MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY
BHOPAL, INDIA, 462003

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

Endometrial Cancer Detection Using Transfer Learning and Hybrid Approaches

Major Project Report

Semester 7

Submitted by:

Vishal Kumar Rajak	191112006
Maneet Kaur Bagga	191112012
Darshini Sharma	191112021
Uday Chaturvedi	191112053

Under the Guidance of
Dr. Nilay Khare
DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
Session: 2021-22

MAULANA AZAD
NATIONAL INSTITUTE OF TECHNOLOGY
BHOPAL, INDIA, 462003

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

CERTIFICATE

This is to certify that the project entitled "**Endometrial Cancer Detection using Transfer Learning and Hybrid Approaches**" submitted by:

Vishal Kumar Rajak	191112006
Maneet Kaur Bagga	191112012
Darshini Sharma	191112021
Uday Chaturvedi	191112053

is the partial fulfillment of the requirements for the award of the degree of Bachelor of Technology in Computer Science and Engineering is an authentic work carried out by them under my supervision and guidance.

Dr. Nilay Khare
(Major Project Mentor)

DECLARATION

We, hereby declare that the following report which is being presented in the Major Project Documentation Entitled as "**Endometrial Cancer Detection Using Transfer Learning and Hybrid Approaches**" is an authentic documentation of our own original work and to best of our knowledge. The following project and its report, in part or whole, has not been presented or submitted by us for any purpose in any other institute or organization. Any contribution made to the research by others, with whom we have worked at Maulana Azad National Institute of Technology, Bhopal or elsewhere, is explicitly acknowledged in the report.

Vishal Kumar Rajak	191112006
Maneet Kaur Bagga	191112012
Darshini Sharma	191112021
Uday Chaturvedi	191112053

ACKNOWLEDGEMENT

With due respect, we express our deep sense of gratitude to our respected guide and coordinator Dr. Nilay Khare, for his valuable help and guidance. We are thankful for the encouragement that he has given us in completing this project successfully.

It is imperative for us to mention the fact that the report of major project could not have been accomplished without the periodic suggestions and advice of our project guide Dr. Nilay Khare and project coordinators Dr. Mansi Gyanchandani and Dr. Bholanath Roy.

We are also grateful to our respected director Dr. N. S. Raghuwnanshi for permitting us to utilize all the necessary facilities of the college.

We are also thankful to all the other faculty, staff members and laboratory attendants of our department for their kind cooperation and help. Last but certainly not the least; we would like to express our deep appreciation towards our family members and batch mates for providing the much-needed support and encouragement.

ABSTRACT

In recent years, machine learning algorithms had shown significant success at picture recognition tasks, coinciding with a period of dramatically increased use of electronic medical records and diagnostic imaging. This report discusses machine learning technique and how they are used in medical picture analysis mainly on endometrial cancer concentrating on convolutional neural networks, with a focus on clinical aspects of the benefit of the field.

In an era of medical big data, one of the challenges of machine learning is that there are large hierarchical links within the data. Without the need for tedious hand-crafting of features, data can be discovered algorithmically. We go over the most important research, categorization, localisation, and detection of medical images: regions and applications.

Computer-aided detection systems use computer technologies to detect abnormalities in clinical images which can assist medical professionals in a faster and more accurate diagnosis. In recent times, the performance of computer-aided diagnosis systems in classification of malignancies has significantly improved. Search and retrieval methods are specifically important as they assist physicians in making the right diagnosis in medical imaging owing to their ability of obtaining similar cases for a query image.

Histopathology is a method used for endometrial cancer diagnosis. Machine learning (ML) methods have achieved success for supervised learning tasks in the medical domain. Supervised classification algorithms are generally more accurate than unsupervised search-based classifications. Endometrial histopathology analysis is the gold standard for diagnosing endometrial cancer. Convolutional neural network-based methods for breast histology image classification have emerged in recent years to make the analysis process simple and fast.

Herein, we have considered a Histopathology image dataset for ET having 3302 images comprising of four classes: EA, EH, EP and NE. We first divide the entire dataset in terms of types of cancer and infections and then classify by dividing the entire dataset into training, validation and test set.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

In clinical medicine, histopathology refers to the examination of a biopsy or surgical specimen by a pathologist, after the specimen has been processed and histological sections have been placed onto glass slides. Cytopathology examines free cells or tissue micro-fragments.

Histopathological analysis is a highly, time consuming specialized task, dependent on the experience of the pathologists and influenced by factors such as fatigue and decrease of attention. There is a pressing need for computer assisted diagnosis (CAD) to relieve the workload on pathologists by filtering obviously benign areas, so that the experts can focus on the more difficult-to diagnose cases.

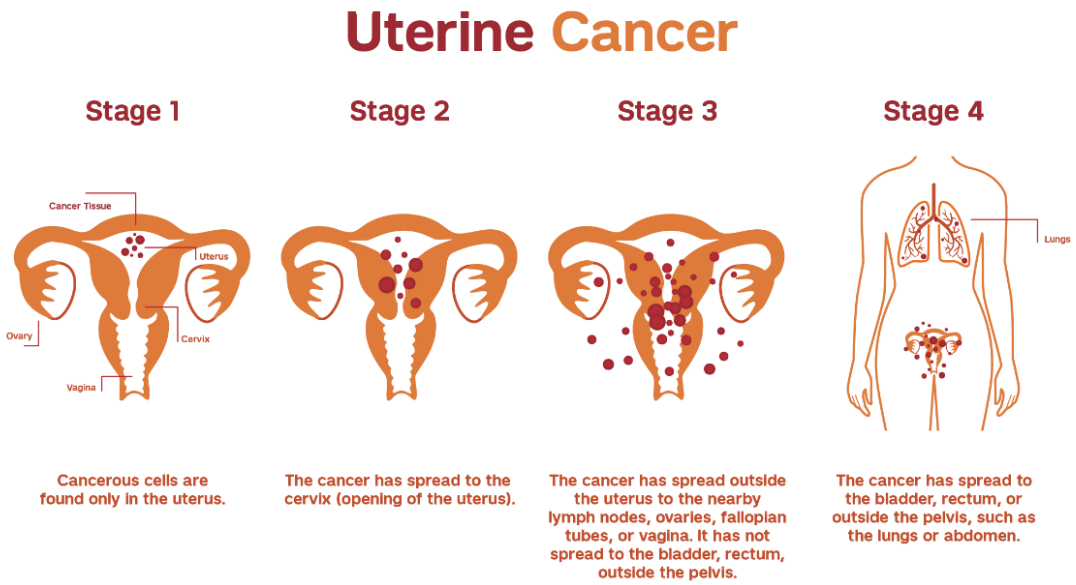

Fig. 1. Stages of Uterine Cancer [1]

Cancer of the uterine corpus (or corpus uteri), also called uterine cancer (as opposed to cervical cancer), was the sixth most frequently diagnosed cancer among women worldwide, with an estimated 319,600 diagnosed new cases in 2012 [2]. Endometrial cancer, which arises from the endometrium of the uterus, is the most common form of uterine cancer. Therefore, endometrial cancer is sometimes known as uterine cancer in a general sense. Incidence and mortality rates for endometrial cancer in females are higher in developed countries, and the incidence rate is increasing [3].

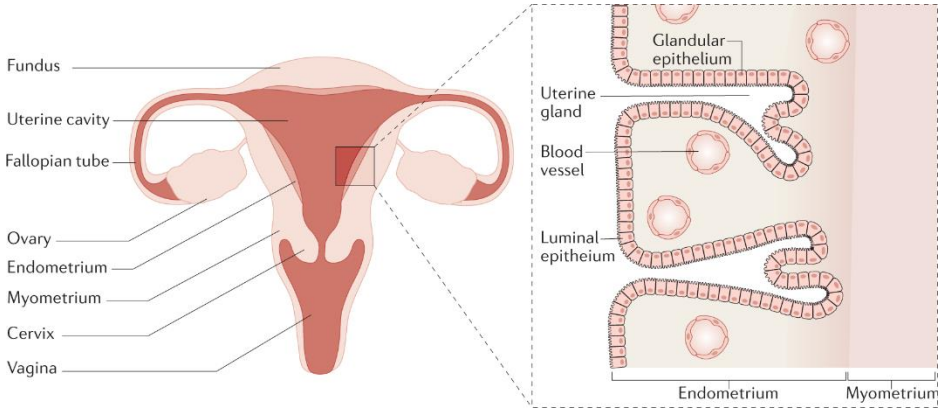

Fig. 2. Endometrial Cancer [4]

According to a report released by the American Cancer Society in 2018 [5], endometrial cancer is the fourth most common cancer overall in the United States, accounting for 7% of all new cancer diagnoses in women. In general, if endometrial lesions can be detected early using commonly-used clinical screening and detection techniques, such as transvaginal ultrasound, hysteroscopy, and hysterosalpingography, the treatment outcome will be favorable. Moreover, the five-year survival rate for endometrial cancer after undergoing appropriate treatment is over 80% [6].

Since computer-aided diagnosis (CAD), also known as computer-aided detection (CADe), can assist doctors in efficiently analyzing and evaluating a vast number of medical images like ultrasound images and X-rays, it has been used in the radiologic diagnosis of some common cancers, including breast cancer, lung cancer, and colon cancer, in clinical environments.

Due to the small size of training samples and limited capability of feature extraction (in other words, heavy dependence on handcrafted image features), these CAD systems, however, failed to achieve an adequate level of overall performance. Thus, the combination of CAD and deep learning has great potential to improve further the efficacy of traditional CAD systems for endometrial cancer using big data in clinical imaging.

The purpose of this study is to develop a CAD approach based on deep learning to assist pathologists in efficiently evaluating histological images from endometrial tissue samples stained with hematoxylin and eosin (H&E). In addition to accurate image classification, we attempt to provide a comparative study between different CAD approaches. Therefore, we expect our work may help improve the efficiency and productivity of pathologists in diagnosing endometrial diseases. [7]

LITERATURE REVIEW AND SURVEY

- Class-Aware Image Search for Interpretable Cancer Identification: It proposes two distinct scenarios to induce "class-awareness" into the search process. In one scenario, it restricts the search to dominant classes, and in the other, it resorts to search results before casting the majority vote by adjusting the matching distances. As there are more than 2 categories in each dataset, considering each label in the classifier, the label of the target category is denoted as positive and the label of all other categories as negative. [8]

- Fine-Tuning and training of DenseNet for histopathology image representation using TCGA diagnostic slides: It proposes the 'Kimianet' architecture training which can be used as "feature extractor" for image analysis. Customizing well-established architectures for specific and sensitive tasks not only appear to be justified but also seem to be necessary for the sake of application-oriented categorization and end-user awareness. It proposes work on TCGA images, Endometrium dataset and Colorectal Cancer Images. Deep network is trained by a variety of tumor types which will provide better image features. Using image patches at 20X magnification for training will help network to focus on cell nuclei distributions and shapes. [9]

- Computer-Aided Diagnosis in Histopathological Images of the Endometrium Using a Convolutional Neural Network and Attention Mechanisms: It proposes a CAD approach based on a convolutional neural network (CNN) and attention mechanisms, called HIENET. The work outperformed three human experts and five CNN-based classifiers regarding overall classification performance. It was also able to provide pathologists better interpretability of diagnoses by highlighting the histopathological correlations of local pixel-level image features to morphological characteristics of endometrial tissue.

- Classification of lung cancer histology images using patch-level summary statistics: The work presents an automated method for NSCLC classification, that consists of a two-part approach. Firstly, they implement a deep learning framework to classify input patches as LUAD, LUSC or non-diagnostic (ND). Next, they extract a collection of statistical and morphological measurements from the labeled whole-slide image (WSI) and use a random forest regression model to classify each WSI as lung adenocarcinoma or lung squamous cell carcinoma. [10]

PROPOSED WORK AND METHODOLOGY

A. DATASET

Endometrial cancer is the most prevalent form of uterine cancer arising from the endometrium (the lining of the uterus). The detection of endometrial and colorectal cancers in their earliest stages often allows for more treatment options and higher patient survivability.

For endometrial cancer, we use a publicly available image dataset containing 498 endometrial tissue samples. Digital images saved in a JPEG format at 10× or 20× magnification. A total of 3,302 histopathologic image patches (640 × 480 pixels) were extracted from each endometrial whole-slide image and labelled by the class number describing the type of disease. The taxonomy of the provided histopathology image patches shows four different classes—normal endometrium (NE), endometrial polyp (EP), endometrial hyperplasia (EH), and endometrial adenocarcinoma (EA). The below table represents these categories and the number of image patches available for each class. The dataset is publicly available. [8]

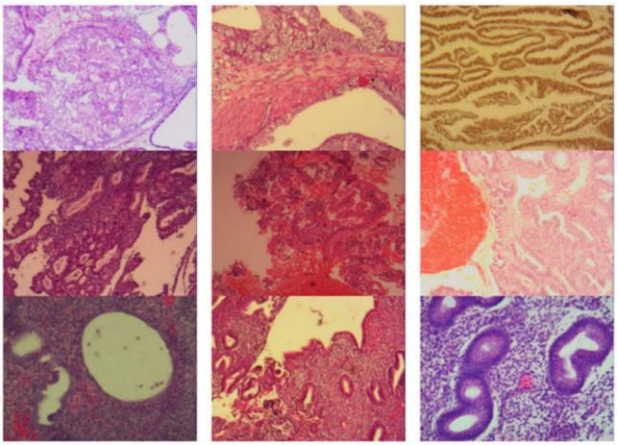

Fig. 3. Sample patches from the endometrial cancer dataset [8]

General class	Fine-grained class	Subtype	#Image Patches
Benign	Normal Endometrium	Luteal phase	600
		Menstrual phase	21
		Follicular phase	712
	Endometrial Polyp	NA	636
	Endometrial Hyperplasia	Simple	516
		Complex	282
Malignant	Endometrial Adenocarcinoma	NA	535
		Total:	3,302

Table 1. Introduction to the experimental dataset [7]

Models used in this study detect these types of cancer or infection. There are 535 images in the set EA. In the second EH set, there are two labels—Simple and Complex. The number of simple EH images is 516, and the number of complex EH images is 282. The third set of EP consists of 636 images. In the last set, NE, there are three labels - follicular, luteal, and menstrual. There are 716 images on the follicular label, 600 images on the luteal label, and 21 images on the menstrual label.

B. ARCHITECTURES

The dataset consists of 3302 images which is split into 2 sets of training and test set with ratio 80:20. Thereafter, test set is further split into test set and validation set in same ratio 90:10. Training set is the input to the proposed models and we get the output in form of binary classification: Benign for labels – NE, EP, EH and Malignant for label – EA.

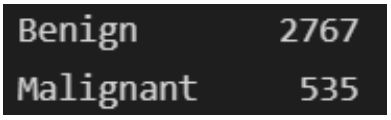

Fig. 4. Total number of images of Benign and Malignant

B.1. Transfer Learning using DenseNet201

We will use the "DenseNet201" as the pre trained weights which is already trained on the Imagenet dataset with a learning rate of 0.0001. Furthermore, globalAveragePooling layer followed by 50% dropouts to reduce over-fitting. Using batch normalization and a dense layer with 2 neurons for 2 output classes i.e., "benign" and "malignant" with softmax as the activation function, Adam as the optimizer and binary-cross-entropy as the loss function.

Let’s look at the output shape and the parameters involved in each layer.

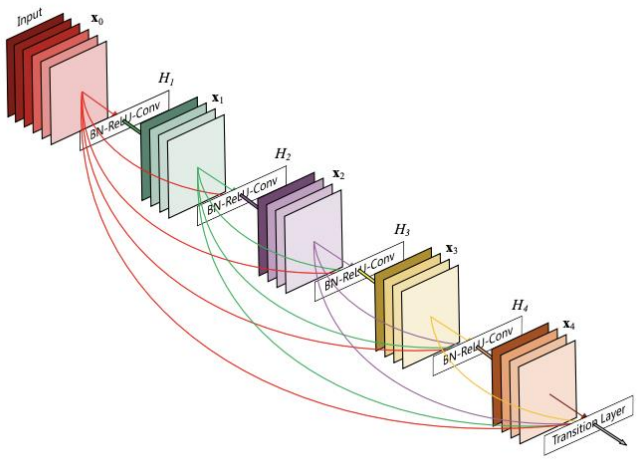

Fig. 5. A visualization of the DenseNet-201 architecture.

Layer (type)	Output Shape	Param #
densenet201 (Functional)	(None, 7, 7, 1920)	18321984
global_average_pooling2d (GlobalAveragePooling2D)	(None, 1920)	0
dropout (Dropout)	(None, 1920)	0
batch_normalization (BatchNormalization)	(None, 1920)	7680
dense (Dense)	(None, 2)	3842

Total params: 18,333,506
Trainable params: 18,100,610
Non-trainable params: 232,896

Fig. 6. Model Architecture of transfer learning using Densenet201

The model is designed using the Sequence class instance. As a first step, the model backbone is added. The backbone is our imported and initialized Densenet201 CNN instance. [11]

Before training the model, it is useful to define one or more callbacks. Pretty handy one, are: ModelCheckpoint and ReduceLROnPlateau.

- **ModelCheckpoint**: When training requires a lot of time to achieve a good result, often many iterations are required. In this case, it is better to save a copy of the best performing model only when an epoch that improves the metrics ends.

- **ReduceLROnPlateau**: Reduce learning rate when a metric has stopped improving. Models often benefit from reducing the learning rate by a factor of 2–10 once learning stagnates. This callback monitors a quantity and if no improvement is seen for a ‘patience’ number of epochs, the learning rate is reduced. [12]

B.2. Transfer Learning using InceptionV3

We will use the "InceptionV3" as the pre trained weights which is already trained on the Imagenet dataset. Furthermore, globalAveragePooling layer followed by batch normalization and a dense layer with 2 neurons for 2 output classes i.e., "benign" and "malignant" with softmax as the activation function, Adam as the optimizer and binary-cross-entropy as the loss function.

Let’s look at the output shape and the parameters involved in each layer.

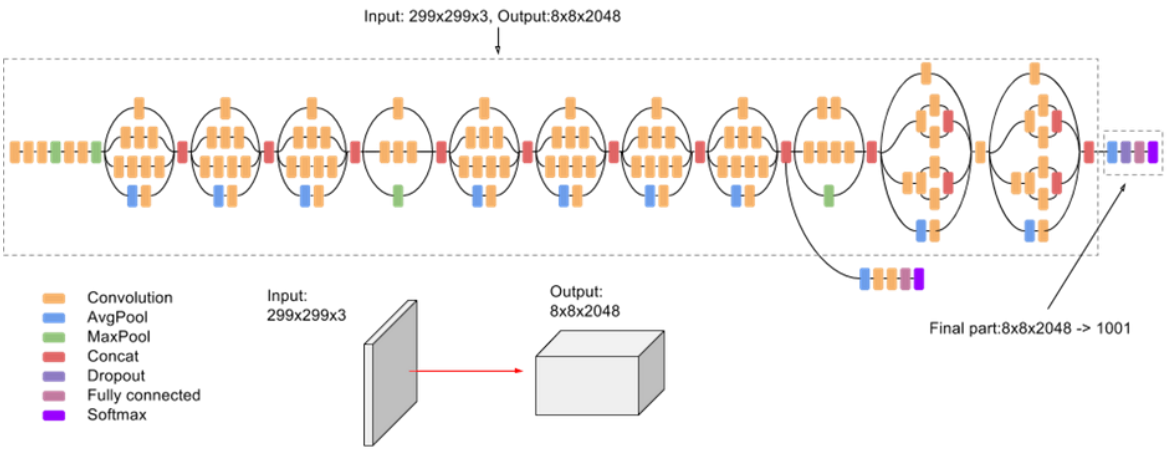

Fig. 7. A visualization of the InceptionV3 architecture. [13]

As a first step, the model backbone is added. The backbone is our imported and initialized InceptionV3 CNN instance.

Before training the model, it is useful to define one or more callbacks. Pretty handy one, are: ModelCheckpoint and ReduceLROnPlateau.

B.3. Hybrid of Resnet + Xception and Modified Xception

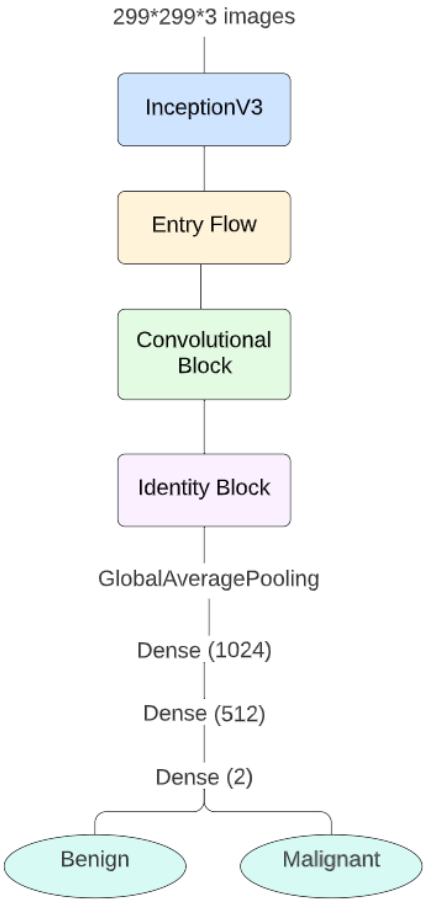

Fig. 8. Hybrid Architecture of Resnet and Xception

The Modified Xception model starts with the images as input. Here, no pre-weights are used instead weights are learned during the time of training itself. Training produces same set of accuracy and loss measures in every epoch. This leads to the AUC (area under curve) of 0.5.

Similarly, the hybrid model of Resnet and Xception produces same set of accuracy and loss measures. The AUC (area under curve) is 0.5. Here, the 299*299*3 image is provided as an input to the InceptionV3 backbone with no pre-defined weights of 'Imagenet'. The output from the backbone is now fed to another block which is the entry flow. This passes through 2 more blocks: convolutional and identity blocks. The output from the last layer is passed through globalAveragePooling layer and then fully connected layers which gives the result in form 2 classes: benign or malignant.

C. PERFORMANCE METRICS

The most common metric for evaluating model performance is the accuracy. However, when only 2% of your dataset is of one class (malignant) and 98% some other class (benign), misclassification scores don't really make sense. We can be 98% accurate and still catch none of the malignant cases which could make a terrible classifier.

For a better look at misclassification, we often use the following metrics to get a better idea of true positives (TP), true negatives (TN), false positive (FP) and false negative (FN).

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

$$Precision = \frac{TP}{TP + FP}$$

Recall is the ratio of correctly predicted positive observations to all the observations in actual class.

$$Recall = \frac{TP}{TP + FN}$$

F1-Score is the weighted average of Precision and Recall. [12]

$$F - measure = \frac{2*Recall*Precision}{Recall + Precision}$$

RESULTS AND DISCUSSION

In this work, we present results obtained from the work of endometrial cancer classification on the DL4ETI dataset.

A. Transfer learning using Densenet201

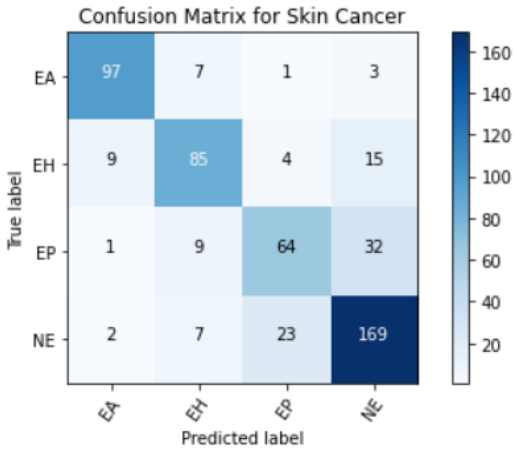

Fig. 9. Confusion matrix for 4-class classification.

	precision	recall	f1-score	support
0	0.89	0.9	0.89	108
1	0.79	0.75	0.77	113
2	0.7	0.6	0.65	106
3	0.77	0.84	0.8	201
accuracy			0.79	528
macro avg	0.79	0.77	0.78	528
weighted avg	0.78	0.79	0.78	528

Table 2. Classification report for 4-class classification.

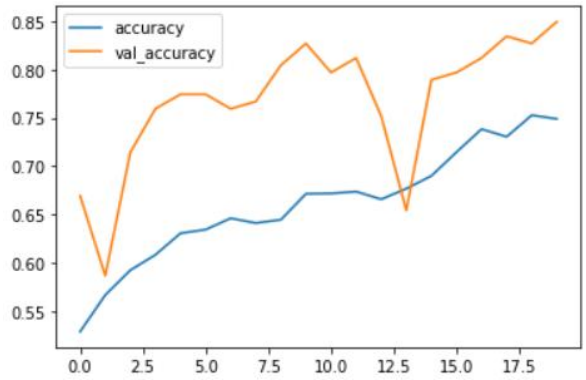

Fig. 10. Visualization of accuracy during training.

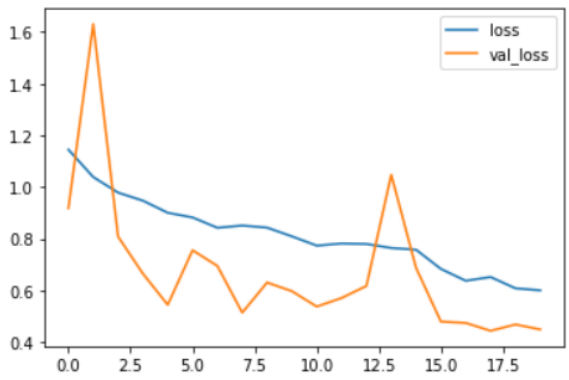

Fig. 11. Visualization of loss during training.

B. Transfer learning using InceptionV3

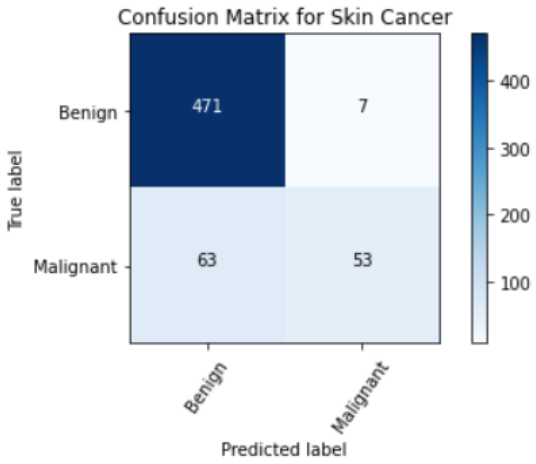

Fig. 12. Confusion matrix for 2-class classification.

	precision	recall	f1-score	support
0	0.88	0.99	0.93	478
1	0.88	0.46	0.6	113
accuracy			0.88	594
macro avg	0.88	0.72	0.77	594
weighted avg	0.88	0.88	0.87	594

Table 3. Classification report for 2-class classification.

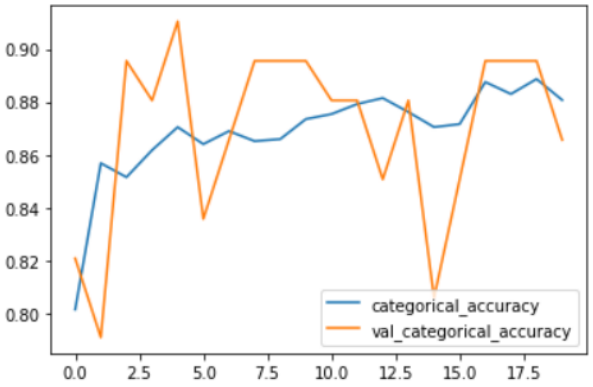

Fig. 13. Visualization of accuracy during training.

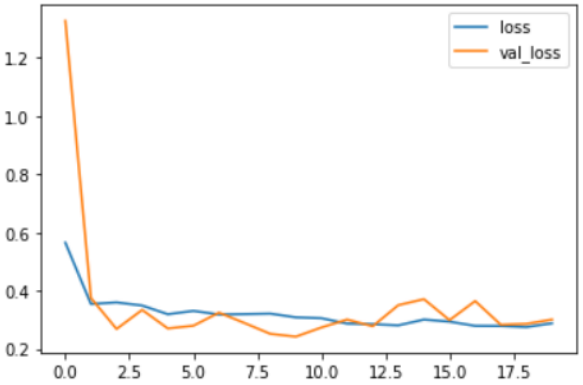

Fig. 14. Visualization of loss during training.

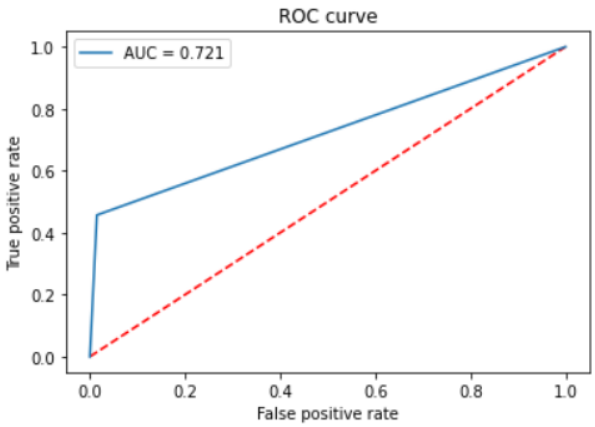

Fig. 15. ROC Curve for the model.

C. Modified Xception

```
    accuracy_score(np.argmax(Y_val, axis=1), np.argmax(Y_val_pred, axis=1))
```

```
0.8059701492537313
```

```
    scores = model.evaluate(X_test, Y_test, verbose=0)
    print("%s: %.2f%%" % (model.metrics_names[1], scores[1]*100))
```

```
categorical_accuracy: 80.47%
```

Fig. 16. Accuracy statistics of Modified Xception model.

CONCLUSION

In conclusion, Convolutional neural networks (CNNs) are the current state-of-the-art architecture for automatic classification of histopathological images.

In our proposed system, we compared the performance of transfer learning model using Inception and Densenet201. Also, we compared the study of hybrid model (Resnet + Xception) and modified Xception model.

From results, we observed the following conclusion:

- The dataset comprises of 3302 images in which number of benign images are 2767 and Malignant images are 353. The size of the data set is small. This enables us to perform data augmentation as a part of pre-processing step for better results.
- The goad of a model is to generalize patterns in training data so that we can correctly predict new data that has never been presented to the model. Overfitting occurs when a model adjusts excessively to the training data, seeing patterns that do not exist, and consequently performing poorly in predicting new data.
- The augmentation technique we used is to augment the data during the training itself batch-wise. We can also use the technique in which augmentation is done beforehand the training.

Thus, the experimental results show state-of-the-art testing accuracy for endometrial cancer detection as compared to existing methods. The presented work demonstrates the applicability and powerful classification capacity of machine learning approaches for the automatic analysis of endometrial cancer histopathology images.

However, there are some limitations in this work. We can improve our performance if we provide larger datasets because limited raw data has an effect on accuracy results.

As future work, we plan to extend the classification problem to more tissue categories i.e., subtypes of benign and malignant classes by using data augmentation techniques. In the long run, we plan to design a complete framework for the analysis of Whole Slide Histopathology (WSI) by adding multi-classification of benign subtypes i.e., NE (Normal Endometrium), EP (Endometrial Polyp), EH (Endometrial Hyperplasia) and Malignant types i.e., EA (Endometrial Adenocarcinoma). Having that in mind, we would like to emphasize the need of data augmentation technique to solve the class imbalance problem.

REFERENCES

[1] Endometrial and Uterine Cancers - https://weillcornell.org/services/obstetrics-and-gynecology/gynecologic-oncology/conditions-we-treat/endometrial-and-uterine-cancer

[2] L. A. Torre et al., "Global cancer statistics, 2012," CA, vol. 65, no. 2, pp. 87–108, 2015.

[3] P. Morice, A. Leary, C. Creutzberg, N. Abu-Rustum, and E. Darai, "En dometrial cancer," Lancet, vol. 387, no. 10023, pp. 1094–1108, 2016.

[4] Endometrial Cancer - https://www.nature.com/articles/s41572-021-00324-8

[5] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," CA., vol. 68, no. 1, pp. 7–30, 2018.

[6] K. A. Nicolaije et al., "Follow-up practice in endometrial cancer and the association with patient and hospital characteristics: A study from the population-based profiles registry," Gynecol. Oncol., vol. 129, no. 2, pp. 324–331, 2013.

[7] Computer-Aided Diagnosis in Histopathological Images of the Endometrium Using a Convolutional Neural Network and Attention Mechanisms: Hao Sun, Xianxu Zeng, Tao Xu, Gang Peng, and Yutao Ma.

[8] Class-Aware Image Search for Interpretable Cancer Identification: ARASH EBRAHIMIAN, HOSSEIN MOHAMMADI, MORTEZA BABAIE, NIMA MAFTOON, AND H. R. TIZHOOSH.

[9] Fine-Tuning and training of densenet for histopathology image representation using TCGA diagnostic slides - https://www.sciencedirect.com/science/article/pii/S1361841521000785

[10] Classification of lung cancer histology images using patch-level summary statistics: Simon Graham, Muhammad Shaban, Talha Qaiser, Navid Alemi Koohbanani, Syed Ali Khurram, et al. https://eprints.whiterose.ac.uk/150422/7/SPIE%201058119%20%281%29.pdf

[11] Convolutional Neural Network based Breast Cancer Classification - https://tkhan11.github.io/blog/cnnbreastcancerclf.html

[12] Convolutional Neural Network for Breast Cancer Classification - https://towardsdatascience.com/convolutional-neural-network-for-breast-cancer-classification-52f1213dcc9

[13] InceptionV3 - https://paperswithcode.com/method/inception-v3