

**Deep Learning**

***Breast Cancer Detection***

***Team 4 - Pink Ribboners***

*Venkata Naga Sai Sri Vaishnavi Appala,*

*Sai Sruthi Avula*

*Namratha Bolar*

*Luna Yang*

**School of Graduate Professional Studies**

Data Analytics

DAAN 570 – Deep Learning

*Spring Semester, 2025*

**TABLE OF CONTENTS**

**Introduction....................................................................................................................................2**

**Problem Statement........................................................................................................................2**

**Challenges.......................................................................................................................................2**

**Related Works…………................................................................................................................3**

**Data Description............................................................................................................................6**

**Deep Learning Process................................................................................................................14**

**Expected Results..........................................................................................................................19**

**References To Date Sources........................................................................................................20**

**Justification For Existing Code..................................................................................................21**

# Introduction

# In the United States, approximately 1 in 8 women will be diagnosed with invasive breast cancer in their lifetime (National Breast Cancer Foundation). According to the National Breast Cancer Foundation, an estimated 310,720 women and 2,800 men are expected to be diagnosed with invasive breast cancer in 2024. Breast cancer remains a leading cause of death among women worldwide. Beyond contributing to high mortality rates, it also imposes significant psychological and emotional burdens on those undergoing treatments. However, advancements in early detection and cancer therapies have significantly improved the outcomes, leading to a remarkable 42% decline in breast cancer-related deaths among women between 1989 and 2021 (National Breast Cancer Foundation). Notably, deep learning technologies can be designed to analyze key factors influencing cancer detection and treatment, enhancing early-stage diagnosis and intervention (Nassif et al., 2022). This, in turn, improves treatment efficacy and increases patients’ chances of survival.

# Problem Statement

# Our objective is to develop a deep learning model for breast cancer detection with an accuracy rate that meets or exceeds the results of related studies. As highlighted earlier, achieving highly accurate early detection can significantly reduce mortality rates, enable more effective treatment, and minimize unnecessary overtreatment which helps alleviate the psychological and emotional burden on patients. Additionally, if genetic sequencing data is available, our project will compare the accuracy of deep learning models in analysing both genetic sequencing data and imaging data across different stages of cancer development to determine which approach is more suitable for early-stage detection.

# Challenges

# Researchers commonly develop deep learning models to analyze tumor characteristics or protein status/types for breast cancer detection using either medical imaging or genetic sequencing data. According to the systematic literature review, both approaches have demonstrated promising results in previous studies (Nassif et al., 2022). While genetic sequencing offers higher accuracy and reliability, access to relevant datasets remains a significant challenge. For our project, we have submitted a request to obtain circulating tumor DNA data for breast cancer detection and are currently awaiting approval.

# Label noise presents a significant challenge for our deep learning models. As breast cancer progresses, it spreads across different regions of breast tissue, leading to varying cancer stages within the same sample. Consequently, a single image may contain multiple sub-category annotations, creating inconsistencies that complicate multi-class classification. To address this issue, we can leverage attention-based multi-instance learning (MIL) to focus on the most relevant regions, reducing confusion in classification. Additionally, Bayesian Neural Networks (BNNs) and Monte Carlo Dropout (MC Dropout) can aid in quantifying uncertainty in difficult-to-classify regions, improving model reliability. Furthermore, instead of analyzing each image independently, Graph Neural Networks (GNNs) can be employed to capture spatial relationships between different regions, enhancing classification accuracy and robustness.

## RELATED WORKS

**Shen et al., (2019)** aimed to develop a deep learning algorithm for breast cancer detection on screening mammograms using an end-to-end training approach that eliminates the reliance of the rarely available lesion annotations. The challenge is that very few publicly available mammography databases contain fully annotated lesion data, making it difficult to accurately train deep learning-based diagnostic models. Its key contribution is to propose an automatic deep learning model that can be trained to obtain high accuracy rates with data from multiple mammography platforms. This end-to-end approach improves false positive and false negative screening mammography results by leveraging a hybrid architecture of ResNet and VGG16 networks. Unlike previous studies, this “end-to-end” training approach is quite novel, because it only requires complete clinical lesion annotations in the initial training stage, and subsequent stages require only image-level labels, and hence eliminating the reliance on rarely available lesion annotations. The model achieved a high AUC of 0.91 on the Digital Database for Screening Mammography (CBIS-DDSM) and AUC of 0.98 on the INBREAST dataset. The strength of this study includes high accuracy, innovative architecture (Resnet-VGG16) instead of CNN, and adaptable to various mammography platforms, and end-to-end training approach. The weakness includes large GPU requirements (NVIDIA 8 GB Quadro M4000 GPU card), lack of dataset diversity (only two datasets trained and tested), and the patch-based approach requiring high computation power.

Its application to our research is that we can only retrieve very limited datasets from various MRI breast cancer detection databases (e.g., MIT and Duke MRI Breast Cancer databases). This novel training approach will enable us to develop deep learning models adaptable to various imaging databases. Unlike conventional CNN-based deep learning models requiring full lesion annotations for training, this new architecture (Resnet-VGG16) will enable better feature extraction and generalization and eliminate our dependency on rarely available lesion annotations.

**Iniyan et al., (2024)** aimed to improve breast cancer diagnosis by integrating multi-modality medical imaging with deep learning feature fusion and transfer learning techniques. The study proposes a Computer Vision with Fusion-Based Joint Transfer Learning for Breast Cancer Diagnosis (CVFBJTL-BCD) model to improve breast cancer detection accuracy by combining deep learning models (e.g., DenseNet201, InceptionV3, MobileNetV2) for feature extraction and fusion. They optimized classification using Stacked Autoencoders (SAE) and a Horse Herd Optimization Algorithm (HHOA) to fine-tune parameters. They contributed to breast cancer detection literature by several ways: 1) they leveraged multi modalities from ultrasound and Histopathology databases with different resolutions, contrast, and texture, affecting diagnostic accuracy; 2) they are able to integrate DenseNet201, InceptionV3, and MobileNetV2 to extract complementary features from different modalities with limited samples to deliver a comprehensive cancer diagnosis. This study contributes by using Gabor filtering to reduce noises, combining three deep learning models (DenseNet201, InceptionV3, MobileNetV2) into a single joint transfer learning pipeline, using stacked autoencoder for classification to improve accuracy by learning hierarchical patterns, and using Horse Herd Optimization algorithm as the optimizer. Its strength is that it handles small data very well and balances specificity and sensitivity. Its only weakness is high-performing GPU.

Due to data scarcity, we may use their deep learning architectures (e.g., DenseNet201, InceptionV3, and MobileNetV2) to extract complementary features from different imaging modalities, such as ultrasound, histopathology, MRI, and histopathology, for a more comprehensive breast cancer diagnosis. We can validate their deep learning model on MRI or mammography databases. We can also test their suggested Gabor filtering, Stacked autoencoders for classification, and Horse Herd Optimization Algorithm for fine-tuning hyperparameters.

**Almarri et al.’s (2024)** research study titled *The BCPM Method: Decoding Breast Cancer with Machine Learning* aims to develop and implement a Breast Cancer Predictive Model (BCPM) using machine learning algorithms to enhance early-stage breast cancer detection and diagnosis. By leveraging advanced computational techniques, the study seeks to improve prediction accuracy, thereby facilitating personalized treatment plans and improving patient outcomes. Early detection is critical in breast cancer management, and the BCPM model is designed to refine diagnostic capabilities through data-driven insights. However, the study faces several scientific challenges. One major hurdle is data availability and quality, as obtaining large, well-annotated datasets remains difficult due to inconsistencies in diagnostic factors and regional disparities in medical records. Additionally, high computational infrastructure requirements pose a challenge, as processing large datasets and training deep learning models require significant resources. Lastly, model optimization is a crucial aspect that demands iterative refinements, including adjustments to network architectures and hyperparameter tuning, to achieve optimal performance.

The key contributions of the research include the development of a multi-source data framework, integrating clinical records, health datasets, and medical imaging data to improve the model’s robustness. The study also emphasizes data preprocessing techniques to handle missing values, inconsistencies, and feature selection. Various machine learning models, including Random Forest, K-Nearest Neighbors (KNN), and Support Vector Machines (SVM), are implemented using K-Fold Cross-Validation to enhance accuracy. Additionally, the study promotes personalized treatment strategies, ensuring that predictive insights contribute to individualized recovery plans based on specific prognostic factors.

Unlike prior studies that rely on a single generalized model, this research integrates multi-modal data sources (clinical and imaging) and explores multiple machine learning algorithms, offering a more comprehensive approach. By identifying previously unknown prognostic factors, the BCPM model refines diagnosis and recommends tailored treatment pathways. However, the study acknowledges limitations, such as data accessibility challenges and the need for high-cost computational resources. Despite these constraints, the model demonstrates high accuracy and reinforces the potential of machine learning in personalized medicine and early cancer detection.

This research is highly relevant to our project, which focuses on early cancer detection using bloodstream cell images. The insights gained from BCPM highlight the importance of multi-modal data integration, model optimization techniques, and feature selection to improve predictive performance. By applying ensemble learning and fine-tuning hyperparameters, we aim to develop a highly accurate and generalizable cancer detection model. The study reinforces key lessons, including the significance of combining clinical and imaging datasets and adopting rigorous optimization strategies to mitigate overfitting and improve diagnostic precision.

**Zhang and the colleages’ (2023)** research paper Machine Learning and AI in Cancer Prognosis, Prediction, and Treatment Selection: A Critical Approach explores the role of artificial intelligence (AI) and machine learning (ML) in cancer diagnosis, prognosis, and treatment strategies. The study aims to evaluate the effectiveness of AI-driven models in improving decision-making in oncology while addressing their current limitations and future potential. Unlike research focused solely on predictive modeling, this paper provides a broader perspective on AI integration in clinical settings, examining how data-driven approaches enhance patient survival rate predictions and treatment planning. It also highlights the ethical and technical concerns involved in deploying AI-based solutions, emphasizing the need for transparency and interpretability in decision-making.

One of the main challenges discussed in the paper is data privacy, as large, high-quality datasets are necessary for effective AI models while ensuring patient confidentiality. Another issue is computational complexity, requiring substantial infrastructure to process vast amounts of medical data. The study also acknowledges that AI in cancer research is still in its early developmental stages, with only a few real-world applications currently in practice. Additionally, the complexity of multi-modal medical data presents difficulties in model interpretability and integration. Despite these obstacles, the paper emphasizes how AI and ML models have the potential to refine cancer diagnostics by incorporating techniques such as DNA methylation analysis and Artificial Neural Networks (ANNs) to enhance predictive accuracy and treatment selection.

This research contributes significantly by examining how AI-driven solutions can address bias, imbalanced datasets, and real-world constraints in cancer research. Instead of focusing on a single predictive model, it provides a comprehensive view of AI’s role in various stages of cancer treatment, distinguishing between machine learning and deep learning methodologies. The study also offers insights into future AI advancements in oncology, detailing how robust sampling techniques and improved model transparency can lead to better healthcare outcomes. These insights are highly relevant to our project, as they reinforce the importance of integrating multi-source medical data, optimizing model performance, and ensuring ethical AI practices in cancer detection and prognosis modeling.

The paper of A Deep CNN Technique for Detection of Breast Cancer Using Histopathology Images by **Wadhwa and Kaur (2020)** proposes a deep learning-based Computer-Aided Diagnosis (CAD) system using the DenseNet-201 model to classify breast cancer histopathology images into benign or malignant categories. Using the BreakHis dataset, which comprises 9,109 images across multiple magnification levels, the model achieves a classification accuracy of 95.58%, surpassing previous approaches. The study highlights key challenges such as high intra-class variability and limited training data, addressing them through dense feature extraction and dropout layers to mitigate overfitting. The framework enhances diagnostic efficiency, reducing the workload for pathologists while ensuring reliable cancer detection with high precision (80%) and recall (99%).

This research is particularly relevant to our breast cancer prediction project, as it demonstrates the effectiveness of DenseNet-201 in feature extraction and classification. The study's approach to handling data limitations through feature reuse and dropout mechanisms provides valuable insights for optimizing model performance. Additionally, the results underscore the importance of leveraging deep CNN architectures to improve early cancer detection, aligning well with your project's objective of analyzing image data for precise cancer diagnosis.

The study by **Melek et al.** introduces a spatiotemporal deep learning model for breast cancer risk prediction by integrating mammograms from different time points using a Siamese neural network. Traditional risk models relying on clinical factors often lack consistency, whereas deep learning approaches have shown promise in capturing subtle imaging cues. The proposed model improves risk prediction by leveraging both spatial and temporal data, achieving an AUC of 0.81, outperforming CNN-based models trained on single time-point images. Despite dataset limitations (191 patients, 61 cancer cases), the model demonstrates superior performance, particularly in dense breast imaging and younger patients, where malignancy features are more dynamic.

This research is highly relevant to our project, as the Siamese network's ability to integrate sequential imaging data could be explored for analyzing bloodstream image sequences. Additionally, the study’s strategies for handling small datasets, such as stratified cross-validation and augmentation, provide valuable insights into improving generalizability, which is crucial for your deep learning model.

# Data DESCRIPTION

The **Curated Breast Imaging Subset of the Digital Database for Screening Mammography (CBIS-DDSM)** is a comprehensive, publicly accessible dataset designed to aid in the development and evaluation of computer-aided detection and diagnosis systems for breast cancer. This dataset is an enhanced and standardized version of the original DDSM, particularly defined for for the research and clinical needs.

**Data Collection**

The original DDSM comprises 2,620 scanned film mammography studies, encompassing normal, benign, and malignant cases with confirmed pathology information, the CBIS-DDSM was curated by experienced mammographers who meticulously selected and annotated a subset of the DDSM data. This curation process involved decompressing the images, converting them into the Digital Imaging and Communications in Medicine (DICOM) format, and providing updated region of interest (ROI) segmentations and bounding boxes along with the masses and calcifications data of each patient and the related image path to it. The dataset also includes detailed pathologic diagnoses for the training data, ensuring its utility for machine learning applications.

**Key Components of CBIS-DDSM:**

* Mammograms are stored in DICOM format.
* Metadata includes lesion type, assessment, subtlety and pathology status.
* Region of Interest(ROI) masks indicates areas of abnormalities.

**Relevance to Goal**

Our goal mainly focuses on early-stage breast cancer detection using deeep learning. CBIS-DDSM is an ideal dataset as it provides:

* Labeled mammograms (benign vs. malignant cases), crucial for supervised learning.
* Detailed ROI masks, allowing the application of CNN-based segmentation models.
* Metadata features that can support multi-modal analysis (image + tabular data).

**Exploratory Data Analysis (EDA)**

The brief description of features in the dataset is as follows,

|  |  |  |
| --- | --- | --- |
| **Feature Name** | **Description** | **Type** |
| patient\_id | Unique identifier for the patient | Categorical |
| breast\_density | Density of the breast tissue (1-4) | Categorical |
| left\_or\_right\_breast | Laterality of the breast (Left/Right) | Categorical |
| image\_view | Mammogram view type (CC/MLO) | Categorical |
| abnormality\_id | Unique identifier for detected abnormalities | Categorical |
| abnormality\_type | Type of abnormality (Mass, Calcification) | Categorical |
| mass\_shape | Shape of the detected mass | Categorical |
| mass\_margins | Border properties of the mass | Categorical |
| assessment | BI-RADS assessment category (1-5) | Ordinal |
| pathology | Benign or Malignant label | Binary |
| subtlety | Level of abnormality visibility (1-5) | Ordinal |
| image\_file\_path | File path to the original mammogram | Categorical |
| cropped\_image\_file\_path | File path to the cropped mammogram | Categorical |
| ROI\_mask\_file\_path | File path to the segmented region of interest (ROI) | Categorical |

A screenshot of a computer screen

AI-generated content may be incorrect.

**Interpretation & Explanation:**

* The dataset contains both image and metadata, allowing multi-modal analysis.
* Benign vs. Malignant distribution should be balanced for training models effectively.
* Breast density, mass shape, and assessment scores are the important predictors for classification models.
* Feature correlation helps identify dependencies, e.g., pathology is strongly correlated with assessment scores.

**Dataset Preparation**

The CBIS-DDSM dataset is a refined and curated version of DDSM, making it more suitable for machine learning and deep learning applications in breast cancer detection to ensure:

* Higher quality and accurate labeling.
* Standardized image format (DICOM) for medical imaging.
* Structured metadata for easy integration into AI models.

The provided diagram illustrates the process of preparing the CBIS-DDSM (Curated Breast Imaging Subset of DDSM) dataset, which is derived from the DDSM (Digital Database for Screening Mammography).

A screenshot of a computer

AI-generated content may be incorrect.  
The explanation of flow diagram is as follows:

1. **DDSM Data Input:** The process begins with raw DDSM data, which contains mammography images and related metadata.
2. **Images Decompression:** Next, the mammography images in the dataset are decompressed to a usable format.
3. **Cases Filtered by Mammographers:** Medical experts (mammographers) review the dataset and filter cases based on clinical relevance and quality.
4. **Image Reannotation:** A subset of 118 mammograms undergoes reannotation by mammographers to ensure accuracy in diagnosis and labeling.
5. **Metadata Extraction:** Metadata is extracted from .ics and .OVERLAY files that contain information about the images and patient data.
6. **ROI (Region of Interest) Extraction:** ROI outlines (regions in the mammograms containing masses or calcifications) are extracted from .OVERLAY files.
7. **Automated Mass Segmentation:** The extracted regions of interest undergo mass segmentation, where an algorithm identifies and outlines tumors or calcifications in the images.
8. **Final CBIS-DDSM Dataset:** The processed dataset is saved as CBIS-DDSM, containing, mammograms saved in DICOM format, metadata about masses & calcifications in CSV files,and binary DICOM images containing mass & calcification outlines.

**Dataset Composition**

The CBIS-DDSM dataset comprises 6,775 studies from 6,671 subjects, totaling 10,239 images These images are categorized into three primary conditions: normal, benign, and malignant. Specifically, the dataset includes 753 cases of calcifications and 891 cases of masses, all encapsulated in the industry-standard DICOM format. Each case is divided into three Full\_Mammogram, Cropped Images, and ROI Images.

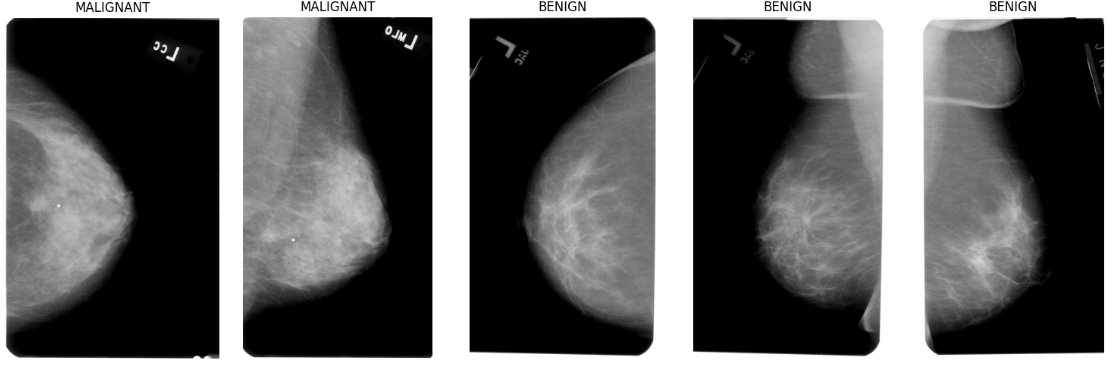


  Figure 1. Full Mammogram Images with Conditions Benign, Malignant.

A close-up of several images

AI-generated content may be incorrect.

Figure 2. Cropped Images with Conditions Benign, Malignant.

A black and white image of a person

AI-generated content may be incorrect.

Figure 3. ROI Mask- Images with Conditions Benign, Malignant.

**Data Preprocessing**

Three different types of image preprocessing techniques were applied to the CBIS-DDSM mammography dataset to improve model performance and enhance breast cancer detection accuracy. The **original images, negative images, and AHE (Adaptive Histogram Equalization)** images were used to train and evaluate the deep learning model.

**Original Images (Baseline Dataset)**

These are the unaltered, raw mammogram images obtained from the dataset in their original grayscale format. Provides a baseline performance for the deep learning model.Allows comparison between enhanced vs. unenhanced images. Ensures that any improvements from enhancement techniques can be quantified. This may contain uneven lighting, noise, or artifacts that can affect important features. Some tumor regions may have low contrast, making them harder to detect.

**Negative Images (Inverted Intensity Transformation)**

This transformation inverts pixel intensities in the mammogram images, making dark regions appear bright and vice versa to overcome the effects of the original images for image enhancement to identify the features reducing the noise. Enhances the visibility of tumor structures by reversing contrast. Some deep learning models may learn different feature representations when applying the negative transformation. Requires model adaptation to interpret inverted features correctly.

**AHE Images (Adaptive Histogram Equalization)**

AHE is an advanced contrast enhancement technique that improves the visibility of small details by adjusting the local contrast in different parts of the image. It divides the image into small local regions (tiles). Enhances contrast individually in each region rather than applying a global contrast change. Helps in making subtle tissue variations more distinct.It mproves contrast in dense breast tissue, and is useful for enhancing calcifications and microstructures in mammograms and reduce the impact of uneven lighting in images.

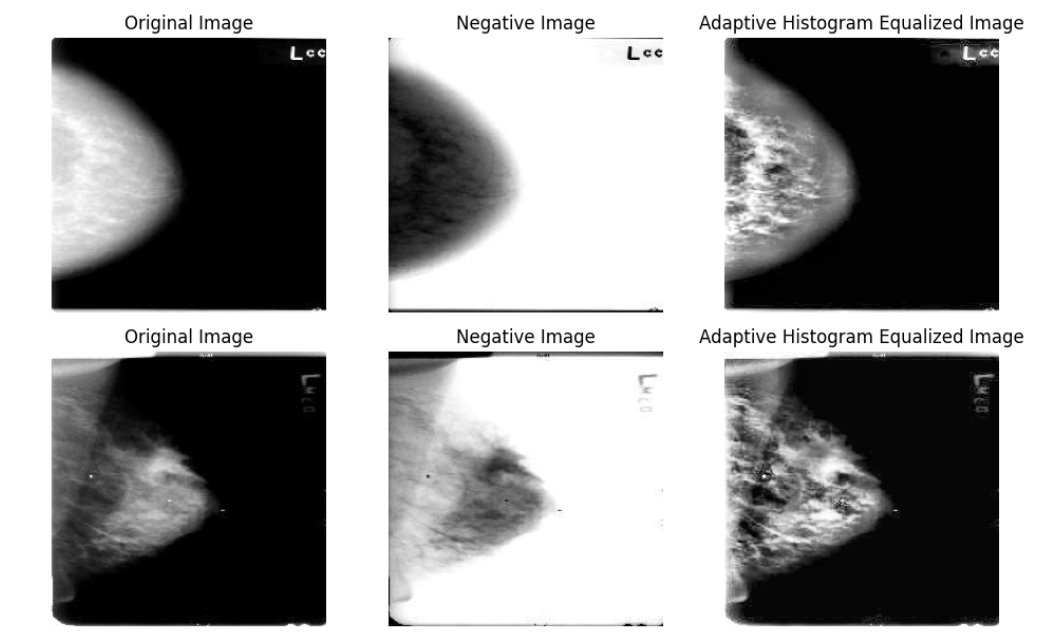


Figure 4. Sample Images ( Original Image, Negative Image, AHE Image)

**Advanced Image Transformation Technique**

**HOG (Histogram of Oriented Gradients)**

It is a feature extraction technique widely used in image processing and helps to capture important edge and shape features by analyzing the gradient directions of pixel intensities. The Gradient Computing Calculates the intensity changes in an image using Sobel filters horizontal and vertical edges. Divide the image into small regions (cells), typically 8×8 pixels. Within each cell, it creates a histogram representing gradient directions. The grouping of multiple cells is done into blocks (e.g., 2×2 cells) and normalizes their histograms which improves the robustness of changes.

The computed histograms are flattened into a feature vector, which serves as input for machine learning models.

A collage of images of symbols

AI-generated content may be incorrect.

A collage of images of symbols

AI-generated content may be incorrect.

A collage of images of a person&#39;s face

AI-generated content may be incorrect.

Figure 5. Sample Images (HOG Original Image, HOG Negative Image, HOG AHE Image)

# DEEP Learning process

Our approach for classifying the five types of leukocytes will start with using a convolutional neural network (CNN). CNNs have become the standard approach when it comes to image identification since the introduction in the 1980s[[1]](#endnote-1). A variety of CNN architectures dominate modern studies for cellular image identification and we intend to implement a CNN from scratch, as well as experiment with existing architectures such as AlexNet, ResNet and VGG. Alexnet could be used for larger datasets, as it allows for training larger datasets while reducing training time[[2]](#endnote-2) and could yield a higher accuracy when it comes to WBC identification. Another CNN architecture that could be used is VGG, which is based off Alexnet. A reason to use this method is that it has smaller convolution filters, which can provide a larger number of weight layers, yielding greater performance than a standard CNN or Alexnet architecture[[3]](#endnote-3). ResNet is a more recent development, with an architecture that adds shortcuts between layers, allowing greater depth without degradation[[4]](#endnote-4). Due to its success in much of the literature we reviewed, we anticipate our ResNet model will likely outperform our AlexNet model. To further enhance performance, transfer learning (which would further reduce training time by utilizing a model pre-trained on other images[[5]](#endnote-5)), batch normalization (which is a technique that stabilizes the learning process and reduces the training time for deep networks) and ensemble methods (which allow for the combination of models into a single classifier) could be pursued (if we can figure them out). We could try to incorporate feature selection as well, as reducing the number of redundancies in will likely improve training time[[6]](#endnote-6).

Our training/testing strategy will center around an 80/20 split. As we evaluate our models, we will make adjustments if needed. When will evaluate our models based on correct classification of each white blood cell type. We will evaluate the individual accuracy of each blood cell type and see how each model performs for the identification of the blood cell types. We will then evaluate the overall accuracy of the blood cell classification by taking the average of all five blood cell type accuracy metrics. Additional metrics such as sensitivity and specificity may also be considered.

A Convolutional Neural Network (CNN) is a type of deep learning model specifically designed for processing structured grid-like data, such as images. It is widely used in computer vision tasks like image classification, object detection, and medical image analysis.

Step 1: Training the data: The preprocessed HOG image which includes Original, Negative, and AHE was normalized to range of [0,1].

Step2: The images are resized to 224\*224 pixels to match the CNN input requirements and have used augmentation techniques on it such as Rotation, Flipping, scaling and Shearing which helped to increase dataset variability and prevent overfitting.

Step3: CNN Model Architecture:

It consists of multiple layers for feature extractions and classification. We have ReLU activation function introduced for non-linearity. CNN Layers:

Input Layer

* Accepts HOG-transformed images (224×224×1) as input.

Convolutional Layers (Feature Extraction)

* Conv2D(32, 3x3, ReLU) – Extracts basic edge features.
* Conv2D(64, 3x3, ReLU) – Learns deeper patterns.
* Conv2D(128, 3x3, ReLU) – Captures complex shapes and tumor structures.
* Conv2D(256, 3x3, ReLU) – Identifies high-level features.
* Conv2D(512, 3x3, ReLU) – Enhances fine-grained feature extraction.
* Batch Normalization applied after each layer to stabilize training.

Pooling Layer (Dimensionality Reduction)

* MaxPooling2D(2x2) applied after every few convolutional layers to reduce spatial dimensions while retaining key features.

Flatten Layer (Vector Transformation)

* Converts extracted feature maps into a 1D feature vector for classification.

Fully Connected (Dense) Layers

* Dense(256, ReLU) – Captures abstract representations.
* Dense(128, ReLU) – Reduces complexity and refines learned features.
* Dropout(0.5) – Prevents overfitting by randomly disabling neurons.
* Dense(2, Softmax) – Outputs final probability for benign vs. malignant classification.

Step 3 : Training the CNN model: The loss function used here is Categorical Crossentropy since we have a binary classification task involved (Benign vs Malignant). Adam Optimizers are used to ensure stable and fast convergence.

The major metrics used to evaluate are Accuracy, Precision, recall and F1 score.

The model was trained for **10 epochs** with a **batch size of 32**, ensuring a balance between computational efficiency and learning stability and the model did not show any overfitting.

To overcome all the problem with the CNN and for the accuracy, we have used the ResNet50 Model for deeper analysis.

Step 4: Model Validation

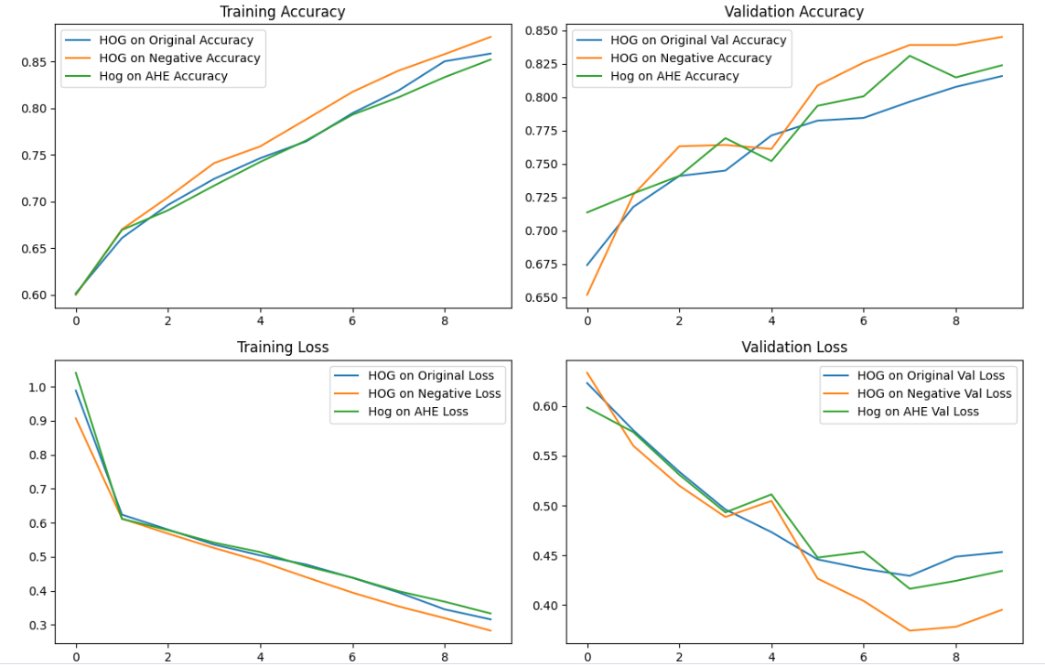
* Validation Accuracy: Measured after each epoch to track model improvement.
* Validation Loss: Analyzed to detect overfitting and fine-tune hyperparameters.

Results Observed:

* HOG Negative images outperformed other transformations, achieving the highest validation accuracy.
* HOG AHE images showed stable accuracy, indicating effective contrast enhancement.
* HOG Original images had slightly higher loss, possibly due to the presence of more noise.

Step 5: Model Testing

* After training and validation, the model was tested on the remaining 10% test data.
* Confusion matrix and classification report were generated to analyze performance.
* The test accuracy for each transformation was recorded:
  + HOG Negative: ~81.16%
  + HOG AHE: ~79.89%
  + HOG Original: ~80.48%
* Key Observation:  
  HOG Negative transformation provided the best classification results, likely due to its ability to enhance edge features and contrast differences.



The graph illustrate accuracy of training and validation with their respective losses over multiple epochs.

HOG Negative images consistently performed the best, achieving higher accuracy and lower loss. HOG AHE provided stable accuracy but showed slight validation loss fluctuations.

HOG Original had the lowest accuracy and highest loss, indicating the need for more preprocessing improvements. The CNN model successfully learned from all three transformations but displayed overfitting tendencies, necessitating the use of ResNet50 for improved generalization.

**ResNet50 (Residual Network with 50 layers)**

ResNet50 (Residual Network with 50 layers) is a deep convolutional neural network (CNN) designed to overcome the vanishing gradient problem using skip (residual) connections. It enables deep extraction of features, making it highly effective for medical image analysis.

Step 1: To improve breast cancer classification, the ResNet50 model was applied to the dataset containing Original, Negative, and AHE images, excluding HOG-transformed images due to overfitting issues observed in CNN training.

Step 2: Machine learning Process using ResNet50: Unlike the CNN model, ResNet50 leverages residual learning and pre-trained weights (from ImageNet) for improved feature extraction.

Step 3 : training the ResNet50 Model Categorical Crossentropy loss function used as its a binary classification with Adam learning rate 1e-4. The evaluation metrics used are accuracy, precision, recall and F1-score. The model was trained for 10 epochs with a batch size of 32, leveraging transfer learning with pretrained ImageNet weights.

Step 4: Model Validation

Validation Accuracy: Measured after each epoch.

Validation Loss: Monitored to detect potential overfitting.

Results Observed:

* Negative Images achieved the highest validation accuracy (~97%).
* AHE Images showed stable accuracy (~96%).
* Original Images had some fluctuations (~95%), indicating more sensitivity to noise. Step 5: Model Testing
* After training and validation, the model was tested on 10% unseen test data.
* The test accuracy for each preprocessing method was recorded:
  + Negative Images: ~97.84%
  + AHE Images: ~96.71%
  + Original Images: ~96.67%

Key Observation:

* Negative Images resulted in the highest classification accuracy, reinforcing that contrast inversion enhances feature differentiation.
* AHE Images provided slightly lower accuracy, potentially due to noise introduced by aggressive contrast enhancement.
* ResNet50 outperformed CNN by effectively generalizing across training and validation data.

A graph of different levels of performance

AI-generated content may be incorrect.

This figure presents the training and validation performance of the deep learning model across three different image preprocessing techniques. Training accuracy all three preprocessing methods show high training accuracy, especially with Negative and AHE images. Validation accuracy in terms of negative images consistently maintains the highest validation accuracy indicating the contrast whereas AHE images show a stable accuracy and low accuracy with Original Images. Training loss, all models exhibit a steady decline with negative and AHE images having the lowest loss indicating efficient learning. Validation loss, contains some fluctuations that indicate some degree of variance and potential overfitting.

# Expected Results

We anticipate at least one of our models will identify the classification of the pathology cases comparable accuracy to the previous studies that have achieved this task. We hope to get an accuracy of between 95.04% to 99.52% for the classification, which is the accuracy that the previous studies have achieved. Based on the advances that have been made since these studies have come out, we tried to get those results considering various processes and models, which was accomplished by the ResNet50 model. We were able to get excellent accuracy on two models that we applied CNN and ResNet50 and the feature extraction techniques worked quite well and did not lead to any overfitting conditions.

**Conclusion**

Breast cancer remains one of the most prevalent and life-threatening diseases worldwide, making early detection crucial for improving patient outcomes. Our project leverages deep learning techniques, specifically CNN architectures such as ResNet50, to classify breast cancer using imaging data from the CBIS-DDSM dataset. By incorporating advanced image preprocessing techniques, including histogram equalization, negative transformation, and feature extraction using HOG, we have aimed to enhance model performance and classification accuracy.

**CNN Model Conclusion**

✔ The CNN model was trained on HOG-transformed images (HOG Original, HOG Negative, and HOG AHE).  
✔ HOG Negative transformation produced the highest accuracy (81.16%), while HOG AHE achieved 79.89% test accuracy.  
✔ Overfitting issues were observed in HOG-transformed images, leading to high training accuracy but lower validation accuracy.  
✔ The model performed well but had limitations in deep feature extraction, requiring an alternative approach such as ResNet50.

2 ResNet50 Model Conclusion

✔ ResNet50 was applied to Original, Negative, and AHE images (excluding HOG images due to overfitting).  
✔ Negative Images achieved the highest test accuracy (97.84%), followed by AHE Images (96.71%) and Original Images (96.67%).  
✔ ResNet50 outperformed CNN in terms of generalization, due to deeper feature extraction and residual learning.  
✔ Batch Normalization and Global Average Pooling (GAP) reduced overfitting, leading to better validation accuracy.

ResNet50 significantly outperformed CNN in classification accuracy, making it a more robust and reliable approach for mammogram analysis.  
HOG transformation led to overfitting issues, making direct feature extraction using deep learning a better choice for breast cancer detection.  
Negative Image transformation consistently improved classification accuracy, making it an effective preprocessing technique for deep learning models.  
 Further improvements can be explored, such as fine-tuning ResNet50, ensemble modeling, or using advanced architectures like EfficientNet.

To conclude

* ResNet50 is the preferred model for breast cancer detection using Mammogram images, as it has the higher accuracy, generalization and with improved robustness and that negative image preprocessing is the best approach among all transformations as it significantly enhances tumors visibility and classification accuracy**.**

# References TO DATA SOURCES

1. Acevedo, Andrea; Merino, Anna; Alferez, Santiago; Molina, Ángel; Boldú, Laura; Rodellar, José (2020), “A dataset for microscopic peripheral blood cell images for development of automatic recognition systems”, Mendeley Data, V1, doi: 10.17632/snkd93bnjr.1  
   Data download link: https://data.mendeley.com/datasets/snkd93bnjr/1
2. Mousavi Kouzehkanan, Zahra & Saghari, Sepehr & Tavakoli, Sajad & Rostami, Peyman & AbbasZadeh, Mohammadjavad & Shahabi Satlsar, Esmaeil & Mirzadeh, Farzaneh & Gheidishahran, Maryam & Gorgi, Fatemeh & Mohammadi, Saeed & Hosseini, Reshad. (2021). Raabin-WBC: a large free access dataset of white blood cells from normal peripheral blood. 10.1101/2021.05.02.442287.  
   Data download link: <http://dl.raabindata.com/WBC/Cropped_double_labeled/>
3. Almarri, A., Gupta, G., Kumar, R., Vandana, V., Asiri, F., & Khan, S. B. (2024). *The BCPM method: Decoding breast cancer with machine learning*. *BMC Medical Imaging*. <https://doi.org/10.1186/s12880-024-01402-5>
4. Breast Cancer Information, Support & Donations. (n.d.). National Breast Cancer Foundation. Retrieved February 2, 2025, from <https://www.nationalbreastcancer.org/>
5. Iniyan, S., Raja, M. S., Poonguzhali, R., Vikram, A., Ramesh, J. V. N., Mohanty, S. N., & Dudekula, K. V. (2024). Enhanced breast cancer diagnosis through integration of computer vision with fusion based joint transfer learning using multi modality medical images. *Scientific Reports*, *14*(1), 28376. <https://doi.org/10.1038/s41598-024-79363-6>
6. Lee, R. S., Gimenez, F., Hoogi, A., Miyake, K. K., Gorovoy, M., & Rubin, D. L. (2017). A curated mammography data set for use in computer-aided detection and diagnosis research. *Scientific Data,* 4, Article number: 170177. Retrieved February 19th, 2025, from <https://doi.org/10.1038/sdata.2017.177>
7. Mohi ud din, N., Dar, R., Rasool, M., & Assad, A. (2022). Breast cancer detection using deep learning: Datasets, methods, and challenges ahead. Computers in Biology and Medicine, 149. <https://doi.org/10.1016/j.compbiomed.2022.106073>
8. Nassif, A. B., Talib, M. A., Nasir, Q., Afadar, Y., & Elgendy, O. (2022). Breast cancer detection using artificial intelligence techniques: A systematic literature review. Artificial Intelligence in Medicine, 127, 102276. <https://doi.org/10.1016/j.artmed.2022.102276>
9. Shen, L., Margolies, L. R., Rothstein, J. H., Fluder, E., McBride, R., & Sieh, W. (2019). Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Scientific Reports*, *9*(1), 12495. <https://doi.org/10.1038/s41598-019-48995-4>
10. Wadhwa, G., & Kaur, A. (2020). A deep CNN technique for detection of breast cancer using histopathology images. *2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)*. <https://doi.org/10.1109/ACCTHPA49271.2020.9213192>
11. Zhang, B., Shi, H., & Wang, H. (2023). *Machine learning and AI in cancer prognosis, prediction, and treatment selection: A critical approach*. *Journal of Multidisciplinary Healthcare, 16*, 1779–1791. <https://doi.org/10.2147/JMDH.S410301>
12. Melek, A., Fakhry, S., & Basha, T. (2023). *Spatiotemporal mammography-based deep learning model for improved breast cancer risk prediction.* 2023 International Conference on Advanced Machine Learning (ICAML), <https://ieeexplore.ieee.org/abstract/document/10340602>.

# Justification for USING Existing code

When it comes to existing code, we will use basic CNN, and ResNet50 coding models. These neural networks will be used to provide us with a baseline code, in which we can run our data through to evaluate our performance. This starting point will allow us to understand how the basic models work and we can see how much pre-processing our data needs when using these models. We can then modify these neural networks with feature selection, residental learning, and some more advanced feature extraction techniques toachieve greater performance and accuracy. Overall, this will allow us a baseline and provide us with code that can be modified to gain better performance than previous studies.

1. “Convolutional Neural Network.” *Wikipedia*, Wikimedia Foundation, 17 September 2021, <https://en.wikipedia.org/wiki/Convolutional_neural_network>. [↑](#endnote-ref-1)
2. Wei. J. (2019). “Alexnet: The Architecture that Challenged CNNs”. TowardsDataScience. Retrieved September 25, 2021, from <https://towardsdatascience.com/alexnet-the-architecture-that-challenged-cnns-e406d5297951> [↑](#endnote-ref-2)
3. Q. A. Al-Haija and A. Adebanjo, "Breast Cancer Diagnosis in Histopathological Images Using ResNet-50 Convolutional Neural Network," *2020 IEEE International IOT, Electronics and Mechatronics Conference (IEMTRONICS)*, Vancouver, BC, Canada, 2020, pp. 1-7, doi: 10.1109/IEMTRONICS51293.2020.9216455.  
   <https://ieeexplore.ieee.org/document/9216455?utm_source=> [↑](#endnote-ref-3)
4. Patil, R.A., Dixit, V.V. Detection and classification of mammogram using ResNet-50. *Multimed Tools Appl* (2025). <https://doi.org/10.1007/s11042-025-20679-4>

   <https://link.springer.com/article/10.1007/s11042-025-20679-4?utm_source> [↑](#endnote-ref-4)
5. Browniee. J. (2019). “A Gentle Introduction to Transfer Learning for Deep Learning”.

   MachineLearningMastery. Retrieved September 25, 2021, from

   https://machinelearningmastery.com/transfer-learning-for-deep-learning/ [↑](#endnote-ref-5)
6. “Feature Selection.” *Wikipedia*, Wikimedia Foundation, 21 September 2021, <https://en.wikipedia.org/wiki/Feature_selection> [↑](#endnote-ref-6)