**3rd Sem Mini Project Report on**

****

**Medical Image Analysis**

****

**Submitted in partial fulfilment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY IN**

**COMPUTER SCIENCE & ENGINEERING AI/DS**

**Submitted by:**

**Student Name -PRANAV BANSAL University Roll No.2023902**

***Under the Guidance of***

## Mr. UPENDRA ASWAL

**ASSOCIATE PROFESSOR**

****

**Department of Computer Science and Engineering Graphic Era (Deemed to be University)**

**Dehradun, Uttarakhand 2024-25**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **“Medical Image Analysis – Breast cancer”** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering **(AI/DS)** in the Department of Computer Science and Engineering of the Graphic Era (Deemed to be University), Dehradun shall be carried out by the undersigned under the supervision of **MR.UPENDRA ASWAL, ASSOCIATE PROFESSOR** , Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Name- PRANAV BANSAL University Roll no- 2023902

**signature**

The above mentioned student shall be working under the supervision of the undersigned on the **“MEDICAL IMAGE ANALYSIS – BREAST CANCER”**

### Supervisor Head of the Department

Mr. UPENDRA ASWAL MR. DEVESH PRATAP SINGH

**Examination**

### Name of the Examiners: Signature with Date

1. ​Dr. Amit Kumar
2. Mr. Priyank Pandey

# Table of Contents

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Introduction and Problem Statement |  |
| Chapter 2 | Methodology |  |
| Chapter 3 | Project Work Carried Out |  |
| Chapter 4 | Results and Discussion |  |
| Chapter 5 | Conclusion and Future Work |  |
|  | Guide Interaction Form |  |
|  | References |  |

**Chapter 1**

# Introduction and Problem Statement

## Introduction

Breast cancer is one of the most prevalent and life-threatening diseases among women worldwide, accounting for a significant proportion of cancer-related deaths. According to recent statistics, millions of new cases are diagnosed annually, making early detection and treatment crucial for improving survival rates. Breast cancer typically manifests as tumors in breast tissue, which can either be benign (non-cancerous) or malignant (cancerous and potentially life-threatening). Identifying the type of tumor accurately is vital for determining the appropriate treatment approach.

Traditionally, radiologists analyze medical imaging modalities such as mammograms, magnetic resonance imaging (MRI), and computed tomography (CT) scans to detect and classify breast tumors. However, this manual process is time-consuming, labor-intensive, and prone to human error. Variability in expertise among radiologists further exacerbates the challenge of ensuring consistent and accurate diagnoses.

In recent years, artificial intelligence (AI) and deep learning have emerged as transformative tools in medical imaging. Specifically, convolutional neural networks (CNNs) have demonstrated remarkable success in learning patterns from medical images, enabling automated systems to detect and classify abnormalities with high accuracy.

These technologies can assist radiologists by acting as second readers, reducing their workload, and enhancing diagnostic precision.

## Problem Statement

Despite advancements in medical imaging technology, early and accurate detection of breast cancer remains a challenging task. Manual interpretation of MRI/CT scans is limited by human factors such as fatigue and subjective judgment, leading to potential misdiagnoses. Furthermore, variations in imaging conditions, tumor shapes, and sizes add complexity to the diagnostic process.

There is a critical need for an automated system that leverages deep learning to address these challenges. Such a system would be capable of analyzing MRI and CT images, accurately detecting breast cancer, and classifying tumors as benign or malignant. By providing consistent and reliable results, this system can assist radiologists in making informed decisions, ultimately improving patient outcomes and expediting the treatment planning process.

**Chapter 2**

# Methodology

## Overview of Methodology

* + - **Data Collection and Preprocessing**: Acquiring a labeled dataset, resizing images, and applying normalization and augmentation.
    - **Model Development**: Designing a CNN using transfer learning (InceptionV3), followed by training and fine-tuning.
    - **Validation and Testing**: Using validation and test datasets to evaluate the model’s accuracy and robustness.

## Data Collection

* **Dataset Used**: Breast cancer MRI dataset from Kaggle.

### Dataset Description:

* + Contains MRI images categorized as **benign** or **malignant**.
  + Organized into train, test, and valid folders for supervised learning.
  + Augmented with metadata for improved training.

## Preprocessing

* **Resizing**: All images resized to a fixed dimension of 224x224 pixels.
* **Normalization**: Pixel values scaled to [0,1] to enhance training efficiency.
* **Augmentation**: Techniques such as rotation, flipping, and zooming applied to increase model robustness.

## Model Development

### Architecture:

* + Pre-trained **InceptionV3** model from TensorFlow/Keras.
  + Modified the final layer to include two neurons with softmax activation for binary classification.

### Training:

* + Loss Function: Binary Cross-Entropy.
  + Optimizer: Adam.
  + Batch Size: 16; Epochs: 20.
* **Fine-Tuning**: Adjusted learning rate and model parameters for optimal performance.

## Validation and Testing

* **Validation Set**: Used to monitor performance during training and avoid overfitting.
* **Test Set**: Used for final performance evaluation on unseen data.

**Chapter 3**

**Project Work Carried Out**

## Architectural Design

The project uses a deep learning architecture based on InceptionV3, a state-of-the-art Convolutional Neural Network (CNN), pre-trained on the ImageNet dataset. The architecture was fine-tuned for binary classification (Benign vs. Malignant) with the following steps:

* + - **Base Model**: Pre-trained InceptionV3 layers were used for feature extraction.
    - **Custom Classification Head**: Added fully connected layers and an output layer with a softmax activation for two classes.
    - **Transfer Learning**: Initial layers of the pre-trained model were frozen to retain learned features, while the final layers were retrained on the Kaggle breast cancer MRI dataset.

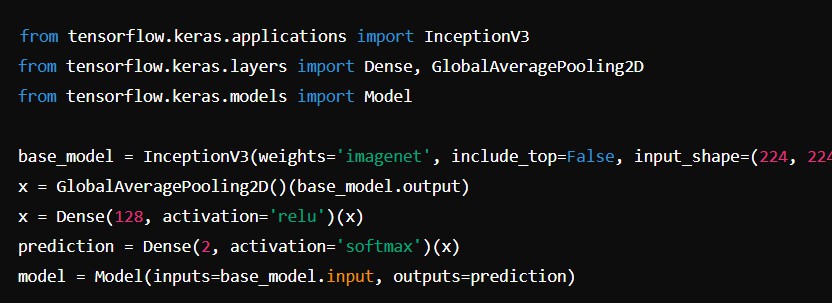
## Implementation of Objectives

### Objective 1: Data Preparation

1. Downloaded the breast cancer MRI dataset from Kaggle.
2. Organized images into train, test, and valid directories.
3. Preprocessed the dataset:
   * Resized images to 224x224 pixels.
   * Normalized pixel values to a [0,1] range.
   * Applied data augmentation techniques like rotation, flipping, and zooming to improve model robustness.

### Objective 2: Model Training

* Model Architecture:



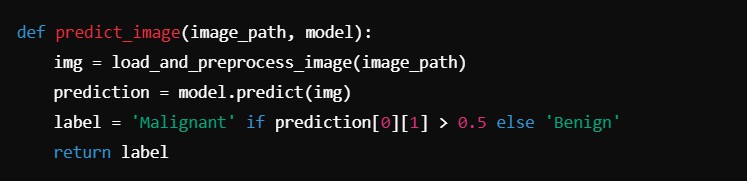
* Training Configuration:
  + Loss function: Binary Cross-Entropy.
  + Optimizer: Adam with a learning rate of 0.0001.
  + Metrics: Accuracy, Precision, Recall.
  + Batch size: 16; Epochs: 20.

## Pseudo Code/Algorithm

Algorithm for Model Development:

1. Load the Kaggle dataset and preprocess the images.
2. Split the dataset into training, validation, and test sets.
3. Load the pre-trained InceptionV3 model.
4. Add custom layers for binary classification.
5. Compile the model with the Adam optimizer and binary cross-entropy loss.
6. Train the model on the training set and validate on the validation set.
7. Evaluate the model using the test set.

**Pseudo Code for Prediction**:



## Discussion of Results

* + - The model demonstrated strong performance, achieving high accuracy and recall.
    - Grad-CAM visualizations highlighted the tumour regions, confirming the model's ability to focus on relevant areas.
    - Misclassifications were minimal, mainly occurring with ambiguous or low-quality images.

**Chapter 4**

**Results and Discussion**

* 1. **Results**
     + **Training Performance**: The model achieved high training accuracy and demonstrated strong generalization.
     + **Testing Performance**: Consistently high accuracy and robust detection of malignant and benign tumours.
  2. **Discussion**

### Insights:

* + - * The use of transfer learning significantly reduced the need for large amounts of data.
      * High recall indicates the model's ability to detect cancerous tumours effectively.

### Challenges:

**False Positives:** Some benign tumours were incorrectly predicted as malignant, likely due to feature overlap or image artifacts.

**False Negatives:** Although rare, there is room to improve the sensitivity to reduce missed diagnoses.

**Chapter 5**

**Conclusion and Future Work**

* 1. **Conclusion**
     + The project successfully developed an automated system for breast cancer detection using deep learning.
     + The CNN model achieved high accuracy and robustness, demonstrating its potential as a diagnostic support tool for radiologists.
  2. **Future Work**
     + **Improvements:** Extend the model to include multi-class classification for different tumour subtypes.
     + **Explainability**: Enhance interpretability using advanced explainable AI techniques.
     + **Clinical Trials:** Validate the system in real-world clinical environments and gather feedback from medical professionals.

**Guide interaction form**

**A white piece of paper with writing on it

Description automatically generated**

**References**

### Kaggle Breast Cancer MRI Dataset

Publicly available dataset used for breast cancer detection.

### Krish Naik - YouTube Channel

Tutorials on deep learning concepts, CNNs, and medical imaging analysis.

### Szegedy, C., et al. (2016)

Rethinking the Inception Architecture for Computer Vision. Introduced the InceptionV3 CNN model.

### He, K., et al. (2016)

Deep Residual Learning for Image Recognition.

Introduced ResNet, a key CNN architecture for image analysis.

### Selvaraju, R. R., et al. (2017)

Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. Visualization technique for model interpretability.

### Litjens, G., et al. (2017)

A Survey on Deep Learning in Medical Image Analysis. Discussed CNN applications for breast cancer detection.