Group #5

BREAST CANCER ULTRASOUND IMAGE ANALYSIS

Jack McMorrow, Alejandra Mejia

Source: Google

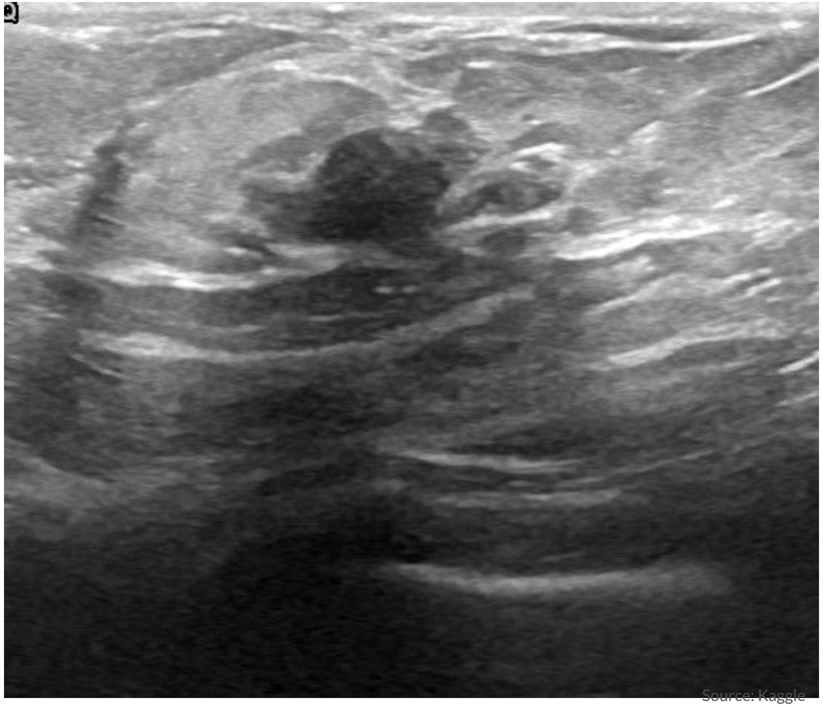
Why breast cancer detection is so relevant?

- Every **29 secs** breast cancer is diagnosed *(National Context)*
- **1 in 8 women** will develop breast cancer in her lifetime *(Norcal Think Pink)*
- **25% of women** with breast cancer are younger than 50 years old *(Practis inc)*
- **85% of women** diagnosed have no family history *(Virginia Breast Cancer Foundation)*
- **98% of women** survive with early detection of breast cancer*(Practis inc)*
- Ultrasound is highly accurate, research suggesting it can correctly detect around **80% of breast cancers** *(nbcf)*

Source: Norcal Think Pink

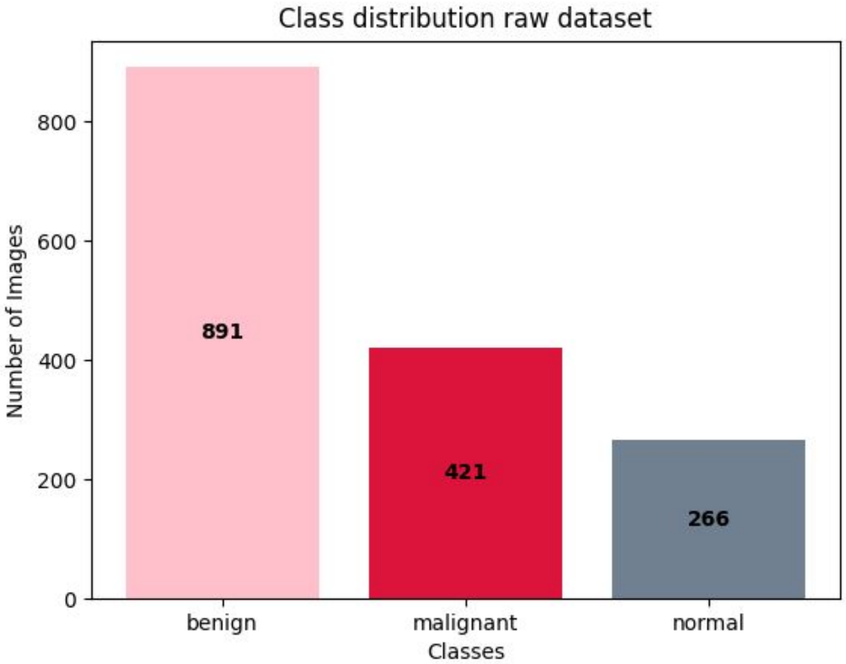
About our dataset

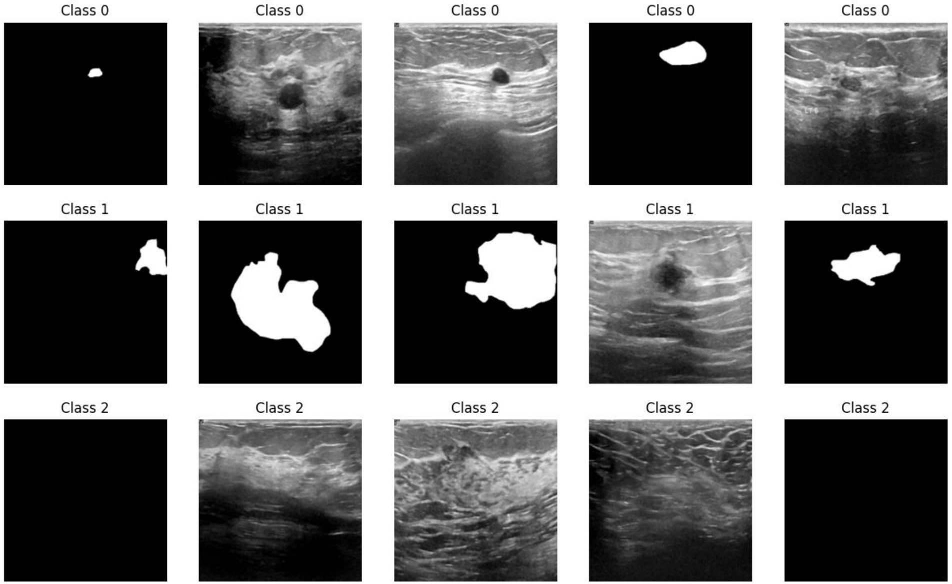
- The dataset used in the current project was retrieved from Kaggle and contains 780 images collected from 2018.
- Ground truth images are also contained named masks
- Overall, the dataset contains more than 1,500 images
- Unbalanced dataset
- Breast ultrasound images among women in ages between 25 and 75 years old
- Images are categorized in three categories: normal, benign, malignant



Source: Kaggle

Raw dataset overview



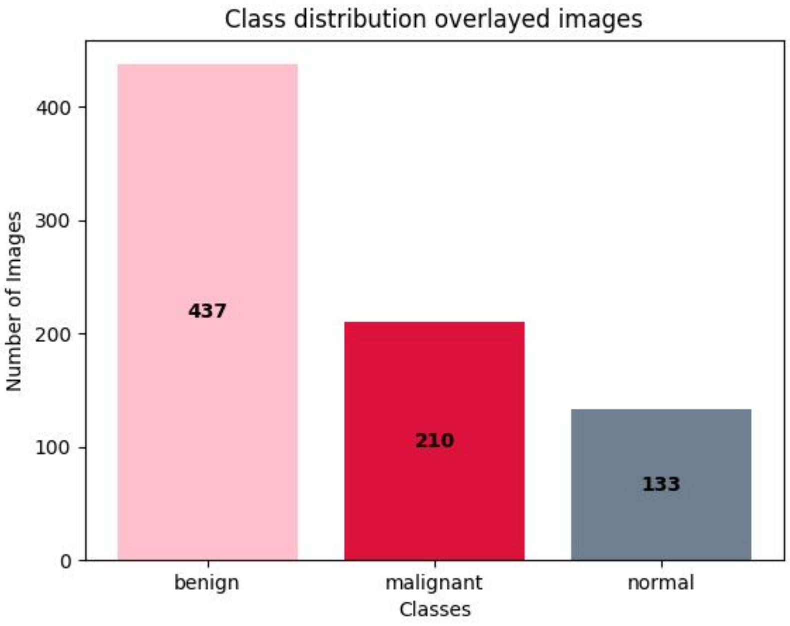


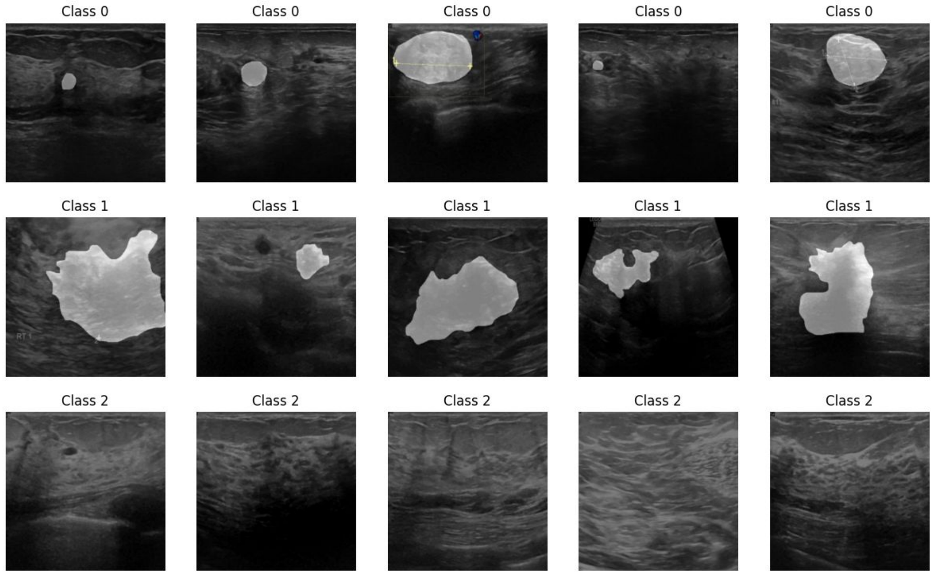
Pre-processing

- Dataset Creation:
  - Dataset were separated into original, masked and a new dataset was created overlaying the images
- Data was split into training, validation (0.2) and test sets (0.3)
- Dataset exploration:
  - Class distribution
  - Sample Images
- Resizing of images to 224 (ResNet alignment)
- Std and mean exploration for normalization
- Labels changed to:
    - 0: Benign
    - 1: Malignant
    - 2: Normal

Source: Google

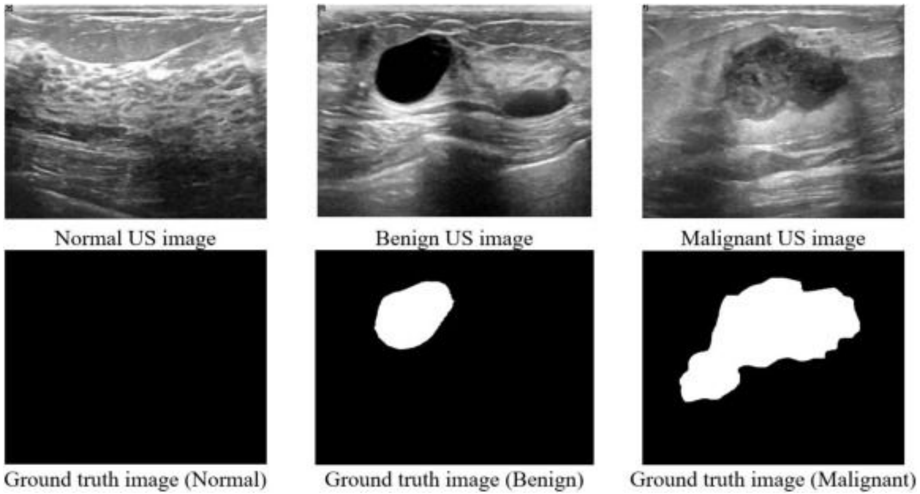
Overview of overlayed dataset





What are masks and why are they important?

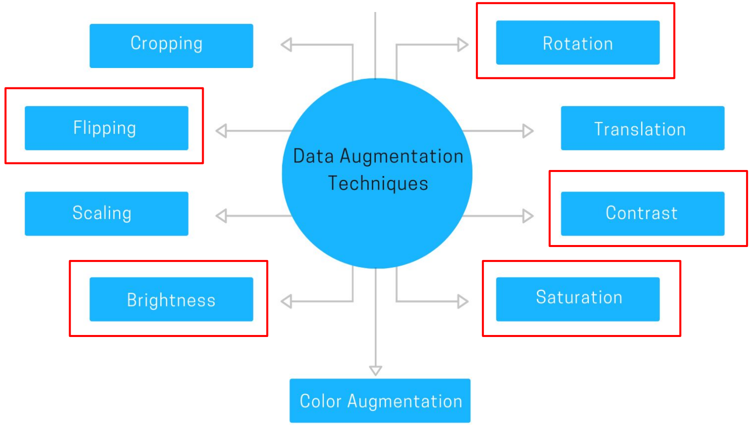
- **Mask:** Ground truth (image boundary) used Matlab an consisting of a freehand segmentation for each image.
- Using mask advantages:
    - Provide information about the spatial extent and location of abnormalities within medical images.
    - Enhance the interpretability of model predictions
    - Helps models generalize well to new patients and unseen variations in imaging conditions.



Source: Walid Al-Dhabyani et Al.(2020)

Augmentation techniques implemented

- When properly implemented augmentation techniques help improve model's performance, generalization, and robustness.
- Five techniques were used
- Given the unbalanced data set augmentation techniques were only implemented in the minority classes (malignant,normal) of the training dataset

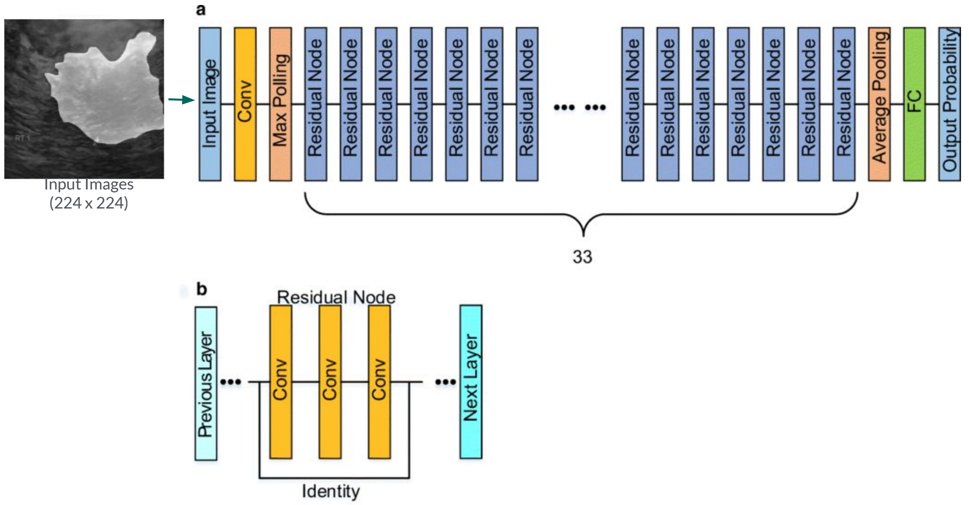


Source: AIMultiple

Our approach

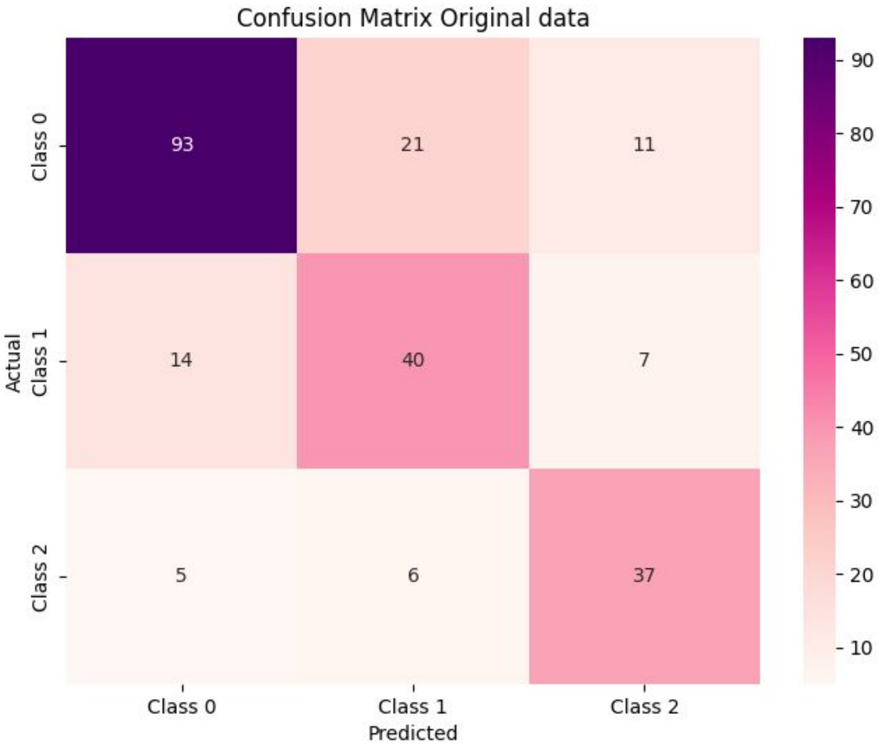
- Capitalize in Pytorch recent years popularity
- Tested out two different approaches:
  - Only use the actual image files and ignore the mask files to classify.
  - Overlay the actual image and the masks and use the overlayed images to classify to try improve model performance
- Used pretrained Resnet 101, benefits include:
  - Hierarchical feature learning
  - Transfer learning
  - Performance
  - Spatial Hierarchies
- Model architecture:
  - Adam, 20 epochs, CrossEntropyLoss, ReduceLROnPlateau

RESNET-101 Architecture



Source: Research Gate + Authors

Model 1: Original images with no mask

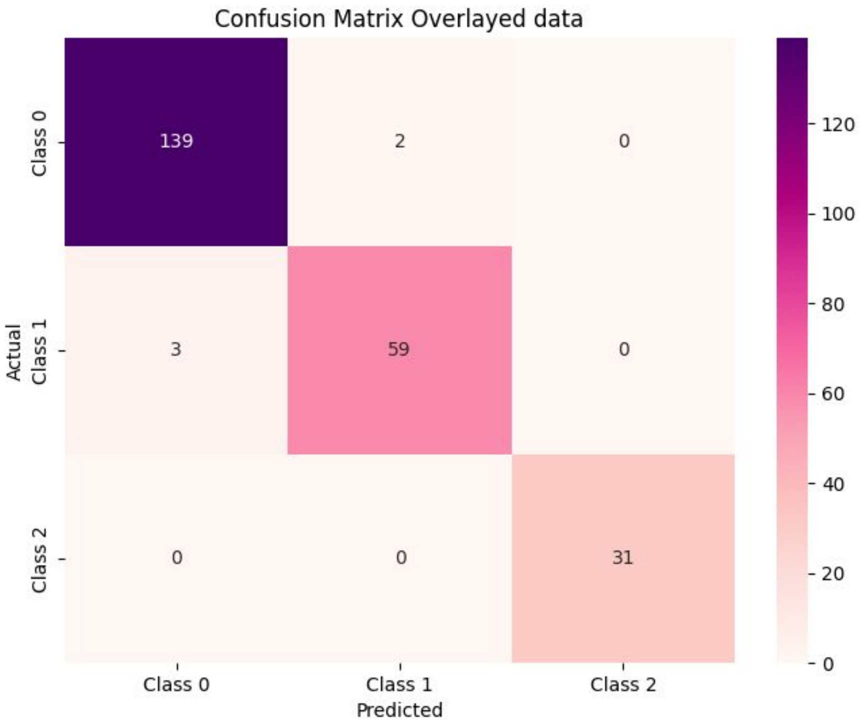



```
METRICS ON TEST DATASET
              precision    recall  f1-score   support

           0       0.83      0.74      0.78       125
           1       0.60      0.66      0.62        61
           2       0.67      0.77      0.72        48

    accuracy                           0.73       234
   macro avg       0.70      0.72      0.71       234
weighted avg       0.74      0.73      0.73       234
```

```
Accuracy: 0.7265, Precision: 0.7372, Recall: 0.7265, F1: 0.7295
```

Model 1: Original images with no mask

Model 2: Original with mask

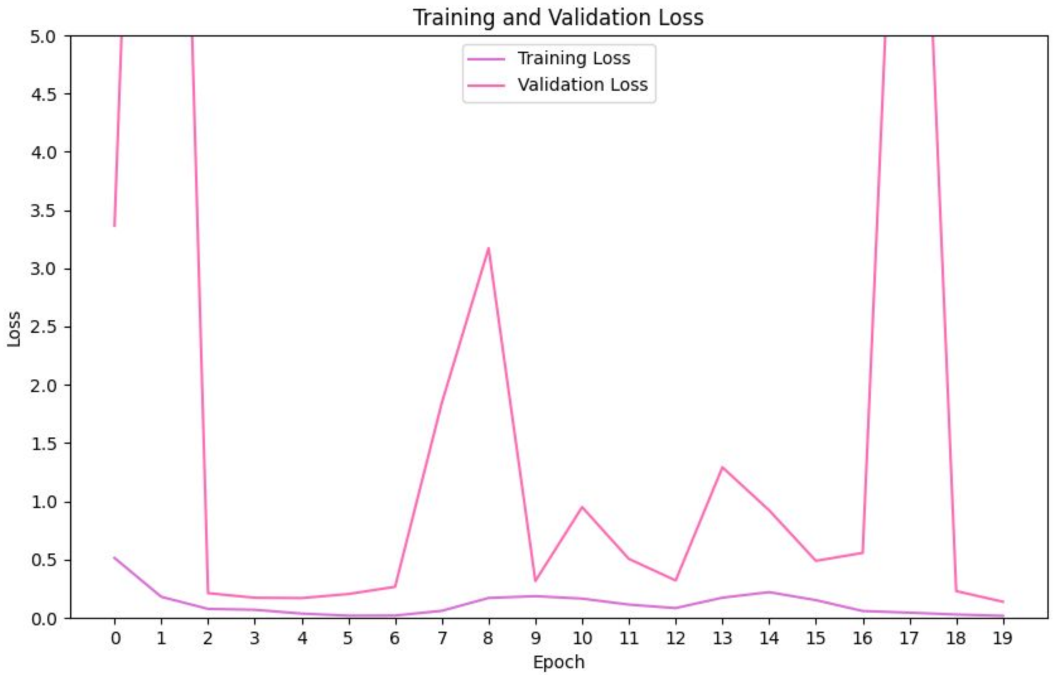


METRICS ON TEST DATASET

	precision	recall	f1-score	support
0	0.98	0.99	0.98	141
1	0.97	0.95	0.96	62
2	1.00	1.00	1.00	31
accuracy			0.98	234
macro avg	0.98	0.98	0.98	234
weighted avg	0.98	0.98	0.98	234

Accuracy: 0.9786, Precision: 0.9786, Recall: 0.9786, F1: 0.9786

Model 2: Original with mask



Limitations and future work

- Strategies such as class weighting or data augmentation may be insufficient to address severe imbalances.
- The model's performance could be constrained by a limited dataset size, especially in medical imaging where acquiring labeled data is often resource-intensive.
- The model may struggle to generalize to diverse and rare cases not well-represented in the training set.
- Model performance may be sensitive to the quality of input images, and variations in imaging conditions might affect predictions.

- Exploring ensemble learning techniques, combining predictions from multiple models, can enhance overall model performance and provide more robust predictions.
- Incorporating explainability techniques such as SHAP (SHapley Additive exPlanations) or Grad-CAM (Gradient-weighted Class Activation Mapping) for better model interpretability.
- Evaluating the model's performance in real-world clinical settings and ensuring its seamless integration into the existing healthcare infrastructure.
- Consider ethical implications, privacy concerns, and regulatory compliance in healthcare AI applications.

Streamlit Demo