BREAST CANCER CLASSIFICATION & MACHINE LEARNING

By: Eduardo Olmos

The Epidemic

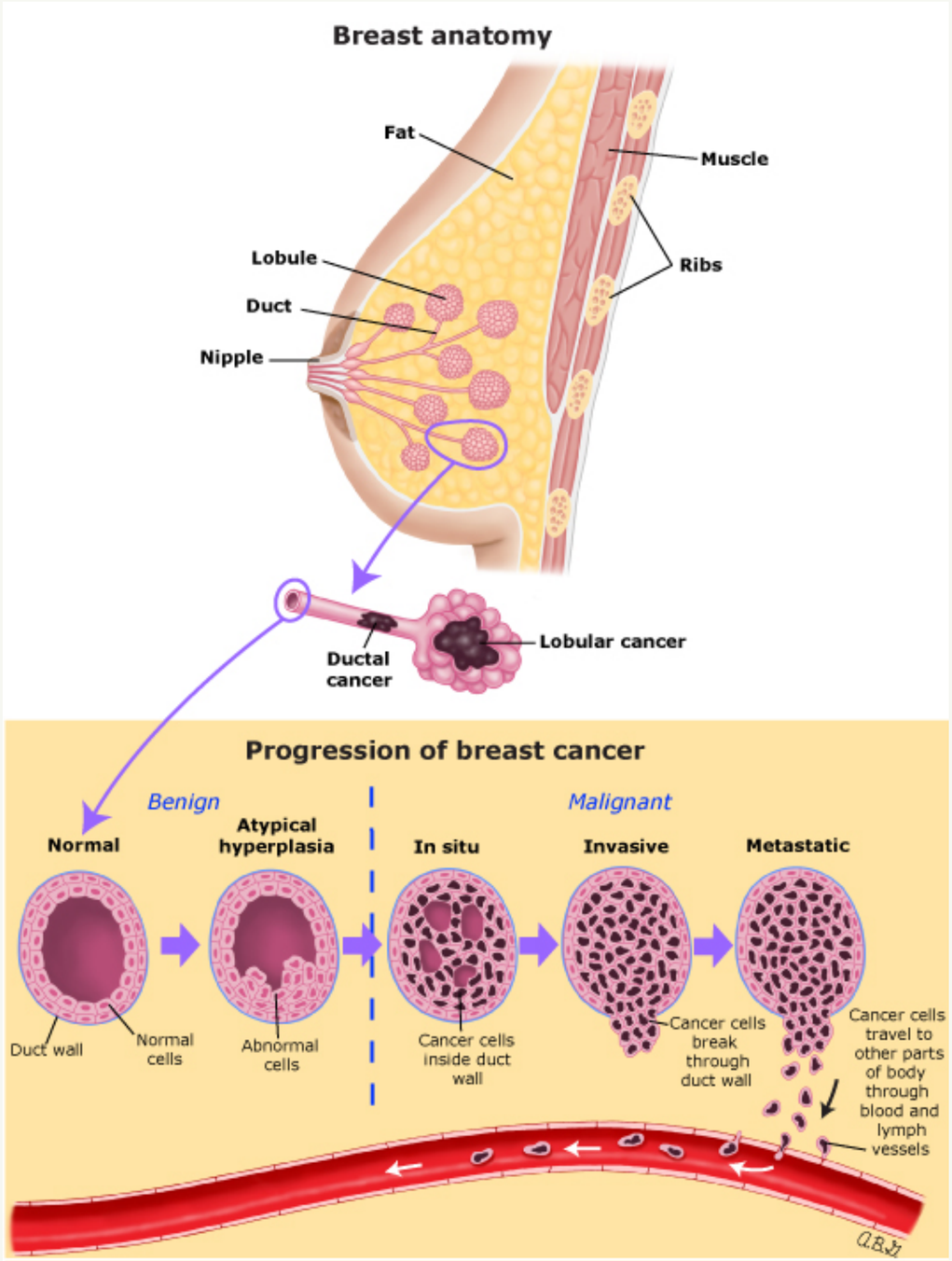
Background

- Affecting 1 in 8 U.S. women (13%).[1]
- 43.6k deaths expected among them in 2021.[1]
- Death rate decreased by 1% per year from 2013 to 2018.[1]
    - Expected to be the result of treatment advances and earlier detection through screening.[1]

What is Breast Cancer?

Background

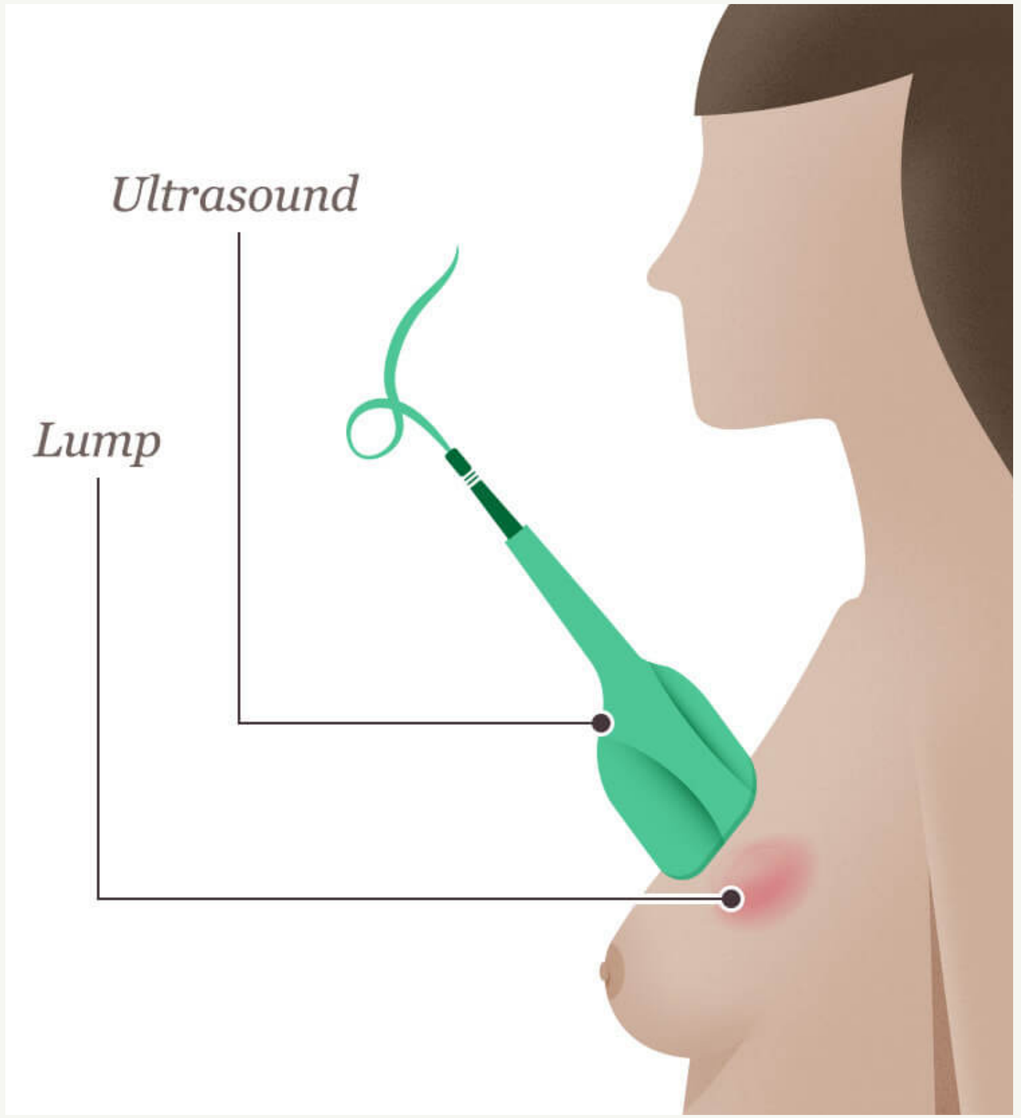
- Breast cancer cells usually form a tumor that can often be seen on an x-ray or felt as a lump.[2]
- Most breast lumps are benign and not cancer (malignant).[2]
- Any breast lump or change needs to be checked by a health care professional to determine if it is benign or malignant (cancer).[2]



Screening - Imaging

Background

- **Mammogram**
  - Uses X-rays that checks for cancer in breast tissue.[3]
- **MRI**
  - Uses magnetic fields, not x-rays, to produce detailed images of the body.[3]
- **Ultrasound**
  - Uses sound waves to create a picture of the breast tissue.[3]



Screening - Biopsy

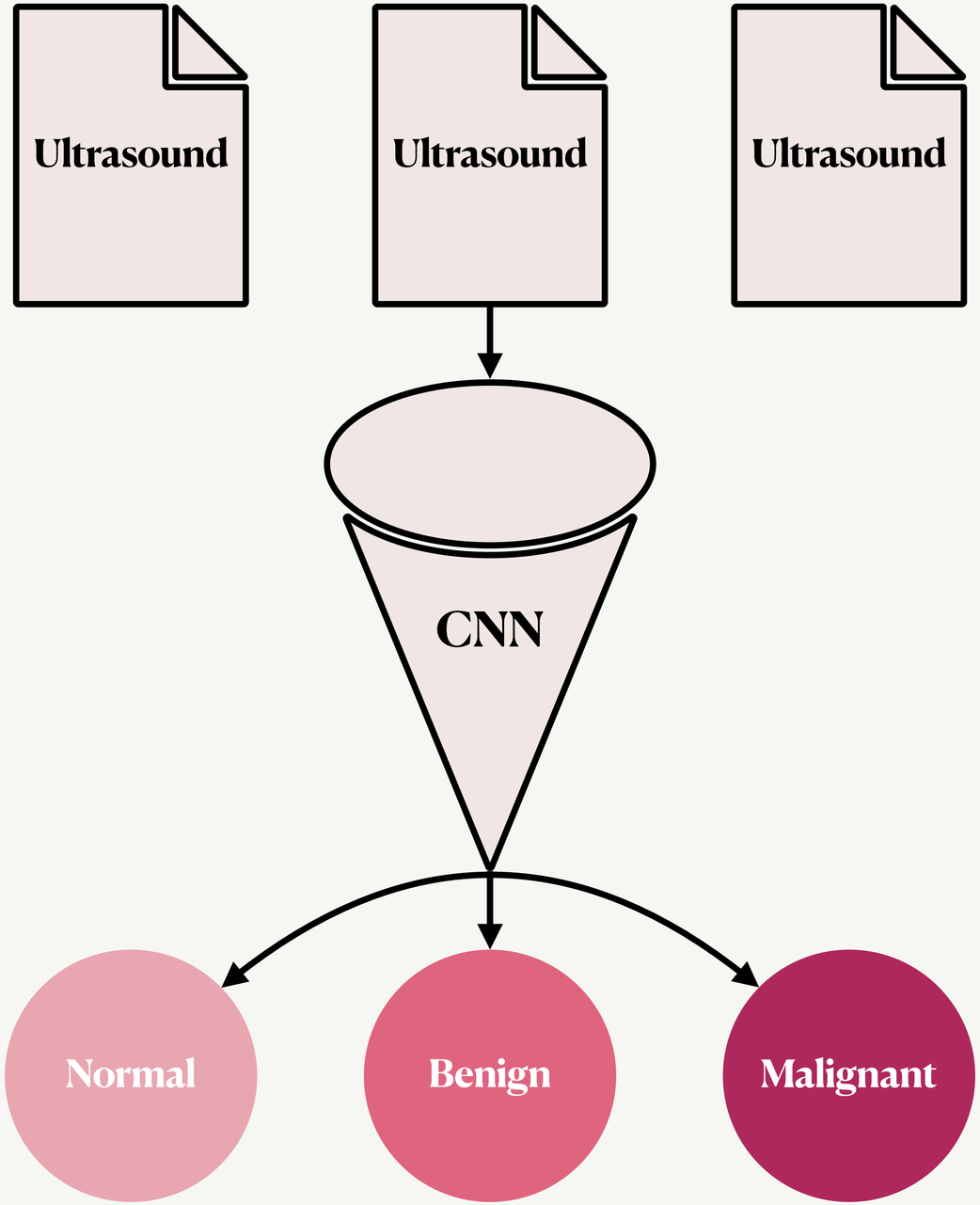
Background

- A biopsy is usually the only sure way for the doctor to know if an area of the body has cancer.[3]
- Used to determine if it is invasive or non-invasive (in situ); ductal, lobular, or another type of breast cancer.[3]
- Results of these tests will help determine treatment options.[3]

Normal, Benign, or Malignant?

Problem Statement

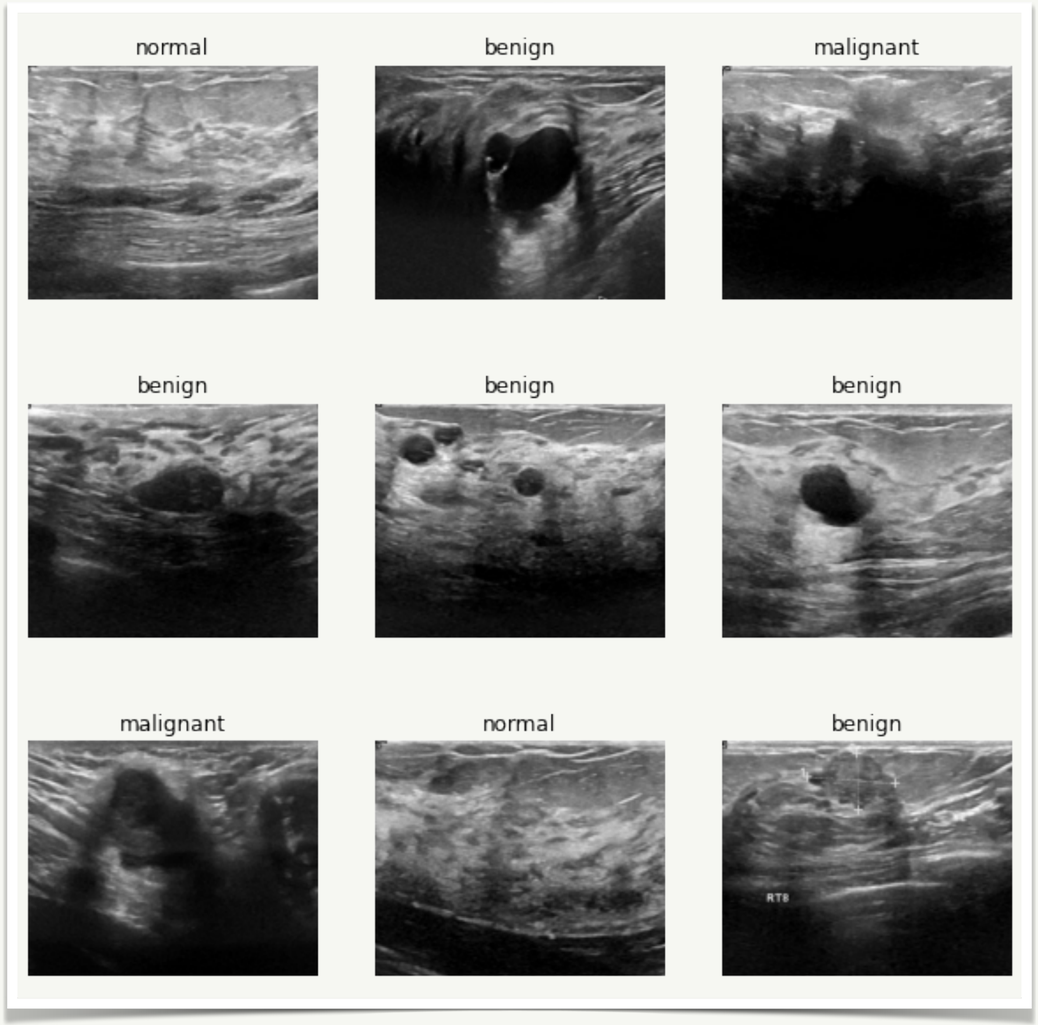
- Build and train a Convolutional Neural Network (CNN) to classify a breast ultrasound as either normal, benign, or malignant.



Breast Ultrasound Images Dataset (BUSI)

Data Collection

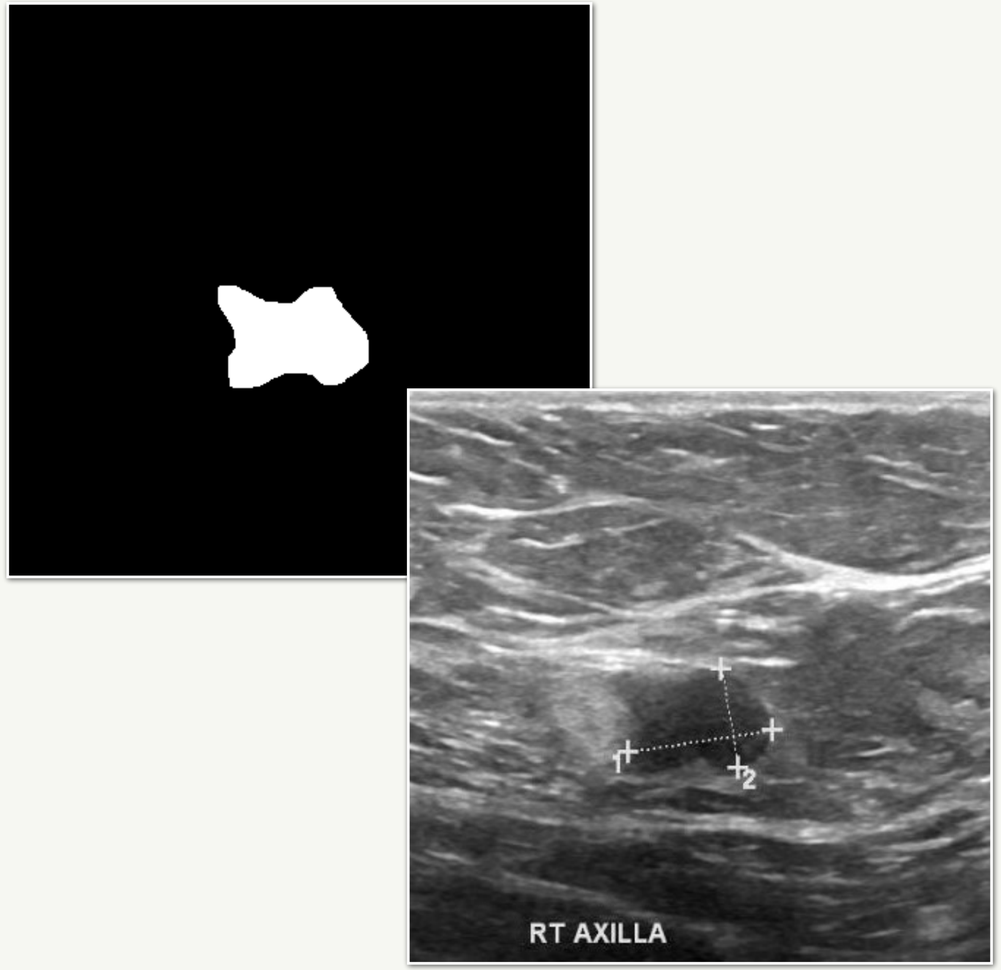
- Collected by Dr. Aly Fahmy and team in 2018.[4]
- Includes breast ultrasounds from 600 women from ages 25 to 75 years old.[4]
- Contains 780 images “with an average image size of 500*500 pixels.”[4]
- Labeled as normal, benign, or malignant.[4]



Preliminary Observations

Data Cleaning & EDA

- Formatted by proficient data authorities.
- Average pixel size was about 500*615 pixels.
- Writing and tags provided on some images.
- Ground truth images also provided.



Rationale for Decisions and Assumptions

Preprocessing and Modeling

- Utilizing the Keras package from TensorFlow.
- Going with the calculated average height and width of the images.

TensorFlow

Rationale for Decisions and Assumptions

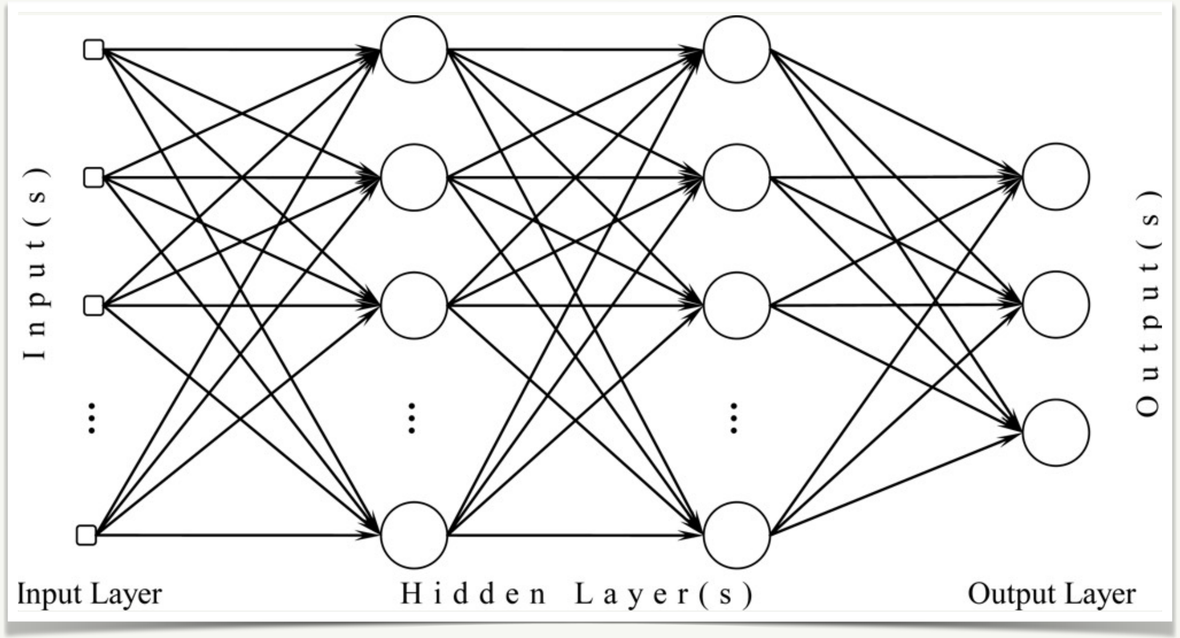
Preprocessing and Modeling

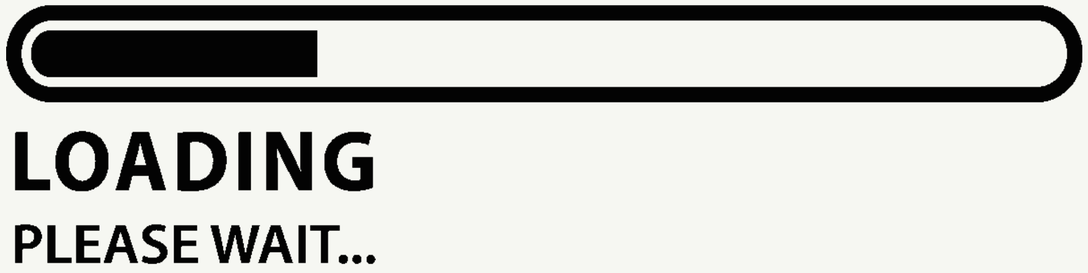
TensorFlow

Building the Model

Preprocessing and Modeling

- Rescale pixel values to adhere to CNN.
- Add hidden layers:
  - Convolution
  - Dense
- Add output layer for the three classes.
- Train the model

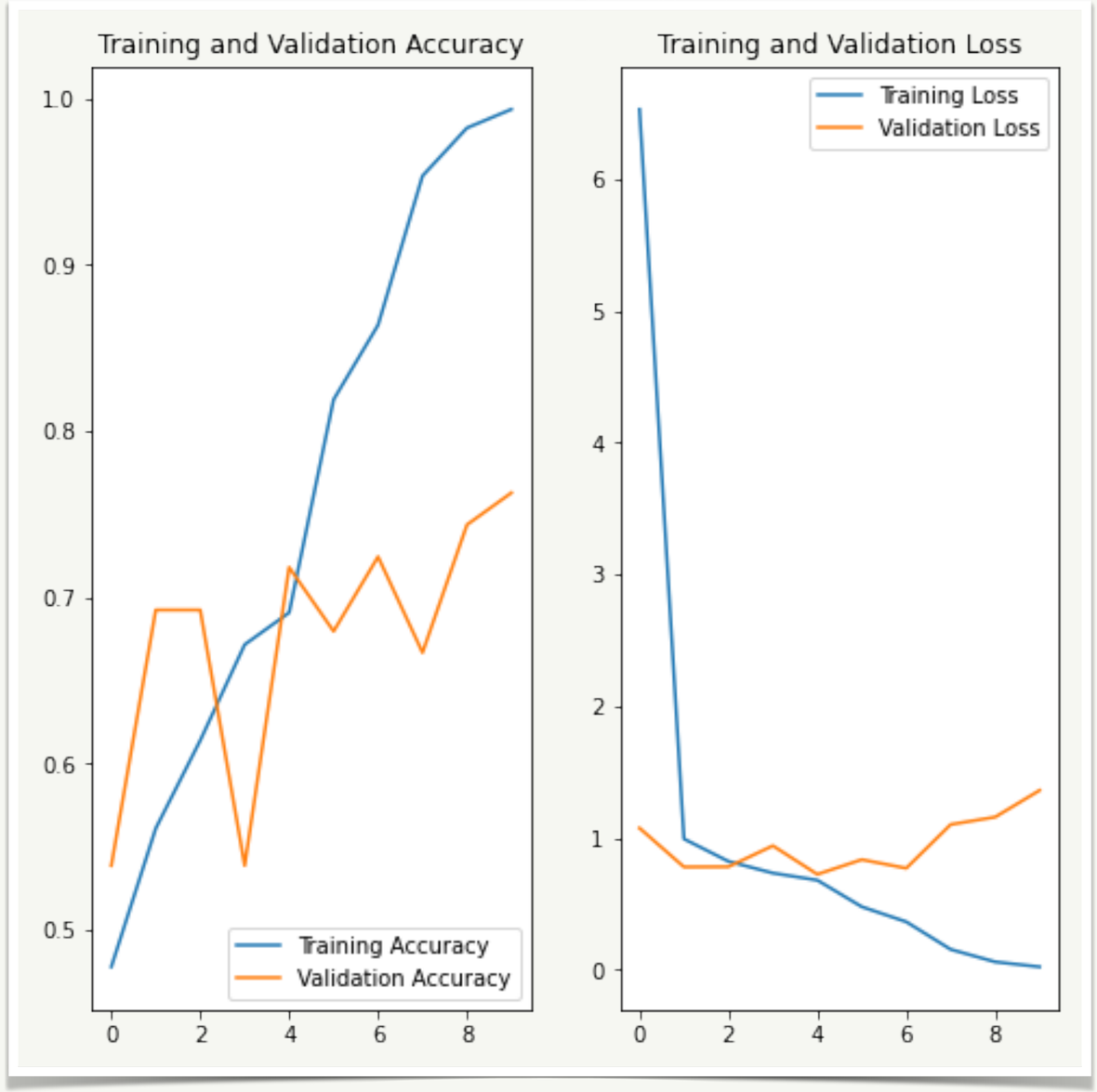




Assessing the Model

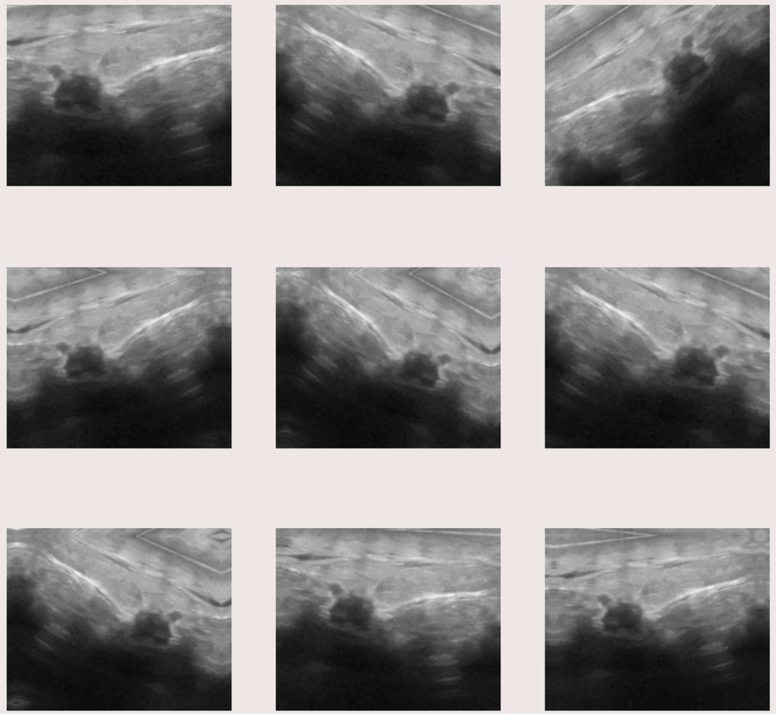
Evaluation

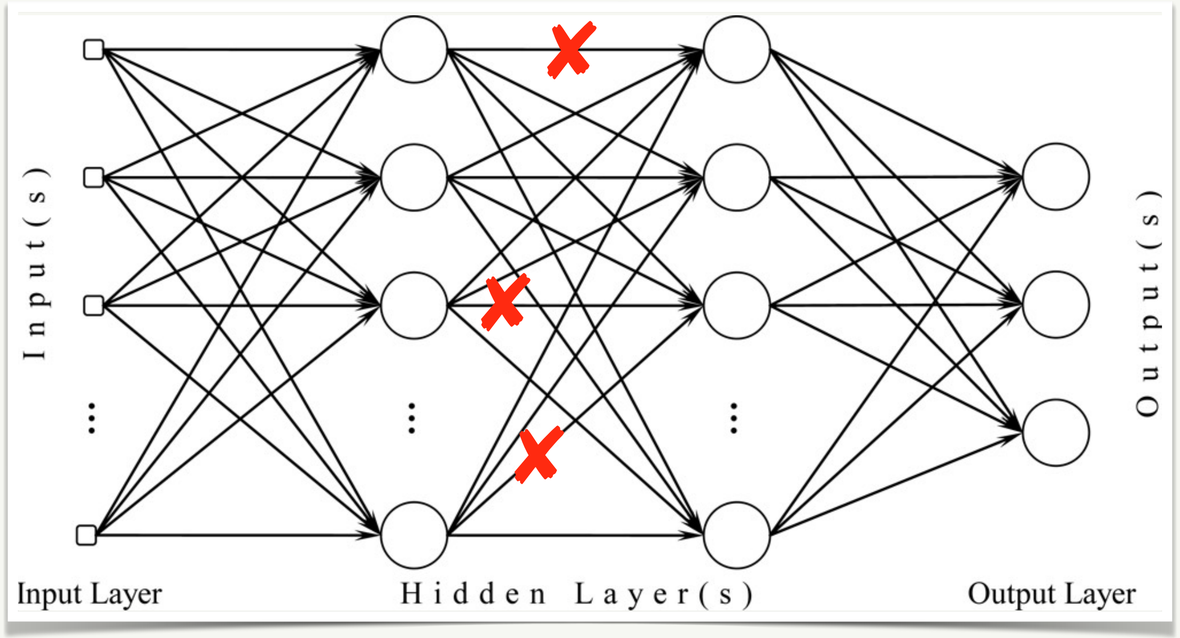
- Model is severely overfit.
- Results are trending well, but not good enough.
- We can do better.

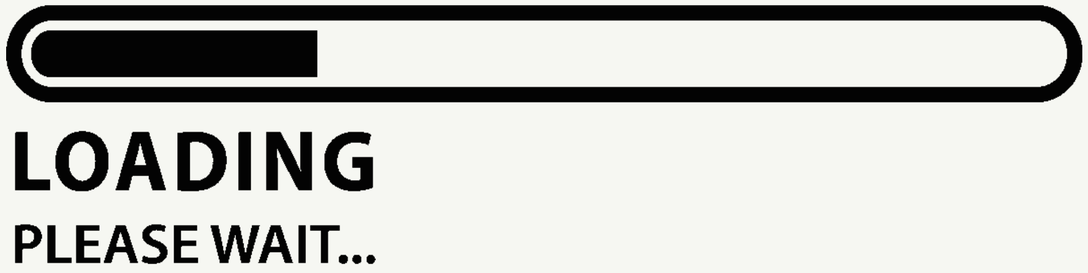


Data Augmentation & Dropout Layers

Preprocessing and Modeling



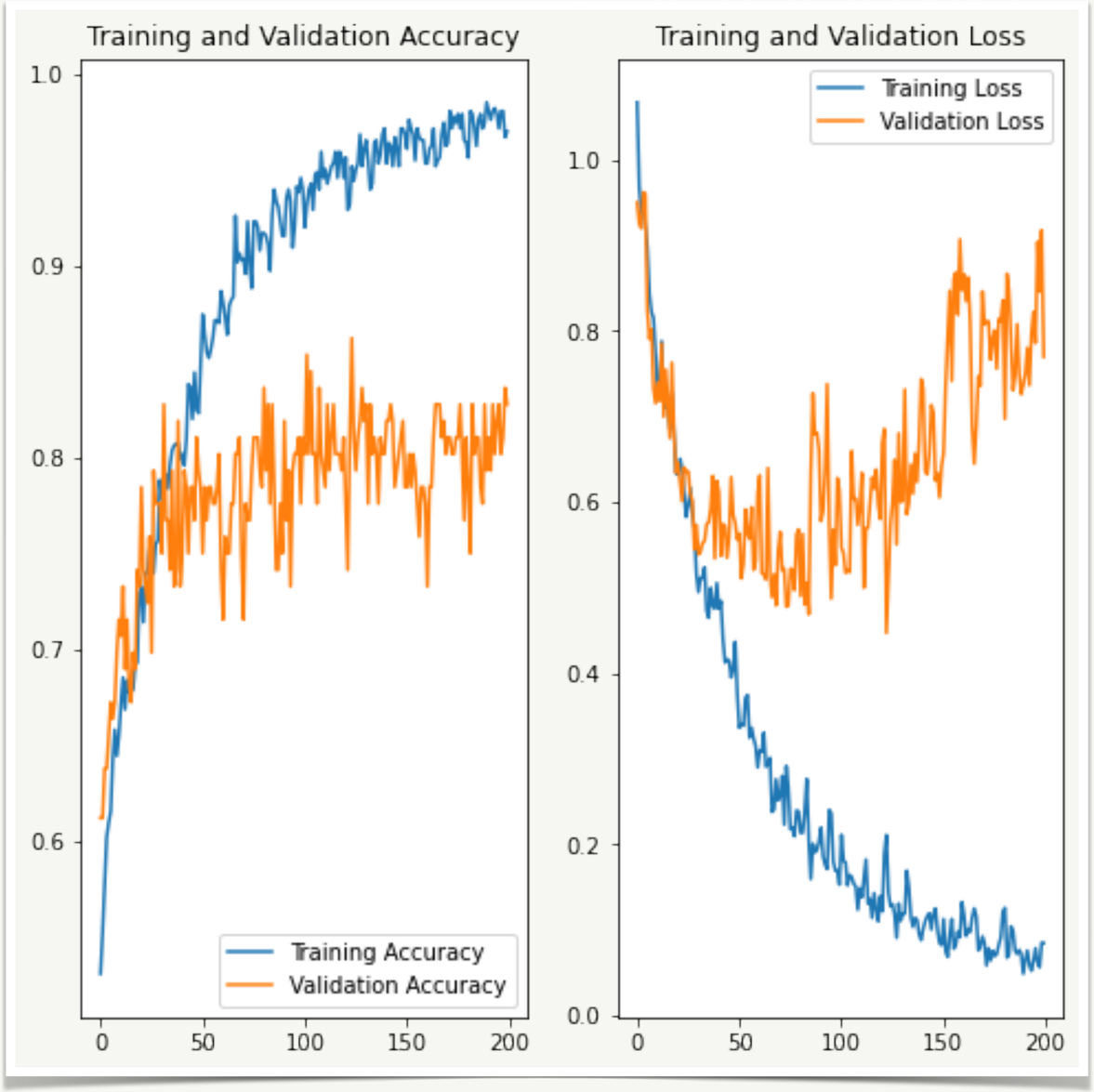




Assessing the Model

Summary

- Model is still overfit, but took substantially more epochs to get there.
- Validation data seemed to perform better than training data for the first 50 epochs.
- Validation and training loss lower than previous version.
- We can *always* do better.



Desired Improvements

Next Steps

- Implement a GridSearch
- Create a confusion matrix
- Connect with a SME


THANK YOU!

Sources

References

1. https://www.breastcancer.org/symptoms/understand_bc/statistics
2. https://www.cancer.org/cancer/breast-cancer/about/what-is-breast-cancer.html
3. https://www.cancer.net/cancer-types/breast-cancer/diagnosis
4. https://scholar.cu.edu.eg/?q=afahmy/pages/dataset
5. Dataset provided by: Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. Data in Brief. 2020 Feb;28:104863. DOI: 10.1016/j.dib.2019.104863.

Sources

Images

- https://oakland.edu/medicine/news/auto-list-news/2021/More-needs-to-be-done-to-help-Black-women-fight-breast-cancer,-says-OUWB-prof
- https://chicago.suntimes.com/well/2021/10/8/22704713/breast-cancer-subtypes-treatments-developments-wellness
- https://www.uptodate.com/contents/image/print?imageKey=PI%2F53453
- https://www.nationalbreastcancer.org/breast-ultrasound
- https://www.sanovadermatology.com/skin-cancer-blog-cat/what-is-a-skin-biopsy/
- https://www.istockphoto.com/vector/breast-cancer-awareness-with-realistic-pink-ribbon-on-a-white-background-women-gm1176663746-328162763
- https://www.mdpi.com/1424-8220/10/9/8363/htm
- https://blog.hellojs.org/create-a-very-basic-loading-screen-using-only-javascript-css-3cf099c48b19
- https://rethinkbreastcancer.com/women-with-mbc-need-you-to-be-their-ally/
- https://www.nasa.gov/centers/wstf/about_us/safety_and_mission_assurance/continual_improvement.html