**Breast Cancer Classification**

**Project Overview:**

This project aims to leverage deep learning techniques for the accurate detection of cancerous cells in histopathological images. Histopathological analysis plays a crucial role in diagnosing cancer, but manual evaluation by pathologists can be time-consuming and prone to human error. By developing an automated system, this project seeks to assist medical professionals in making faster and more reliable diagnoses.

The dataset used for this project is the Breast Cancer Histopathological Dataset, sourced from Kaggle. It contains labeled histopathological images of cancerous and non-cancerous tissues. The dataset has been preprocessed and augmented to ensure a balanced and diverse representation of images for training, validation, and testing.

The project employs Convolutional Neural Networks (CNNs), which are widely used for image analysis tasks. Additionally, transfer learning is explored by fine-tuning pre-trained models like ResNet50 and VGG16 to enhance performance. To improve the interpretability of predictions, visualization techniques such as Grad-CAM are implemented to highlight cancerous regions in the images.

The ultimate goal of this project is to develop a deep learning model that not only detects cancerous cells with high accuracy but also provides visual explanations of its predictions. This tool could serve as a valuable resource for pathologists, improving the accuracy and efficiency of cancer diagnosis.



**Objectives**

The objectives of this project, **Cancer Detection Using Histopathological Images**, are as follows:

1. **Automate Cancer Detection:**  
   Develop a deep learning-based system to accurately detect cancerous cells in histopathological images, reducing the reliance on manual analysis.
2. **Dataset Utilization and Augmentation:**  
   Utilize the **Breast Cancer Histopathological Dataset** from Kaggle and apply data augmentation techniques to balance the dataset and improve the generalization of the model.
3. **Build and Optimize CNN Models:**  
   Implement Convolutional Neural Networks (CNNs) for image classification, leveraging their capability to extract complex patterns from histopathological images.
4. **Integrate Transfer Learning:**  
   Explore and fine-tune pre-trained models like ResNet50 and VGG16 to enhance performance and efficiency, minimizing the computational cost of training from scratch.
5. **Visualize Predictions:**  
   Use techniques like Grad-CAM (Gradient-weighted Class Activation Mapping) or image segmentation to highlight cancerous regions, improving the interpretability of model predictions.
6. **Aid Medical Diagnosis:**  
   Create a reliable and efficient tool to assist pathologists and medical professionals in diagnosing cancer, aiming to save time and improve diagnostic accuracy.

**Dataset:**

Dataset is used from the Kaggle**(**[**https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images**](https://www.kaggle.com/datasets/paultimothymooney/breast-histopathology-images)**)**

**Content**

The original dataset consisted of 162 whole mount slide images of Breast Cancer (BCa) specimens scanned at 40x. From that, 277,524 patches of size 50 x 50 were extracted (198,738 IDC negative and 78,786 IDC positive). Each patch’s file name is of the format: u\_xX\_yY\_classC.png — > example 10253\_idx5\_x1351\_y1101\_class0.png . Where u is the patient ID (10253\_idx5), X is the x-coordinate of where this patch was cropped from, Y is the y-coordinate of where this patch was cropped from, and C indicates the class where 0 is non-IDC and 1 is IDC.

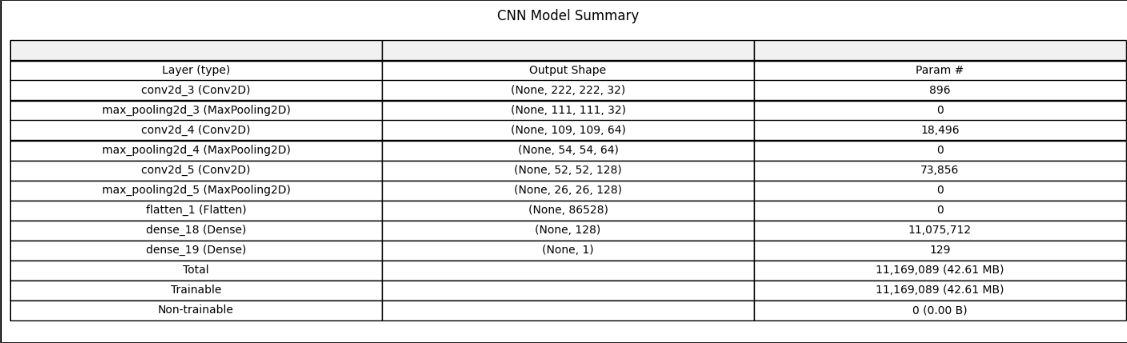
**Libraries**

**What libraries that we used in this project actually.**

* Pandas Ipython
* Numpy
* Sklearn
* Matplotlib
* Seaborn
* Shutil
* Random
* Os
* Tensorflow

**Models**

* **Convolutional Neural Network (CNN)**: Implement a custom CNN using TensorFlow or PyTorch for detecting cancerous cells.
* **Pretrained Models for Transfer Learning**:
* **ResNet** (e.g., ResNet50)
* **VGG16**



**Environment:**

Google colab

**Project Access**

**Google colab link :**

<https://colab.research.google.com/drive/1yK_vwA49riEFmUT55kwsw1K7Vl3TRluA?usp=sharing>