# Machine Learning and Tomosynthesis for Breast Cancer Detection

Project 1: “Building 3D scenes from 2D images using ML algorithms”

Molly Housego

EC601

# Introduction: 3D visualization methodology

3D Imaging from 2D images

- “Z stack”
  - Using numerous 2D images as slices of a 3 dimensional object and then compiling to digitally render
- Multi orientation and machine learning
  - Using a set of training data, utilize context clues such as perspective and shade variation to enable 3 dimensional construction

3D Imaging from Measurement

- Utilizes equipment to gather depth of field measurements from a real 3D scene or object in order to digitally render.
  - LiDAR (light detection and ranging)

Generating a 3D rendering from 2D images utilizes already evolved and optimized photo-video technology - In comparison LiDAR is expensive and time consuming.

# Introduction: 3D visualization Significance

The continual need to increase the efficiency of  human centered work - creates the necessity of technological advancement allowing machines to process data from the natural world and then communicate that information in a way humans can understand.

 This is seen in advancements in speech recognition and facial expression analysis.

Sight being one of mankind's most pertinent senses led to the heavy emphasis on image processing.

# 2 to 3D Visualization Applications

2d to 3d technology can be found in almost every aspect of business. Examples include but aren't limited to:

- Transportation - Tesla self driving cars
- Security - Facial recognition now widely used not only for mobile access but 3rd party applications such as mobile banking
- Social media - interactive filters
- Military - Self navigating drones
- Healthcare - Mammogram imaging technology

# Area of Focus Selection

Began Researching basics of 3D visualization → Began looking into applications specifically beginning with two dimensional images → Discovery of Tomosynthesis → Discovery of combination Tomosynthesis and Machine Learning for diagnosis

My area of focus for this project did not end up fitting the project category perfectly, however it ties in the key elements of 2D to 3D visualization and machine learning for a noteworthy purpose.

# Background: Machine Learning and Tomosynthesis for Breast Cancer Detection

- Breast cancer was the US's most prevalent malignancy in terms of new cases in 2020 [1]
- One in eight women with undergo diagnosis at some point in her lifetime [2]
- Major technological breakthroughs have been critical for early detection screening specifically in mammogram imaging and analysis

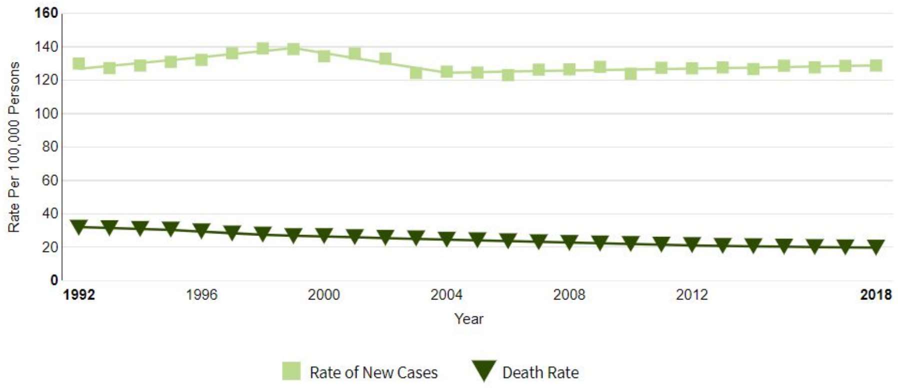


Figure 1. US Breast Cancer Case statistics [1]

# Background: Machine Learning and Tomosynthesis for Breast Cancer Detection

- Tomosynthesis (3-D mammography)
  - after taking x-ray images at slightly varying positions a computer generates a 3-D image of the breast composed of milimeter slices allowing the radiologist to view each layer of tissue.
  - The procedure reduces pictured overlapping tissue reducing false positive call backs by 15% and increases true detection of cancer by 25% [2]

- Computer Aided Detection (CAD)
  - A machine learning software utilizes digitized mammogram images to search for abnormalities that may indicate cancer.
  - CAD is an increase in efficiency, drawing the radiologists attention to areas of concern much faster than human analysis [2]

From here I looked at literature attempting to combine the two existing technologies

# Literature Review

- “Artificial Intelligence for Mammography and Digital Breast Tomosynthesis: Current Concepts and Future Perspectives” [3]
  - Images that are produced from tomosynthesis machines vary more greatly from vendor to vendor than do normal mammograms adding a level of complexity to analysis and comparison over time.
  - Most effective machine learning application is convolutional neural networks. CNN has difficulty in identifying masses and calcifications simultaneously, so often the two features are trained separately and combined at final step for diagnosis
  - Due to the relatively small data set of Digital Breast Tomosynthesis (DBT), transfer learning from existing mammogram data sets have been attempted.
  - Most important future application is to use the AI over a period time to be able to compare changes in a patients which is vital in breast cancer diagnosis.

# Literature Review

- "Deep Learning Framework for Digital Breast Tomosynthesis Reconstruction" [4]
  - Focuses on data driven approach for image reconstruction of DBT named Deep Breast Tomographic Reconstruction (DBToR) - goes beyond existing learned Primal-Dual algorithm by adding a breast tissue thickness
  - Test and training data comes from phantom breast models not real patients
  - Resulting DBToR resolution and noise outperforms the baseline DBT iterative reconstruction algorithm

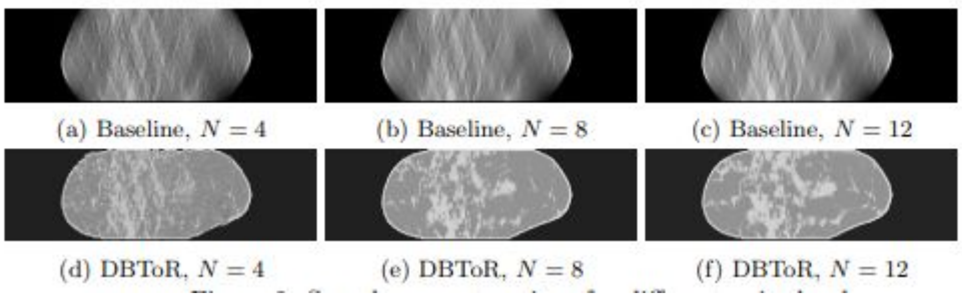


Figure 2. Sample reconstructions for different noise levels [4]

# Literature Review

- "Applying deep learning in digital breast tomosynthesis for automatic breast cancer detection: A review" [5]
  - Provides a summary of all relevant deep learning applications to DBT imaging as of 2021
  - Most existing analyses have non diagnostic classification purposes
  - The majority of analyses include transfer learning with up to 97% accuracy
  - 25/26 studies utilized private patient datasets

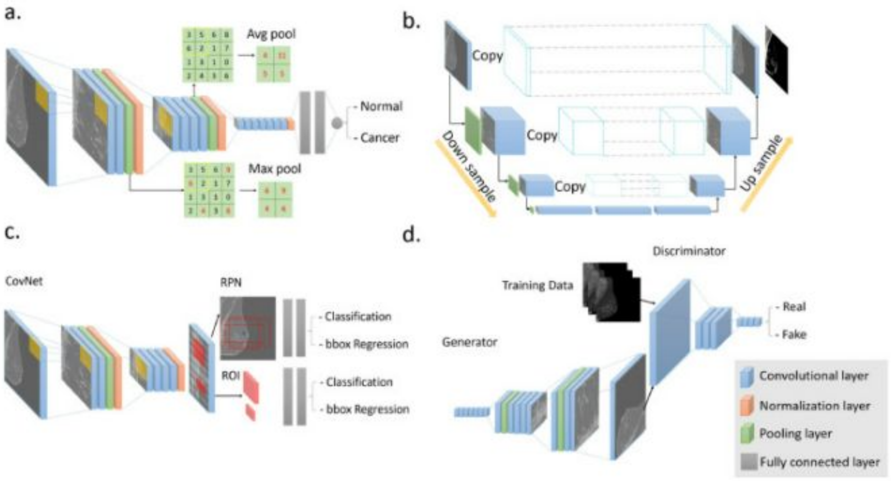


Figure 3. a. Structure of a CNN model b. Structure of U-Net model c. Structure of RCNN model d. Structure of GAN simulation [5]

# Literature Review

- “Detection of masses and architectural distortions in digital breast tomosynthesis: a publicly available dataset of 5,060 patients and a deep learning model” [6]
    - Common challenges for validating AI systems for DBT are existing datasets usually come from limited institutions and devices limiting reproducibility
    - Manually sorted dataset into biopsied and non biopsied categories for training purposes
    - Utilized You Only Look Once: Unified, Real-Time Object Detection a CNN with denseNet architecture
    - Used free receiver operating characteristics curve - depicting the relationship between false positives predictions and model sensitivity with best sensitivity result of 67%

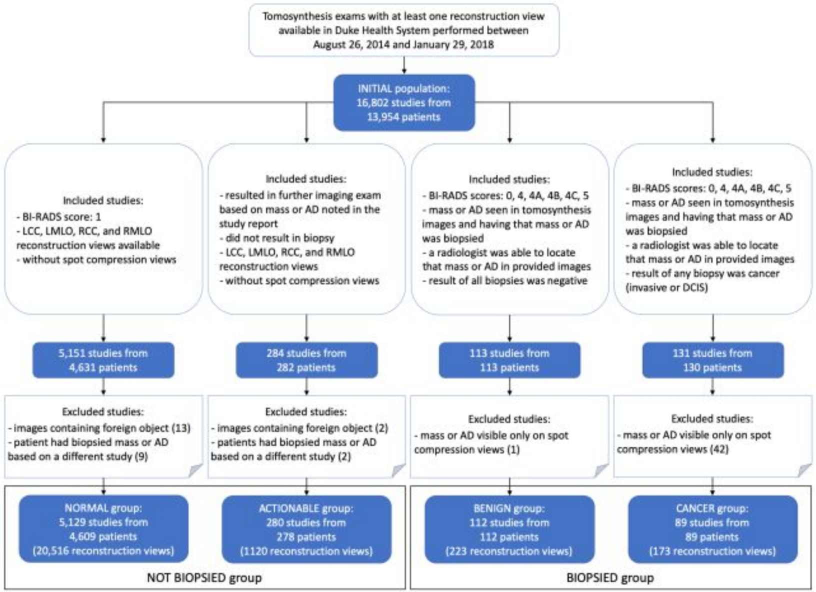


Figure 6. Patient flowchart. BI-RADS = Breast Imaging-Reporting and Data System; AD = architectural distortion; LCC = left craniocaudal; RCC = right craniocaudal; LMLO = left mediolateral oblique; RMLO = right mediolateral oblique [6]

# Open Source Resources

- This CBIS-DDSM (Curated Breast Imaging Subset of DDSM)[7]
  - From the cancer imaging archive
  - Includes a database of over 2,000 mammography studies for ML test and train purposes
  - Must abide by TCIA usage policy
- CBIS-DDSM machine learning tools: https://github.com/fjeg/ddsm_tools/blob/master
  - Utilized to generate test and train datasets
  - License use not available for trademark, liability or warranty
- NYU open-sourced breast cancer screening model: https://github.com/nyukat/breast_cancer_classifier [8]
  - Includes ML and test/train set- reproducible results
  - trained on over 200,000 non DBT mammography exams
  - License not available for liability or warranty
- Breast Cancer Screening – Digital Breast Tomosynthesis (BCS-DBT) [9]
  - Over 5,000 samples of DBT images including normal benign and malignant cases.
  - For ML test and train purposes
  - Creative Commons Attribution-NonCommercial 4.0 International License

# References


[1] Cancer statistics. National Cancer Institute. (2020, September 25). Retrieved September 19, 2021, from https://www.cancer.gov/about-cancer/understanding/statistics.

[2] Tomosynthesis (3D mammography). Stanford Health Care (SHC) - Stanford Medical Center. (2019, May 14). Retrieved September 19, 2021, from https://stanfordhealthcare.org/medical-tests/m/mammogram/tomosynthesis-3d-mammography.html.

[3] Geras KJ, Mann RM, Moy L. Artificial Intelligence for Mammography and Digital Breast Tomosynthesis: Current Concepts and Future Perspectives. Radiology. 2019 Nov;293(2):246-259. doi: 10.1148/radiol.2019182627. Epub 2019 Sep 24. PMID: 31549948; PMCID: PMC6822772.

[4] Moriakov, N., Michielsen, K., Adler, J., Mann, R., Sechopoulos, I., & Teuwen, J. (2019). Deep Learning Framework for Digital Breast Tomosynthesis Reconstruction. Medical Imaging 2019: Physics of Medical Imaging. https://doi.org/10.1117/12.2512912

[5] Bai, J., Posner, R., Wang, T., Yang, C., & Nabavi, S. (2021). Applying deep learning in digital breast tomosynthesis for Automatic Breast Cancer Detection: A Review. Medical Image Analysis, 71, 102049. https://doi.org/10.1016/j.media.2021.102049

[6] Buda, M., Saha, A., Walsh, R., Ghate, S., Li, N., Swiecicki, A., Lo, J. Y., & Mazurowski, M. A. (2021). A data set and deep learning algorithm for the detection of masses and architectural distortions in digital breast tomosynthesis images. JAMA Network Open, 4(8). https://doi.org/10.1001/jamanetworkopen.2021.19100

[7] CBIS-DDSM. CBIS-DDSM - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki. (n.d.). Retrieved September 20, 2021, from https://wiki.cancerimagingarchive.net/display/Public/CBIS-DDSM.

[8] Wiggers, K. (2019, March 21). NYU open-sources breast cancer screening model trained on over 200,000 mammography exams. VentureBeat. Retrieved September 19, 2021, from https://venturebeat.com/2019/03/21/nyu-open-sources-breast-cancer-screening-model-trained-on-over-200000-mammography-exams/.

[9] Breast cancer screening – Digital Breast Tomosynthesis (BCS-DBT). Breast Cancer Screening – Digital Breast Tomosynthesis (BCS-DBT) - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki. (n.d.). Retrieved September 19, 2021, from https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=64685580.