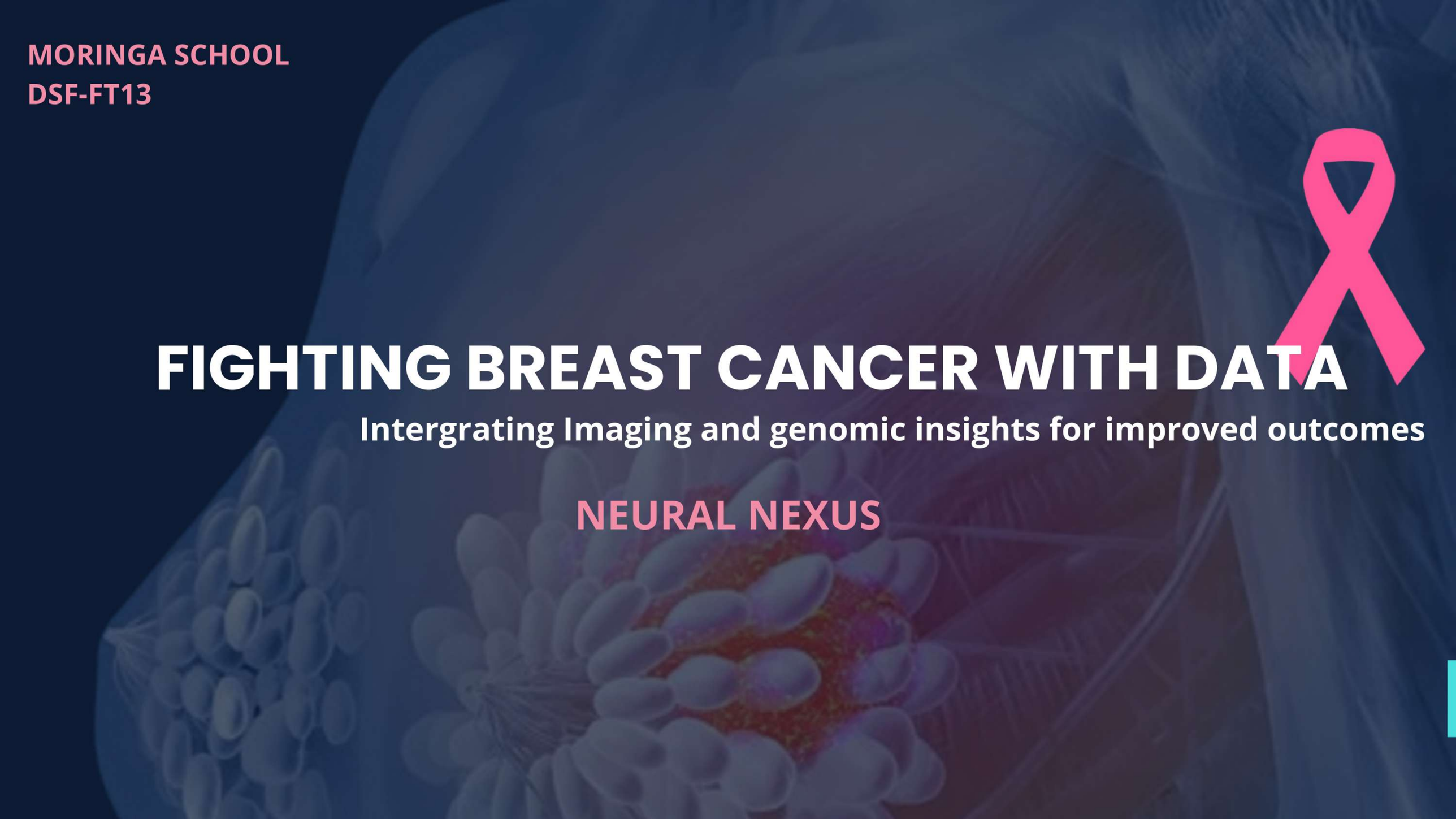
MORINGA SCHOOL
DSF-FT13



FIGHTING BREAST CANCER WITH DATA

Intergrating Imaging and genomic insights for improved outcomes

NEURAL NEXUS

The Team

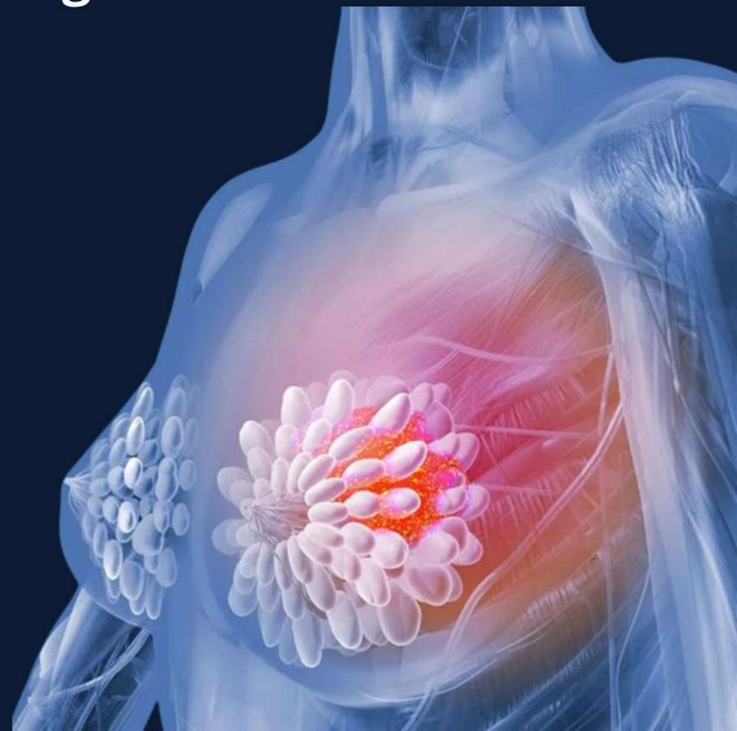
- Elsie Almadi- Scrum Master
- Khalid Abdulkadir
- Tiffany Cheruto
- Jeremy Koome
- Meshack Mboya

Introduction

- Every 14 seconds,somewhere in the world, a woman is diagnosed with breast cancer. Breast cancer remains one of the leading causes of cancer related deaths worldwide.
- Despite advances in screening, late detection and limited predictive tools hinder timely intervention.
- Since early and accurate prediction can significantly improve survival rates and optimize better treatment outcomes,this project intergrates imaging with both clinical and genomic models to support data driven medical decisions

Objectives

- To develop a cnn model that can accurately distinguish histopathological images
- To build a predictive model for surviaval and treatment outcomes using clinical and molecular data.
- To demonstrate how intergrating imaging and clinical based models enhance decision making in breast cancer diagnosis and prognosis.



Data Understanding

This project integrates histopathological imaging with clinical and molecular data to improve prediction.

Imaging (BreakHis)

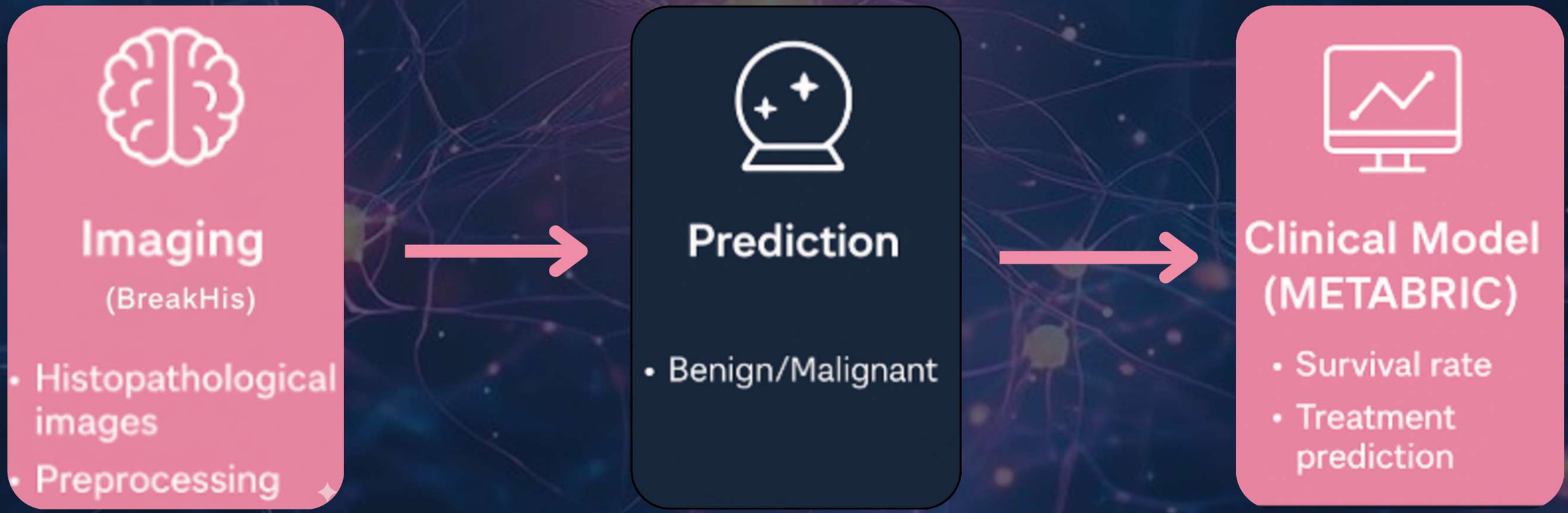
- **Source:** Histopatthological tissue images
- **Samples:** ~8,000 images (benign / malignant)
- **Classes:** Benign, Malignant
- **Goal:** Train CNN to classify tissue
- **Key Steps:** Resizing, normalization, augmentation

Clinical Data (Metabric)

- **Source:** METABRIC dataset
- **Patients:** 1,980 individuals
- **Treatments:** Chemotherapy, Radiotherapy. Hormone therapy
- **Goal:** Predict survival and treatment outcomes
- **Key Steps:** Cleaning, encoding, feature selection

Imaging data + Clinical data → Integrated Model

Approach


Imaging
(BreakHis)
• Histopathological images
• Preprocessing
Prediction
• Benign/Malignant
Clinical Model (METABRIC)
• Survival rate
• Treatment prediction

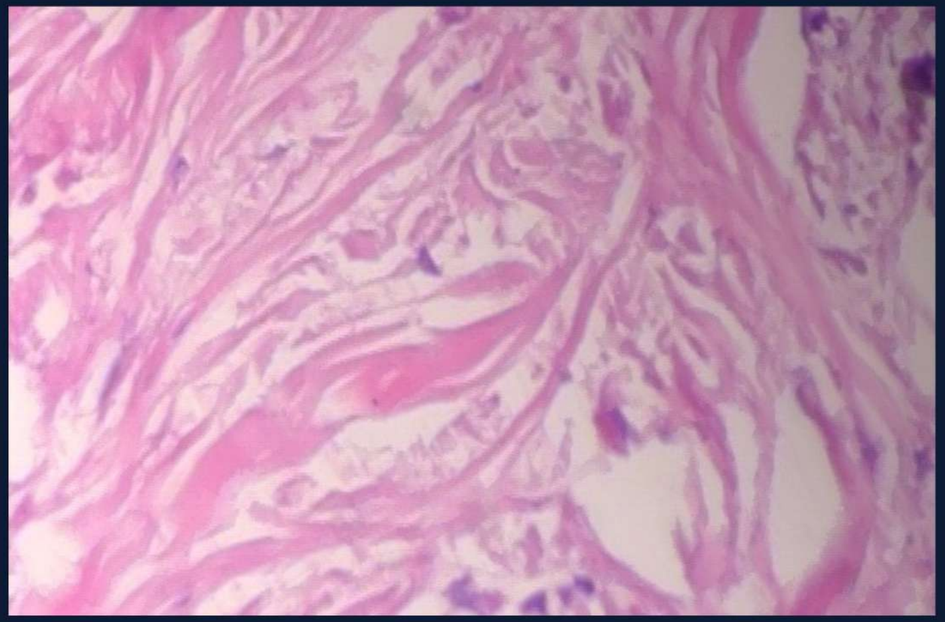
Tumor classification: Malignant vs Benign

Malignant:

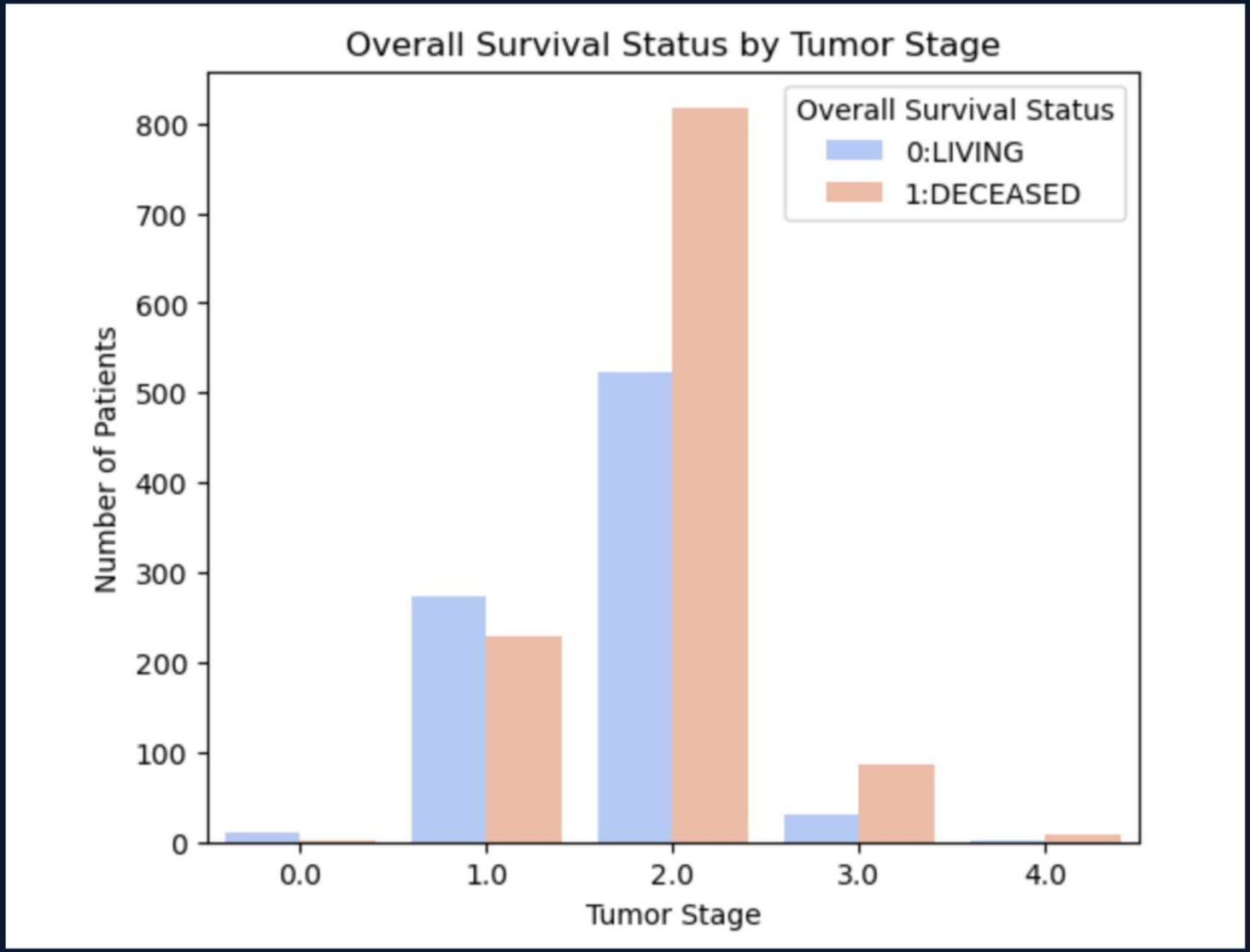
- Cancerous
- Invades surrounding tissue

Benign:

- non-cancerous
- does not invade surrounding tissue

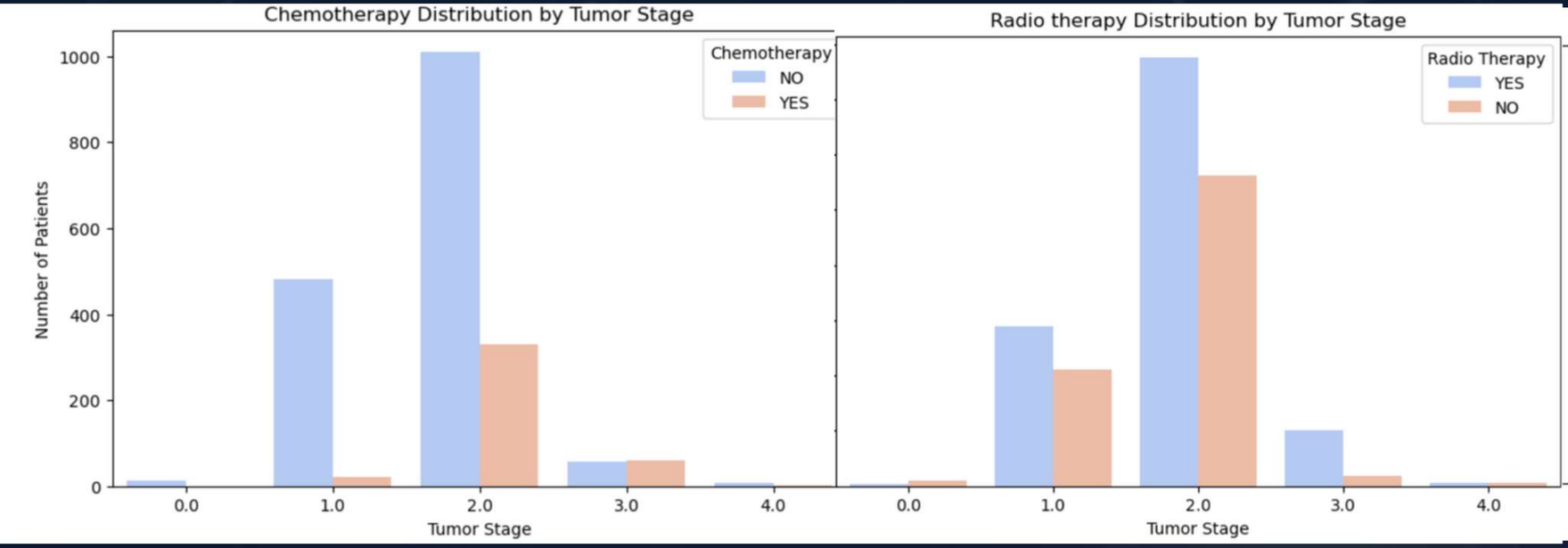


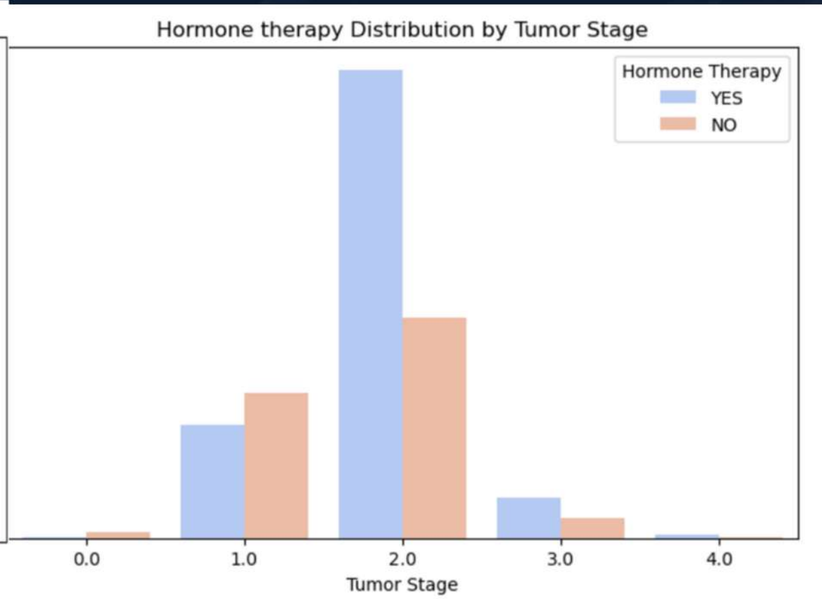
Overall Survival Status



- Survival is higher in the early stages
- Mortality increases with progression
- Stage 2 has the most number of patients

Treatment by Tumor Stage





- **Chemotherapy: works on all stages**
- **Hormone Therapy works on 2,3 and 4**
- **Radiotherapy : works better on 1,2 and 3**

Modeling

CNN for Image Classification

Image Preprocessing

- resizing, normalization
- data augmentation

CNN Architecture

- convolution → pooling → fully connected layers → softmax output

Evaluation Metrics

- accuracy, precision, recall, confusion matrix

Clinical Model for Survival & Treatment Prediction

Data preprocessing

- missing values handled
- categorical encoding

Model comparison

- Logistic Regression
- Random Forest
- Gradient Boosting

Evaluation

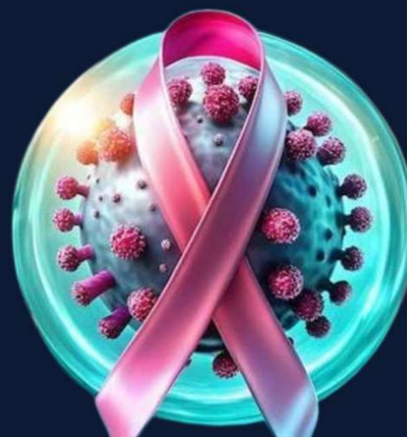
- accuracy, F1-score, ROC-AUC

Key Findings

- Image Classification (CNN) Insights:
  - CNN models were able to differentiate between benign and malignant histopathology images, showing promising predictive performance.
  - Certain tissue patterns are strongly predictive of malignancy, confirming the value of automated image analysis.
- Clinical Data (METABRIC) Insights:
  - Tumor stage, receptor status (ER/PR/HER2), and genetic markers are significant predictors of patient outcomes.
  - Integration of genetic and clinical features enhances predictive accuracy for survival and treatment response.
- Combined Insights:
  - Multi-modal analysis (images + clinical data) shows promise for more personalized and precise treatment planning.
  - Early detection and data-driven intervention can potentially improve patient prognosis.

Conclusions

- CNN models can distinguish benign from malignant tissues, supporting faster, more reliable diagnosis.
- Clinical features like tumor stage, receptor status, and genetic markers are strong predictors of patient outcomes.
- Integrating image-based and clinical data improves predictive performance and offers potential for personalized treatment.
- The findings highlight the importance of early detection, data-driven decision-making, and targeted interventions in breast cancer care



Recommendations

- **Early Detection:** Leverage predictive models to enable timely interventions.
- **Personalized care:** Tailor treatments based on individual characteristics
- **Model Enhancement:** Utilize Deep Generative Adversarial Network
- **Awareness:** Promote early screening

Next Steps

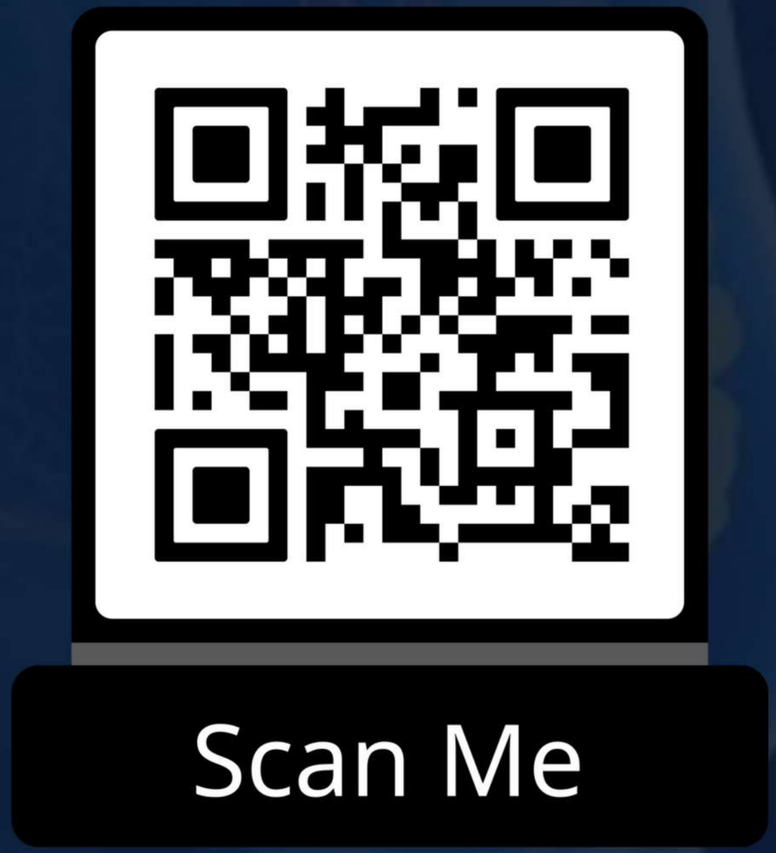
- **Model Optimization**: Fine-tune CNN and explore hybrid models combining images with clinical data.
- **Data Expansion**: Incorporate more diverse patient data, including African patient data to improve generalizability and relevance. Additional biomarkers should also be included
- **Validation**: Test models on external datasets or in a clinical setting for reliability.

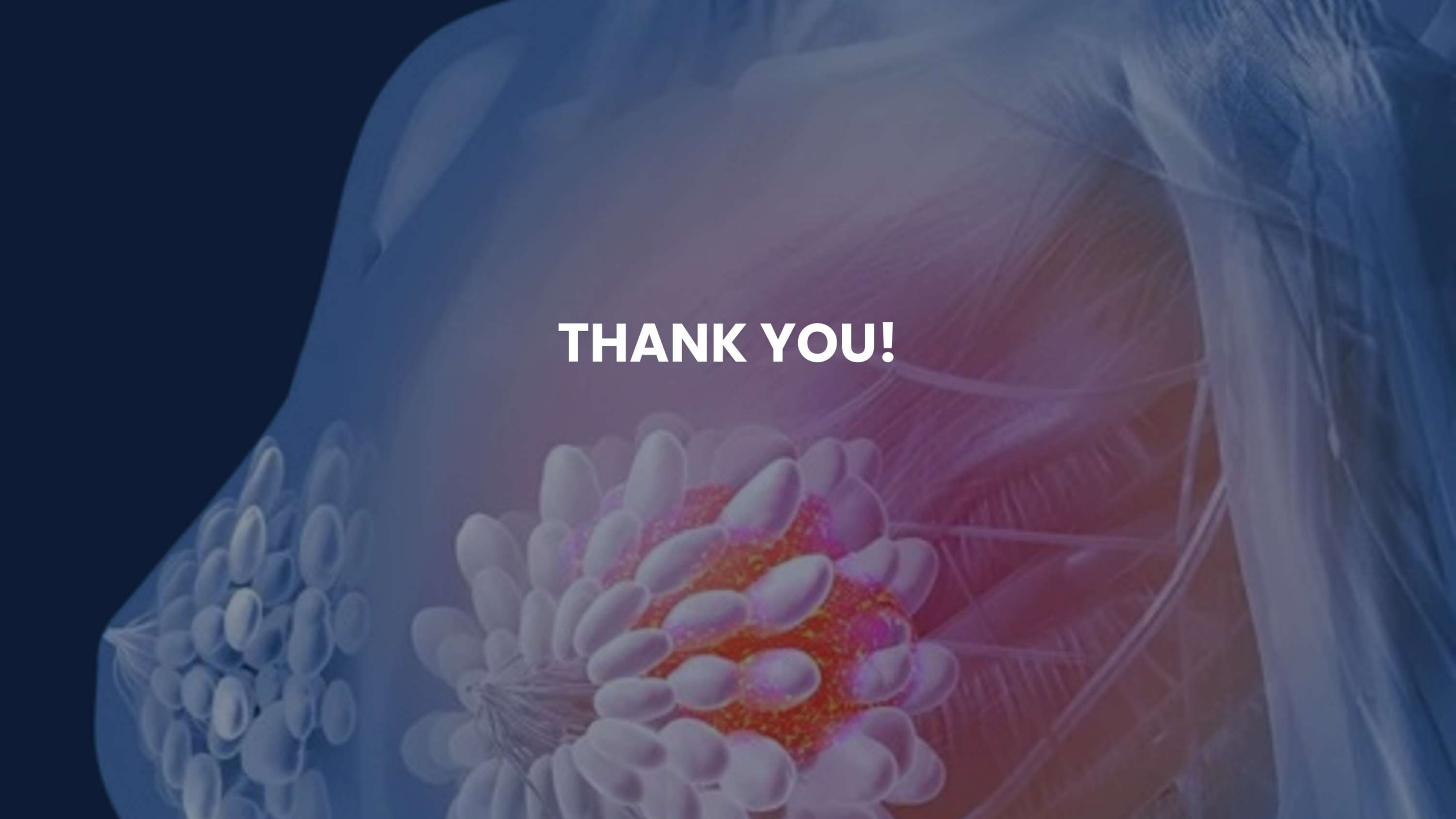
Research & Innovation: Explore new imaging techniques, feature selection methods, and explainable AI for better transparency.

Deployment

Neural Nexus- AI powered Cancer Treatment Platform

Backend: Django
Frontend : React
Version Control: GitHub Large File Storage
Cloud Platform : Render.com (25$/plan), Vercel




THANK YOU!