number of pregnancies, hormonal contraceptives, HPV, STD, IUD and more. The paper acquired the dataset from the online repositories and pre-processed the data. Crucial features and parameters are extracted from the pre-processed data. Then the data is split into training and test sets in the ratio of 4:1, respectively. Three machine learning algorithms were considered in the paper namely, Decision Tree Classifier, Random Forest Classifier, and K-Nearest Neighbour Classifier. During the evaluation of the model, the KNN algorithm outperformed all the other classifiers. The performance of the model was also evaluated by the AUROC curve value.

2.3 Critiques of Ensemble Learning Algorithms

In the paper by Lu et al. (2020), an ensemble learning approach for diagnosing cervical cancer was proposed. The paper primarily focused on three programs of the model to enhance the approach and make it more efficient. The three programs include the correction of data, implementation of ensemble learning, and building an assistant gene module. In the study, multiple classifiers have been added through the voting strategy. These classifiers are Logistic Regression, Decision Tree classifier (DT), Support Vector Machine (SVM), Multi-Layer Perceptron (MLP) and K-Nearest Neighbour (KNN). For the implementation, the data has been acquired from a Chinese hospital with various questionnaires. The dataset consisted of different features and properties. The model implemented is evaluated using various evaluation metrics such as precision, recall, f-1 score, and accuracy score. In another paper by Sarwar et al. (2015), a hybrid ensemble learning approach for the screening of cervical cancer through the analysis of the pap smear test was proposed. The primary aim of this study was to find an enhanced approach in comparison to the conventional automated algorithms. For the dataset, the paper here utilized the pap smear benchmarked dataset. This dataset was constructed utilizing the real-world pieces of information of the patients from a hospital. Through the pap test, one can detect the pre-cancerous cell which can lead to extreme conditions of cervical cancer. This paper considered seven different hybrid approach for the implementation of the model. For a better analysis experience, the author considered the workbench software during the model implementation. Furthermore, the evaluation of model showed that the hybrid ensemble learning outperformed all the other individual algorithms.

On the other hand, E. Ahishakiye and Taremwa (2022), proposed an ensemble learning approach for the prediction of cervical cancer by analyzing the risk factors. There are various studies which utilized various machine learning and deep learning algorithms for the detection of cervical cancer. Although, these conventional algorithms had some drawbacks. Therefore, this paper suggested the implementation of an ensemble learning algorithm which could overcome the challenges faced the conventional algorithms. Different machine learning algorithms were combined in the single models such as KNN, CR Tree, Naïve Bayes, and SVM. The dataset was sourced from the online repositories and pre-processed. After the processing, the crucial features were extracted from the data. Then the algorithms were implemented, trained and validated. During the evaluation of the model, it was stated that the accuracy of the model achieved was 87.21%.

2.4 Critiques of Deep Learning Algorithms

The paper by Chandran proposed the diagnosis of cervical cancer based on ensemble deep learning using colposcopy images Chandran et al. (2021). The primary moat of this study is the utilization of colposcopy images as it has a sufficient contribution to the detection of cervical cancer through deep learning algorithms. This paper considered two different image processing deep learning algorithms which are VGG-19 and Colposcopy Ensemble Network (CYENET). Through these algorithms, various parameters and features are extracted for further implementation by the pooling layer. During the training of the model, the dataset is trained upon about 50 epochs. During the validation and evaluation, the CYENET algorithm outperformed the VGG-19 algorithm with an accuracy of 97.1% against 87%, respectively. On the other hand, Zaid proposed a fully automated deep learning pipeline for cervical cancer pipeline Alyafeai and Ghouti (2019). For the implementation of the model, this study considered the images of cervigram in the cervix to classify the risk for cervical cancer. The proposed model is for two tasks namely the classification of the cervix and detection of the cervical tumours through the pre-trained deep learning algorithms. The algorithm utilized is the light-weighted pre-trained deep learning Convolutional Neural Network (CNN) algorithm. Furthermore, the study utilizes the CNN framework for the extraction of the features from the data. During the evaluation of the models, the first model has a faster computational rate than the conventional model by 1000 times whereas the second model has a faster computational rate than the conventional model by 20 times. In addition to this, the accuracy rate for the first model was 0.68 and for the second model was 0.82.

Park compared the different machine learning and deep learning algorithms implemented for the detection of cervical cancer Park et al. (2021). These algorithms were implemented in the model using the dataset containing the Cervicography images. Furthermore, about ten crucial features were extracted from several other features of the images. The machine learning algorithms considered are XGB, RF and SVM whereas the deep learning algorithm considered are ResNet-50. The dataset was acquired from a hospital with an approved rating for the implementation. Then these acquired datasets were pre-processed utilizing different approaches based on the requirement. After pre-processing, different and crucial features of the images are extracted. These extracted features support the enhancement of the model output and improvise the targeted output accurately. For the evaluation, these models were validated using the five-fold cross-validation approach. Similarly, Xue et al. (2022) did a systematic review and meta-analysis of the deep learning approach in image-based breast and cervical cancer detection. About 2000 research papers were evaluated in the study which were sourced from different online repositories. These online repositories were MEDLINE, Embase, IEEE, Cochrane and more. Although with a thorough analysis of the paper, only 71 papers were considered for in-depth analysis and review. Out of these online 35 papers were selected for the qualitative analysis and 20 papers for the meta-analysis. Through the evaluation, different data visualizations were constructed on the parameters. Various parameters were discussed in the paper considering the advantages, disadvantages, implementations, outcome and feature scope of work. In addition to this, the QUADAS-2 software has been utilized to assess the quality of the research paper.

On the other hand, Peng and the team proposed a novel and robust deep learning approach for assessing image focus for automated cervical cancer screening Guo et al. (2019). In this study, various internal bio-parameters are considered during the implementation of the model. Furthermore, the model aims to enhance the poor quality and improper focused image through deep learning. About 4500 images were collected from the different databases and were enhanced. For the implementation, three different deep learning frameworks were considered. These deep learning algorithms were objected detection models consisting of Retina Net, fine-tuned deep learning model consisting of VGG and Inception, and transfer learning models consisting of VGG and Inception Feature Extraction with VGG. In the final evaluation, the accuracy achieved by the model stood at 94%. Koji in the paper proposed the approach and compared the other approaches for predicting the survival rate for the cervical cancer-induced patient Matsuo et al. (2018). This paper compared the deep learning approach and the Cox proportional hazard regression model. The cox model is the linear regression model which does not include or interpret the non-linear models. In the study, five different deep learning algorithms were considered and compared with the cox model. Various evaluation metrics were considered which are PRF score, accuracy score, AUC score, Mean Absolute Error and Concordance Index. Furthermore, the dataset acquired was spilt into three feature sets based on three different features. These three features were age, ethnicity and tumour severity level. Although this was a pilot study which needed further room for improvement for the real-world application.

2.5 Identified Gaps and Conclusion

After reviewing all the prior work for detection of cervical cancer, it has been found that very limited cytology image data is available publicly for research. Accurately detection of cervical cancer using deep neural network architecture is still an open area of research, where the results obtained from most of the common algorithms such as CNN, RNN are not found to effective and prominent. Therefore, in this research Generative Adversarial Networks has been proposed which is mainly popular for unsupervised learning. However, in recent time for many application it has been used as an classifier. The predicted outcomes of GAN are found to be quite higher as compared to other algorithms. Therefore, in this research GAN architecture will be implemented over the Cervical cancer dataset.

3 Cervical Cancer Methodology Approach

Since the last decade, it is observed that various advancement in medical fields leads to a decrement in the mortality rate and various disease incidences are being detected in a timely manner. Cytology is the process in the medical where the analysis of the cell is carried out for the diagnosis of the diseases such as cancer. Cytology slides are being used for the detection of abnormalities by segmentation of nuclei and cytoplasm. This is a very tedious and time taking task therefore, this research utilizes the advance deep neural network architectures such as GAN for the classification and detection of the diseases by using these slides. The primary task of this research is to recognize the superlative deep learning algorithm which can detect and classify the cytology slides into respective classes more accurately by training on the cervical cytology dataset. Our methodology includes the several set of stages which has been explained in more detailed in further subsections and the flow diagram of our methodology is shown in Figure 1.

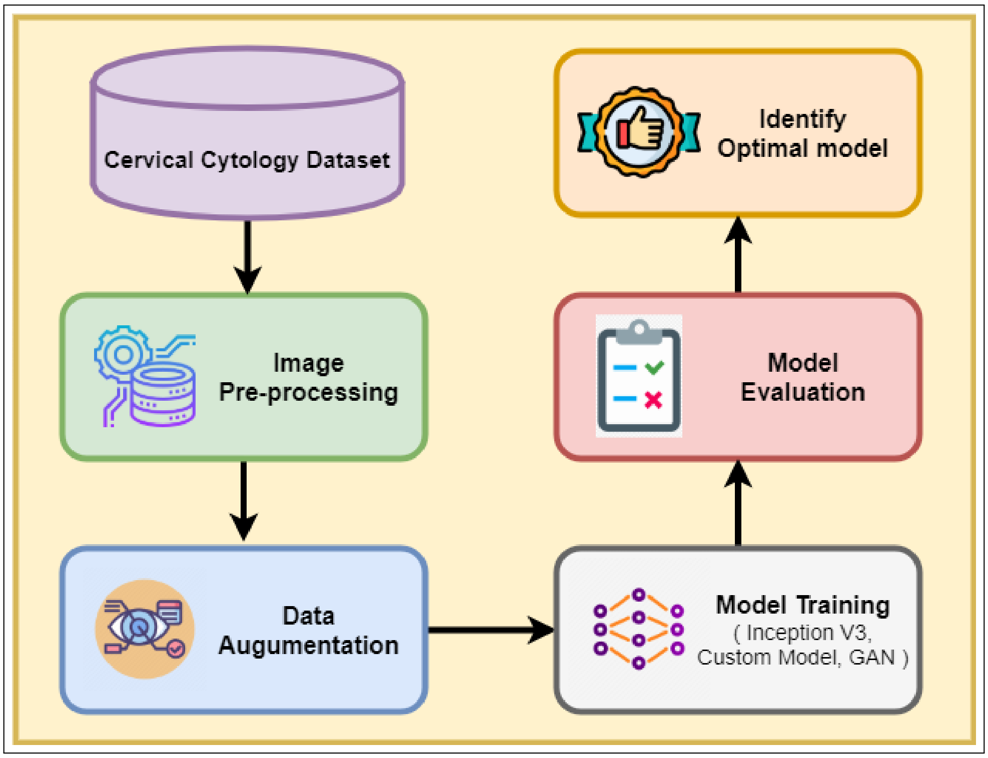

Figure 1: Methodology for Cervical Cancer Detection

3.1 Dataset Description

Cytology is an integral part of the diagnosis of fatal diseases such as cancer because it includes the study of cells, the dataset for this task is collected from the cervical cytology repository on GitHub Phoulady and Mouton (2018). This dataset comprises 93 EDF images. These images are stacked from the slides and categorized into three test categories which are LSIL (Low-grade Squamous Intraepithelial Lesions), HSIL(High-grade Squamous Intraepithelial Lesions), and negative based on the Bethesda method. In this data 2705 nuclei are marked manually and corresponding labels are provided in CSV file format. In this data 16 negative, 46 LSIL, and 31 HSIL images are distributed and in each category marked nuclei are 238, 1536, and 931 respectively which are shown in Figure 2.

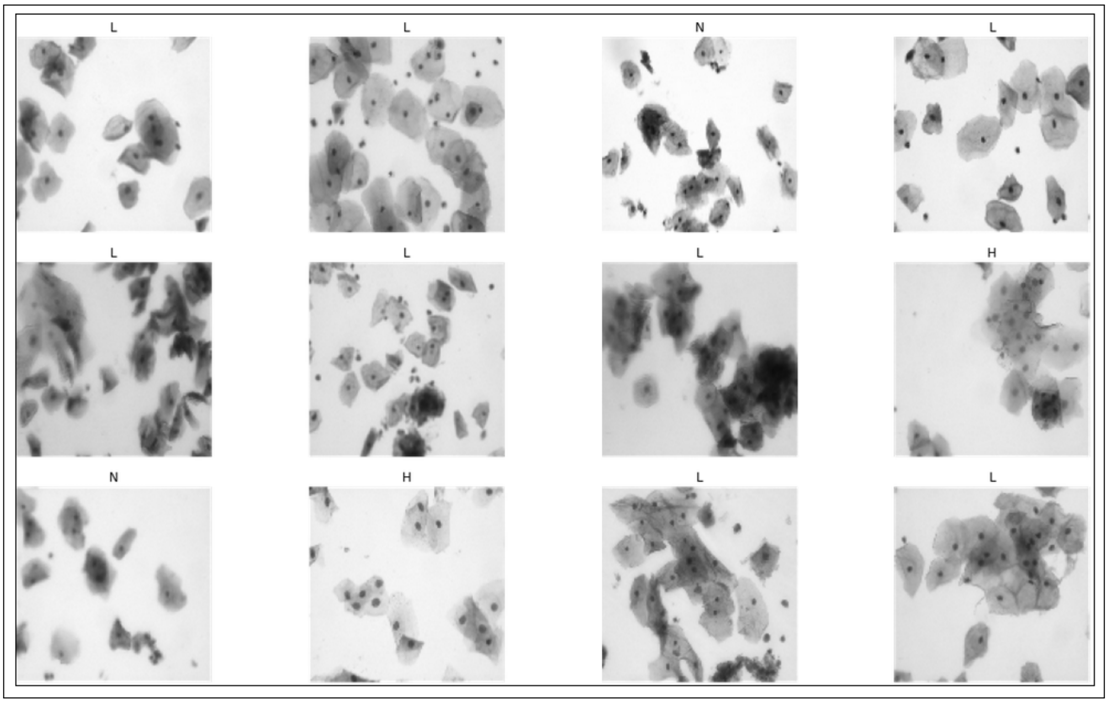

Figure 2: Sample Data Phoulady and Mouton (2018)

3.2 Data Pre-Processing

After collecting the data from a legitimate source, the next step performed is the pre-processing of the data. Pre-processing is required while administering the deep learning-based algorithms for the classification because the rough data may contain noise and variable sizes of images. To check for the noise content and other unnecessary content gaussian filters and sharp filters are implemented on the src images which is shown in Figure 3.

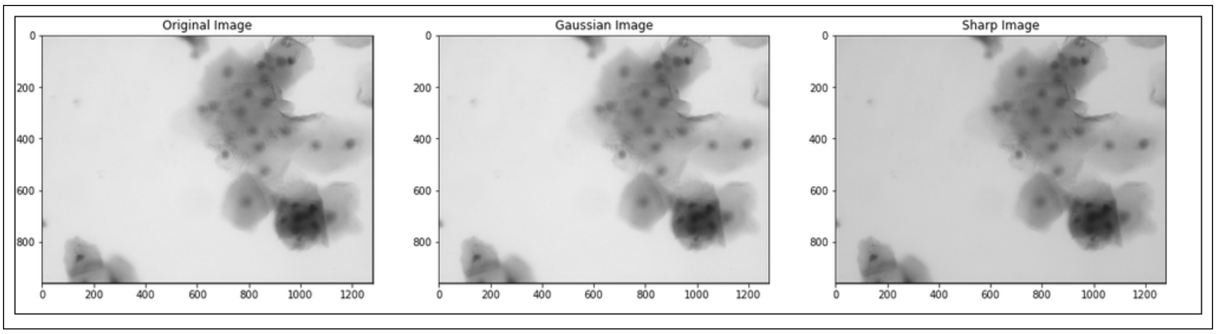

Figure 3: Output of different Filters on Images

The labels are converted into the required format by executing the one-hot encoder method. Next, the images are mapped corresponding to their labels and converted into arrays. In further steps, the photos are augmented. This augmentation includes the methods such as rotation, flipping, zoom, and shear range along with the scaling of the data. This augmentation and pre-processing help in removing biasing of models and results in better outcomes if the dataset is small. The sample of data is shown after applying the data augmentation in Figure 4

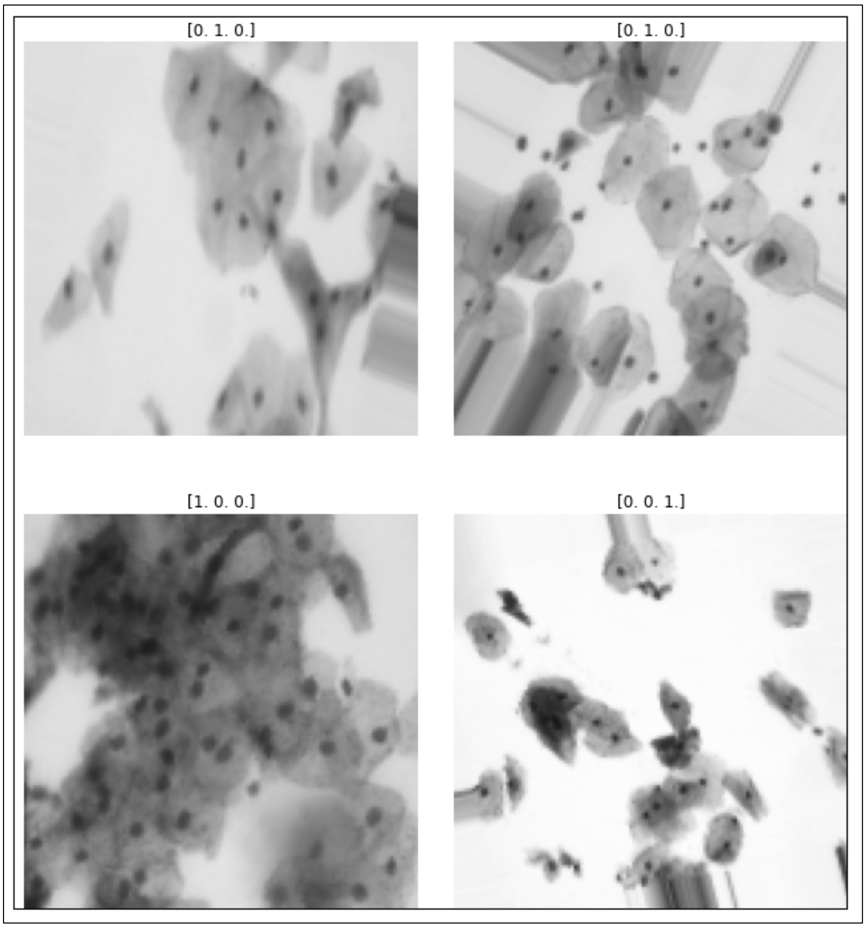

Figure 4: Applying Data Augmentation techniques on Dataset

3.3 Model training

After the pre-processing step, the processed data is split into training and test data in the ratio of 75:25. This training data contain 12 negatives, 34 LSIL, and 23 HSIL frames which include 179,1125, 679 nuclei respectively while the test data contains 4 negatives, 12 LSIL, and 8 HSIL frames which include 59, 411, 252 nuclei respectively. In this classification task, three deep learning-based models are deployed which are divided into categories of transfer learning, custom model, and GANs model. In the transfer learning method here Inception V3 model is executed while the custom model and GANs are made from scratches using various Convolutional layers. These all models are first trained over train data and then tested on test data.

3.4 Model Evaluation

After the successful training of the models on the training data, each model is assessed over the test data. Since this task is a classification task therefore for the evaluation of the model's classification-based metrics is utilized. These metrics are accuracy, precision and recall . The model which achieves the highest value in accuracy, Precision and Recall score is considered the best performing model. Validation loss can also be used for the evaluation of the models, which should be minimum.

4 Design Specification

Algorithms using deep learning are crucial for generating predictions and assessing behavioral traits. Three deep learning algorithms are therefore used to classify cytology slides. Among these, we are using the pre-trained weight for one transfer learning model to get correct results. The last two designs used in this research are CNN-based Custom architectures, one of which is a deep convolutional neural network and the other of which is a GAN-based architecture.

4.1 Inception V3 Model

Convolutional neural networks serve as the foundation for the Inception V3 deep learning model, which is used to classify images. Inception V3 is an enhanced version of Inception V1, the base model was firstly proposed as Google Net. It was created by the Google team, as the name implies. The data was over-fitted when many deep convolutional layers were used in the model. To avoid this, the Conception V1 model employs the concept of many different-sized filters on the same surface. As a result, the initial model employs parallel layers rather than deep layers to make the model wider rather than deeper. The Inception V3 model is simply an enhanced and modified version of the Google Net. Several methods were used in the Inception V3 model to optimize the network and improve model compatibility. It is more effective than the Inception V1 and V2 models, and it has a larger network, but it is slow. We use an auxiliary classifier to reduce the computational cost of normalization. The Inception v3 model, which debuted in 2015, has 42 layers and a lower error rate than its predecessor. Figure 5 illustrates the architecture of Inception V3 model.

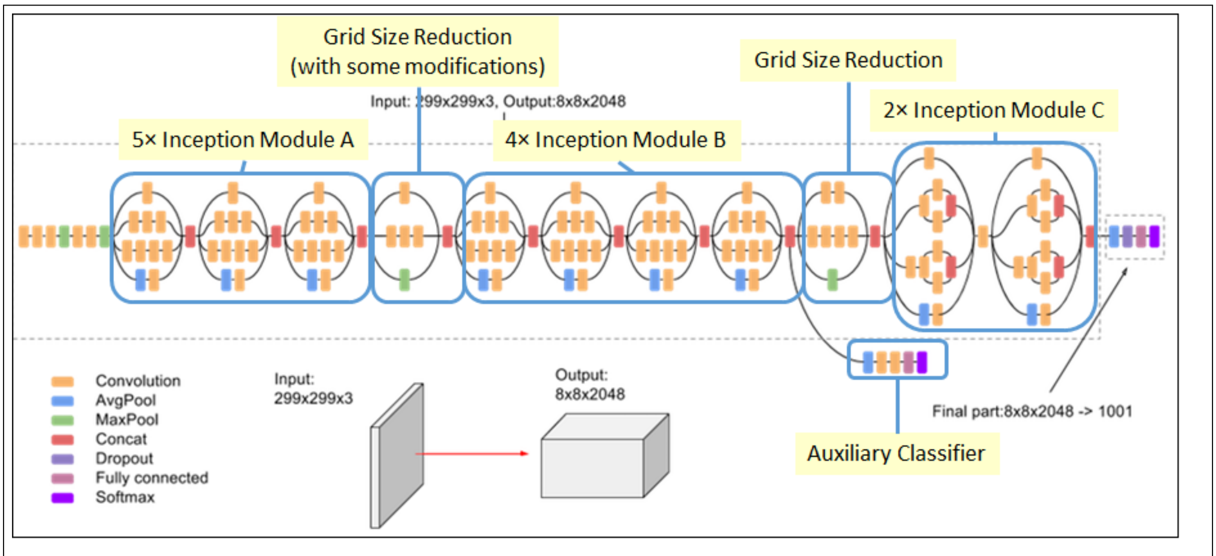

Figure 5: Architecture of Inception V3 Szegedy et al. (2015)

4.2 Custom Model

In this research task, TensorFlow and Keras have been used to create this sequential model. In this model 4 convolutional 2-dimensional layers have been used having 64 neurons with (3,3) filter with the same padding set. ReLu activation function is used by adding the keyword "relu" and pool size of max pool layer is defined as (2,2) is applied followed by the Batch normalization layer. Now in a subsequent layer that is the 2nd layer, everything is similar to the first layer except neurons. In the 2nd layer, 128 neurons have been used followed by the max pool layer with the pool size of (2,2) and the batch normalization layer. After that in the 3rd and 4th layers, the same layers have been added with 256 neurons followed by max pooling and batch normalization. Later on, Flatten layer has been added and after that Dense layer is included with the activation function ReLU having input dimension 128. Furthermore for avoiding over-fitting we have appended the Dropout layer of 0.3 followed by the batch normalization. Since then Dense layer is again implemented having128 neurons with ReLU activation function. At last Dense layer is added with 3 neurons having softmax activation function in the output layer of neural network models that can predict a multinomial probability distribution. In this model categorical cross entropy is used for the loss function and adam optimizer for the optimization purpose. Custom model architecture is shown in Figure 6.

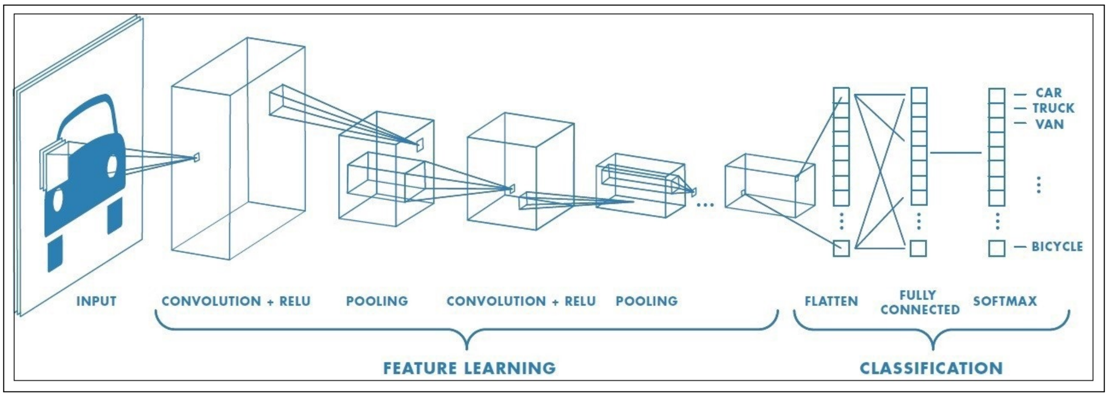

Figure 6: Basic Block Architecture of Custom Model

4.3 Generative Adversarial Network(GAN)

Generative Adversarial Networks (GANs) is one of the most powerful deep neural network architecture, mainly used in unsupervised learning. GANs learn, generate, and improve, allowing them to generate whatever you give them. Understanding GANs necessitates a basic understanding of convolutional neural networks. When an image is fed into a CNN, it analyses it pixel by pixel and passes it through the nodes in the CNN's hidden layers to tell us what the image is. shows. A CNN is taught to classify images based on labels. If a CNN is trained to distinguish between a dog and a cat, for example, the CNN can determine whether the image contains a dog or a cat. As a result, it is also known as a classification algorithm. GANs are made up of two parts: the generator and the discriminator. GAN's recognizer function is comparable to CNN's. Unlike CNN, the output layer of GAN can only have two outputs. A CNN's output can be proportional to the number of labels it trains. The detector consists of a convolutional neural network with several hidden layers and an output layer. The output is either 1 or 0, depending on the activation function that the differentiate carefully selected for this assignment. If the output is 1, the data is considered true; otherwise, it is false. The discriminator is trained on real data to learn what real data looks like and what characteristics real data should have. Generator: As the name implies, this is a generating algorithm. An inverse convolutional neural network (CNN) is a generator that works exactly like a CNN in that a real image is given as input and a classified label is expected as output, but in a generator, random noise (specifically, a vector with a small number of values) is given as input and the actual image is expected as output. Simply put, use your imagination to create data from data. The architecture of Generative Adversarial Network(GAN) is shown in Figure 7.

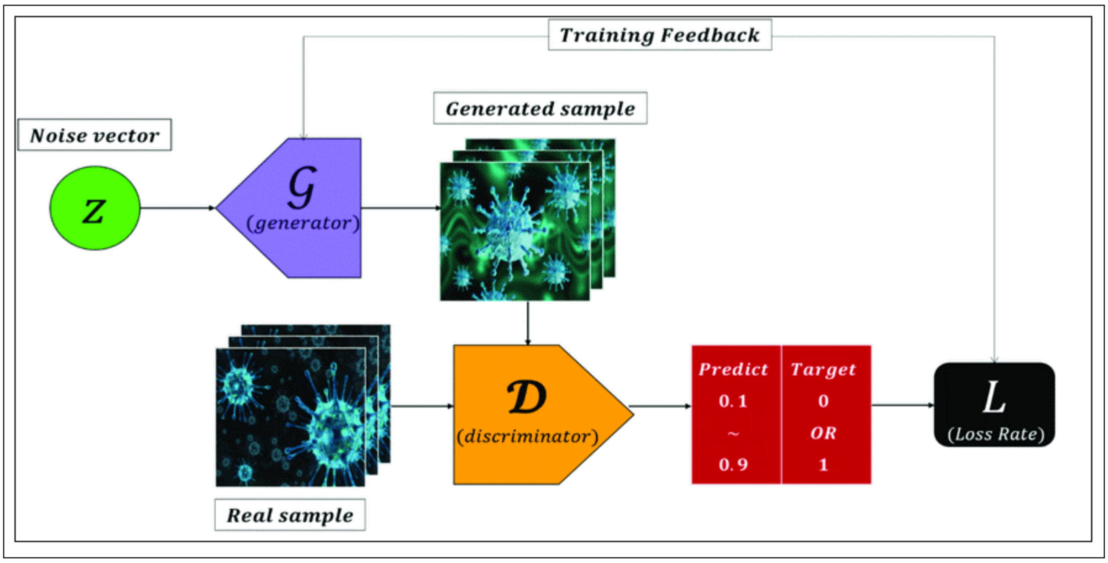

Figure 7: Architecture of GAN Model Goodfellow et al. (2014)

5 Implementations of Cervical Cancer Model

In this work three deep learning-based algorithms are executed which are Inception V3, custom model, and GAN, and the superlative model can be carried out whose value of accuracy, precision, and recall on the test data will be highest. These all models are incorporated with categorical cross entropy as the loss function. To minimize the losses adam optimizer can be used. Each model is trained over the same data and then tested over the same data and each metric is calculated. During the implementation, various libraries are utilized which are NumPy, scipy, plotly, sklearn, OpenCV, TensorFlow, Keras, tqdm, etc. Since the proposed process involves CNN, therefore, google collab is used for the training of the models as it provides free GPU services. Here python language is used for the programming purpose. The following specifications are required in order to implement the model.

Resources	Specification
Operating System (OS)	Windows 10
Main Memory (RAM)	8GB
Hard disk	256GB SSD and 1TB HDD
Programming Language	Python
Platform	Jupyter Notebook
Python Libraries	Numpy, Pandas, matplotlib, Tensorflow, OpenCV, Sklearn, Plotly

6 Results & Evaluation

In this research, the aim is to find an optimal model for the classification of the cervical cell images for the cytology which are categorized into three different classes i.e., negative, LSIL, and HSIL therefore it is essential to assess each model based on the classification metrics. This classification is more accurately a multi-class classification task. All three deep learning-based algorithms are evaluated on test data and metrics such as accuracy, precision, and recall is calculated. Post succession of training, every single model is measured on the biases of Accuracy, Precision-Recall, and validation loss. The model which has the highest value of these matrices is selected for future predictions. A comparative analysis in form of bar charts is also incorporated for it.

6.1 Experiment 1 : Evaluation Based on Accuracy

The Inception V3 model is trained upon 50 epochs. In the process of training, it was seen that the model was starting to overfit as the validation loss was increasing, after 50 epochs. Hence the premature stopping of the model is done to prevent it from overfitting. In the course of training of the model, the top accuracy of the inception v3 model was 90.31%. On analysis of the graph, training accuracy was increasing at every epoch and but validation accuracy was zigzagging during the training, which in turn represents overfitting in the model. The Custom CNN(Convolutional Neural Network) model is trained upon fifty epochs and in the fifty epochs, the model is showing signs of overfitting hence to resolve this situation model was trained only for fifty epochs. The maximum accuracy which was seen using this convolutional model was 90.21%. The GAN model is used for the classification of Cancerous cervical cells since this scratch model is based on the discriminator approach therefore after training the same for 100 epochs, the highest accuracy of the model attained was 97% After the comparison of all implemented models based on the validation accuracy shown in figure 6 and after their analysis it was observed

GAN has a better accuracy score of 97% which was followed by the Custom model at 90.62% and then the inception v3 with an accuracy score of 90.31%. The comparative analysis of accuracy over every epoch for all the models is shown in Figure 8.

Figure 8: Accuracy Comparison of Models

6.2 Experiment 2 : Evaluation Based on Precision and Recall Score

The metrics such as Precision and recall show the Incorrect and correct identified classes between actual and predicted results. The PR(precision and recall)is also calculated for every algorithm used. For Inception V3 the PR score on the validation data are 0.9032, and 0.9031 respectively. For the custom model, the PR score on the validation data is 0.9062, and 0.9062 respectively. For the GAN model, the PR score on the validation data is 0.97 and 0.97 respectively. A comparison of all these implemented models based on PR score is shown in Figures 9 and Figure 10. The highest PR score is scored by the GAN model followed by the custom model.

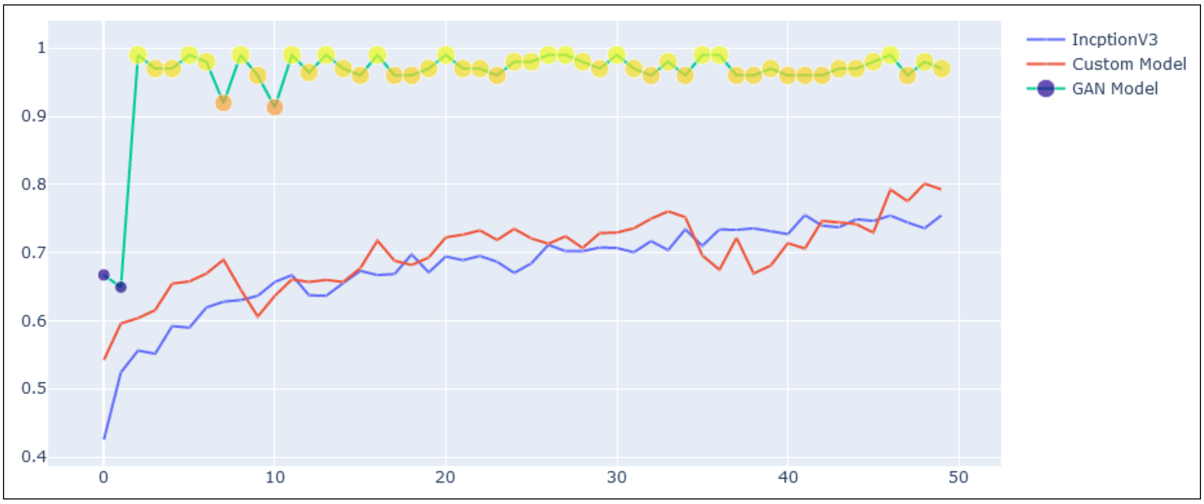

Figure 9: Precision Comparison of Models

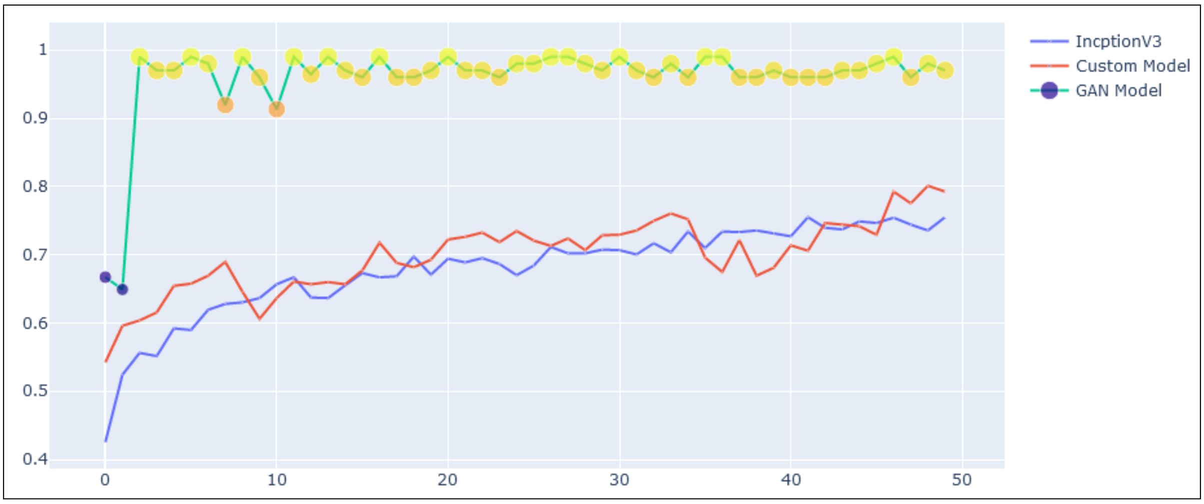

Figure 10: Recall Comparison of Models

6.3 Experiment 3 : Evaluation Based on the Validation Loss

In the following section, the Validation loss is used for the evaluation. For better prediction results values of validation, the loss must be minimum and while training it must be continuously decreasing but if it is increasing it represents the overfitting of the model. For the inceptionV3 model value of validation loss is 0.236. For the custom model value of validation loss is 0.291. on the comparison of all the models, we have a minimum loss in GAN Model of 0.127. The comparison of validation loss is shown in Figure 11.

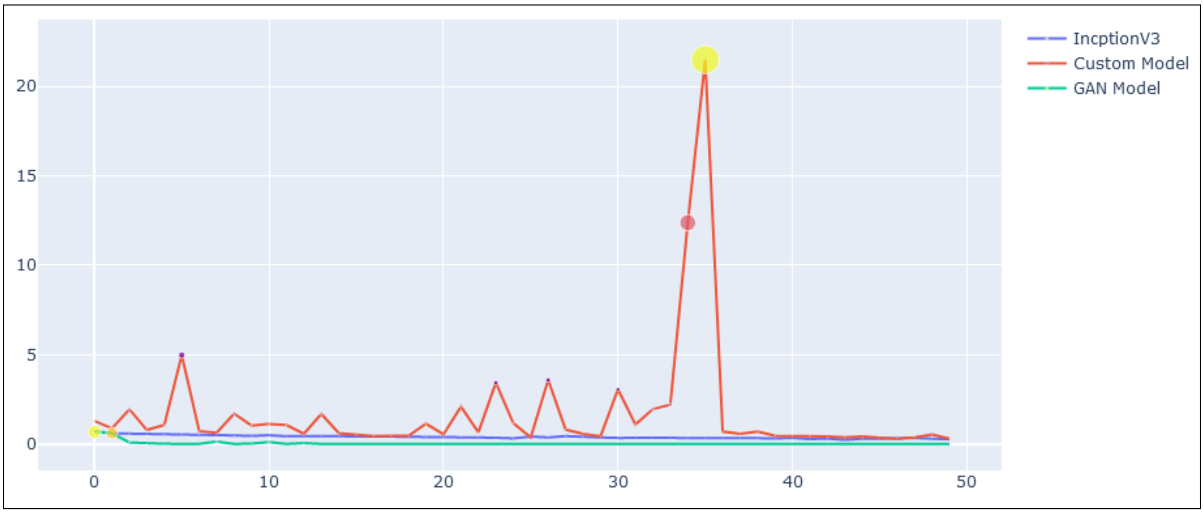

Figure 11: Recall Comparison of Models

6.4 Discussion & Comparison of Developed Models

In this research task after conducting the three experiments on the data by utilizing the deep learning algorithms, it is analyzed that the GAN (Generative Adversarial Network) has surpassed the other executed models. In this task, transfer learning concepts have also been enacted where a model like Inception V3 which is trained on an imagenet dataset is executed but the GAN model exceeds the other models. Here for the evaluation of test accuracy, test Precision and test recall are used in order to make the comparison among the models. The GAN model accomplishes 97% accuracy on test data while the value of precision and recall obtained by this model is 0.97 and 0.97 respectively. Since this model achieved the foremost results therefore it can be utilized further to make outstanding predictions. In real-time scenarios, this model can be deployed for the classification of cervical cells carried out in the medical process of cytology. A custom model is also assessed in this task which performs somewhat lower than the other models because the dataset for the training of the model is small and in such a task large dataset is required for the sequential custom models to perform well. The performance of the inception model is considerable but this model starts overfitting as the value of validation loss increases while training the model. Hence both models can not be utilized for making predictions in real-world tasks.

7 Conclusion and Recommended Future

Classification and detection of cancerous cells into different grades and stages is a very tedious task and a vast area of research because of nuclei and cells overlapping. In this task, different deep learning algorithms are administered to recognize and make predictions on cervical cell frames into corresponding correct classes. This research task is vast and expanding as deep learning is playing a principal role in medical terminologies because these deep learning models are performing well in the segmentation, recognition, and classification of various diseases. In this research work, three deep learning models such as Inception V3, Custom Model and GAN are trained and inspected which results that the GAN model is appropriate for the classification of cervical cells images while the other two models are eliminated because the custom model cannot achieve the highest accuracy due to limitation of the data but still it reaches up to the pre-trained model accuracy which is significant and the pre-trained model inception v3 started overfitting due to same reason. GAN model achieved an highest accuracy of 97% with the precision and recall value of 0.97 and 0.97 respectively. However,the inception model (transfer learning model) also has give the accurate results to some extent, due to high complex architecture. On the other hand, our custom model, which contains only limited number of layers in the architecture was not able to make the prediction on cervical cancer image dataset correctly. Overall, it can be said that GAN is the most accurate and optimal model,able to classify the images into three different classes. In this task, it is observed that the dataset for the purpose is small in size which affects the accuracy of the model and optimizes the performance. In future work, the accuracy of the models can be further enhanced with a large dataset and better predictions can be made. Segmentation of nuclei and cytoplasm can also be incorporated in further works which can be an evolutionary step in the automatic cytology process. Also overall process of detection Cervical cancer can be automated using Artificial intelligence technologies.

8 Acknowledgement

I would like to express my gratitude to my lecturer Dr. Catherine Mulwa for her supervision, guidance and Support throughout the research. I would like to acknowledge my family and friends for encouraging me.

References

Alyafeai, Z. and Ghouti, L. (2019). A fully-automated deep learning pipeline for cervical cancer classification, *Expert Systems with Applications* **141**: 112951.

Asadi, F., Salehnasab, C. and Ajori, L. (2020). Supervised algorithms of machine learning for the prediction of cervical cancer, *Journal of Biomedical Engineering* **10**.

Brüggmann, D., Quinkert-Schmolke, K., Jaque, J. M., Quarcoo, D., Bohlmann, M. K., Klingelhöfer, D. and Groneberg, D. A. (2022). Global cervical cancer research: A scientometric density equalizing mapping and socioeconomic analysis, *PLOS ONE* **17**(1): e0261503.
URL: *https://doi.org/10.1371%2Fjournal.pone.0261503*

Chandran, V., M G, S., Karthick, A. and Subramanian, M. (2021). Diagnosis of cervical cancer based on ensemble deep learning network using colposcopy images, *BioMed Research International* .

E. Ahishakiye, R. Wario, W. M. and Taremwa, D. (2022). Prediction of cervical cancer basing on risk factors using ensemble learning, pp. 1–12.

Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y. (2014). Generative adversarial networks.
URL: *https://arxiv.org/abs/1406.2661*

Guo, P., Singh, S., Xue, Z., Long, L. and Antani, S. (2019). Deep learning for assessing image focus for automated cervical cancer screening, pp. 1–4.

Jahan, S., Islam, M., Islam, L., Rashme, T., Aziz, A., Paul, B. K., Islam, M. and Mosharof, M. (2021). Automated invasive cervical cancer disease detection at early stage through suitable machine learning model, *SN Applied Sciences* **3**.

Kruczkowski, M., Drabik-Kruczkowska, A., Marciniak, A., Tarczewska, M., Kosowska, M. and Szczerska, M. (2022). Predictions of cervical cancer identification by photonic method combined with machine learning, *Scientific Reports* **12**.

Lu, J., Song, E., Ghoneim, A. and Alrashoud, M. (2020). Machine learning for assisting cervical cancer diagnosis: An ensemble approach, *Future Generation Computer Systems* **106**.

Matsuo, K., Purushotham, S., Jiang, B., Mandelbaum, R., Takiuchi, T., Liu, Y. and Roman, L. (2018). Survival outcome prediction in cervical cancer: Cox models versus deep-learning model, *American Journal of Obstetrics and Gynecology* **220**.

Mehmood, M., Rizwan, M., ml, M. and Abbas, S. (2021). Machine learning assisted cervical cancer detection, *Frontiers in Public Health* **9**.

Mugad, N. and Sumana, K. R. (2021). Early prediction of cervical cancer using machine learning algorithms, **09**: 612–620.

Park, Y., Kim, Y. J., Ju, W., Nam, K., Kim, S. and Kim, K. (2021). Comparison of machine and deep learning for the classification of cervical cancer based on cervicography images, *Scientific Reports* **11**.

Phoulady, H. A. and Mouton, P. R. (2018). A new cervical cytology dataset for nucleus detection and image classification (cervix93) and methods for cervical nucleus detection.
URL: *https://arxiv.org/abs/1811.09651*

Raghavendran, R. (2015). Predicting of cervical cancer using machine learning techniques-an analysis.

Sarwar, A., Sharma, V. and Gupta, R. (2015). Hybrid ensemble learning technique for screening of cervical cancer using papanicolaou smear image analysis, *Personalized Medicine Universe* **4**.

Shetty, A. and Shah, V. (2018). Survey of cervical cancer prediction using machine learning: A comparative approach, pp. 1–6.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. (2015). Rethinking the inception architecture for computer vision.
URL: *https://arxiv.org/abs/1512.00567*

William, W., Ware, J., Habinka, A. and Obungoloch, J. (2018). A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images, *Computer Methods and Programs in Biomedicine* **164**.

Xue, P., Wang, J., Qin, D., Yan, H., Qu, Y., Seery, S., Jiang, Y. and Qiao, Y. (2022). Deep learning in image-based breast and cervical cancer detection: a systematic review and meta-analysis, *npj Digital Medicine* **5**: 19.

Zhang, J. and Liu, Y. (n.d.). Cervical cancer detection using svm based feature screening, Vol. 3217, pp. 873–880.

Zhang, S., , Xu, H., Zhang, L. and and, Y. Q. (2020). Cervical cancer: Epidemiology, risk factors and screening, *Chinese Journal of Cancer Research* **32**(6): 720–728.
URL: *https://doi.org/10.21147%2Fj.issn.1000-9604.2020.06.05*