Advancing Breast Cancer Detection with AI and TensorFlow Lite for Real-Time Detection and Improved Patient Outcomes

An Analysis of the Economic Impact and Model Development Plan

Alice Tang and Yinda Chen

Boston College

September 8, 2024

1. Problem Statement

Breast cancer remains a significant global health concern, impacting individuals of all demographics. Most notably, it is one of the most common cancers among women and is the second leading cause of cancer-related deaths in females, following only lung cancer. Detecting breast cancer is a critical and time-sensitive process— early and accurate detection is paramount. Yet current methods, such as manual mammogram analysis, often struggle to identify subtle and inconspicuous malignancy signs. This is particularly problematic in cases with very dense breast tissue, where approximately 50% of breast cancers may go undetected during screenings[1]. Additionally, about a quarter of women with breast cancer are misdiagnosed within two years of screening.

To address these challenges, we propose developing a TensorFlow Lite application that leverages machine learning to improve breast cancer detection. Our objective is to create a tool that complements, rather than replaces, medical professionals. This supplementary tool aims to improve diagnostic accuracy and reduce the time required for analysis. By integrating AI-driven solutions into the diagnostic process, we aspire to create an application that delivers timely and reliable results, thereby enhancing early detection and improving patient outcomes.

2. Articulation of Value

The proposed TensorFlow Lite application offers significant value in enhancing breast cancer detection through machine learning. Key advantages include:

1. **Enhanced Diagnostic Accuracy:** Through machine learning, the application can identify subtle cancer indicators that may be missed in manual examinations— leading to reduced errors and more accurate diagnoses. ML algorithms are particularly effective in healthcare, as they can process large amounts of daily medical data from electronic health records[2], improving the overall diagnostic process.
2. **Timely and Efficient Analysis:** The application streamlines the diagnostic process by providing real-time analysis— lessening diagnosis time. Accelerating diagnosis time alleviates the workload on medical professionals, allowing them to devote more attention to patient care.
3. **Support for Medical Professionals:** Designed to support, not replace, radiologists, this application will play a key role in supporting professionals in making more informed decisions and enhancing overall diagnostic confidence.

---

[1] Rabiei, Reza. “Prediction of Breast Cancer using Machine Learning Approaches.” *Journal of Biomedical Physics and Engineering*, vol. 12, no. 3, July 2022, https://doi.org/10.31661/jbpe.v0i0.2109-1403.
[2] Jafari, Ali. “Machine-learning methods in detecting breast cancer and related therapeutic issues: a review.” *Computer Methods in Biomechanics and Biomedical Engineering Imaging & Visualization*, Jan. 2024, pp. 1–11. https://doi.org/10.1080/21681163.2023.2299093.

4. **Improved Patient Outcomes:** By enabling early and accurate detection, this application paves the way for prevention and treatment, which is crucial for effective care. This will inherently improve patient outcomes and well-being.
5. **Addressing Screening Gaps:** This application is especially useful for detecting cancers in dense breast tissue, which traditional methods may be less reliable. By tackling this specific challenge, it allows a wider range of patients to receive accurate diagnostic practices.

Overall, our application will be a valuable asset to the breast cancer detection process. It provides a reliable, efficient, and supportive solution for medical professionals. By implementing this application, we strive to improve diagnostic accuracy, reduce analysis time, and contribute to improved patient outcomes through timely and effective treatment.

3. Calculation of Economic Value

In essence, implementing our TensorFlow Lite application is anticipated to generate significant economic benefits. This includes an estimated **per-patient cost saving of $31,737.39 per woman** and a **reduction in societal losses of $186.02 million**. These figures underscore the potential impact of the TensorFlow Lite application in improving breast cancer detection while reducing the associated economic burden.

3.1 Assumptions

- **Prevalence of breast cancer:** 13% of women will develop breast cancer in their lifetime.[3]
- **Current detection rate:** 87% of breast cancers are detected by mammograms.[4]
- **False positive rate:** 17% of women undergoing annual 3D mammography have at least one false positive over 10 years.[5]
- **False negative rate:** 12.5% of breast cancers are missed by mammograms.[6]

---

[3] "Breast Cancer Statistics | How Common Is Breast Cancer?" *American Cancer Society*, www.cancer.org/cancer/types/breast-cancer/about/how-common-is-breast-cancer.html.

[4] "Mammogram Accuracy - Accuracy of Mammograms." *Susan G. Komen*, 9 Dec. 2022, www.komen.org/breast-cancer/screening/mammography/accuracy.

[5] "Half of all women experience false positive mammograms after 10 years of annual screening." *News*, 4 June 2023, health.ucdavis.edu/news/headlines/half-of-all-women-experience-false-positive-mammograms-after-10-years-of-annual-screening-/2022/03

[6] "Limitations of Mammograms | How Accurate Are Mammograms?" *American Cancer Society*, www.cancer.org/cancer/types/breast-cancer/screening-tests-and-early-detection/mammograms/limitations-of-mammograms.html.

- **Time saved:** 30 minutes per mammogram using traditional methods[7] - 5 minutes using our TFLite app (current goal) = 25 minutes saved.
- **Cost of medical professionals' time (specifically radiologists):** $353,960 annual salary, $170.17/hour on average[8].
- **Cost of false positives:** $852 per false positive[9].
- **Incremental per-patient lifetime burden of FP results:** $58,900[10].
- **Incremental per-patient lifetime burden of FN results:** $116,000[10].
- **Estimated societal loss based on FP results:** $417 million[10].
- **Estimated societal loss based on FN results:** $575 million[10].

3.2 Calculation Methodology

Direct Cost Savings Calculations:

- **Reduced false positives:** (13% * 17%) * $58,900 = **$12,880.30 per woman**
  - 13%: Represents the prevalence of breast cancer, where 13% of women will develop the disease in their lifetime.
  - 17%: The estimated FP rate for mammograms.
  - $58,900: Incremental per-patient lifetime burden of false positive results. This includes any costs associated with unnecessary follow-up procedures or potential surgeries.

By reducing the false positive rate, our application can significantly decrease the number of women who undergo unnecessary procedures, leading to substantial cost savings.

- **Reduced false negatives:** (13% * 12.5%) * $116,000 = **$18,850 per woman**
  - 13%: Represents the prevalence of breast cancer.
  - 12.5%: This is the estimated false negative rate for mammograms, indicating that 12.5% of women with breast cancer will have a negative mammogram result.

---

[7] "What You Need to Know About Mammograms." *Beth Israel Lahey Health*, 2024, www.bidmc.org/centers-and-departments/cancer-center/cancer-center-programs-and-services/breastcare-center/programs-and-services/breast-imaging-and-procedures/mammograms/faqs-about-mammograms. Accessed 8 Sept. 2024.

[8] "May 2023 National Occupational Employment and Wage Estimates." U.S. Bureau of Labor Statistics, U.S. Bureau of Labor Statistics, May 2023, www.bls.gov/oes/current/oes_nat.htm#29-0000. Accessed 8 Sept. 2024.

[9] Ong, Mei-Sing, and Kenneth D. Mandl. "National Expenditure For False-Positive Mammograms And Breast Cancer Overdiagnoses Estimated At $4 Billion A Year." *Health Affairs*, vol. 34, no. 4, Apr. 2015, pp. 576–83. https://doi.org/10.1377/hlthaff.2014.1087.

[10] Garrison, Louis P., et al. "The Lifetime Economic Burden of Inaccurate HER2 Testing: Estimating the Costs of False-Positive and False-Negative HER2 Test Results in US Patients with Early-Stage Breast Cancer." *Value in Health*, vol. 18, no. 4, June 2015, pp. 541–46. https://doi.org/10.1016/j.jval.2015.01.012.

    - $116,000: Incremental per-patient lifetime burden of false negative results. This pertains to the costs associated with delayed diagnosis, more aggressive treatments, and potentially poorer outcomes.

In reducing the false negative rate, our application can improve early detection— leading to less invasive treatments and better patient outcomes. Big picture-wise, this results in substantial cost savings over the patient's lifetime.

<u>Indirect Cost Savings Calculations:</u>

- **Time saved:** 25 minutes/mammogram * $170.17/hour = **$7.09 per mammogram**
    - 25 minutes: This is the estimated time saved per mammogram by using the TensorFlow Lite application, calculated by 30 minutes (using traditional methods) - 5 minutes (goal for TensorFlow Lite detection).
    - $170.17/hr: This is the mean hourly wage for radiologists.

By lessening the time required for mammogram analysis, our application provides a pathway for radiologists to focus on other tasks or see more patients— manifesting in increased productivity and potentially lower healthcare costs.

<u>Total Economic Value:</u>

- **Total economic value** = $12,880.30 + $18,850 + $7.09 = **$31,737.39 per woman**
    - $12,880.30: The calculated direct cost savings from reduced false positives.
    - $18,850: The calculated direct cost savings from reduced false negatives.
    - $7.09: The calculated indirect cost savings from time saved by radiologists.

The total economic value per woman represents the combined direct and indirect benefits of using the TensorFlow Lite application. This calculation quantifies the potential savings in healthcare costs and improved outcomes for each patient, as doctors utilize the application.

<u>Societal Losses:</u>

- **Reduced societal loss from FP results:** (13% * 17%) * $417 million = **$92.27 million**
    - 13%: Represents the prevalence of breast cancer.
    - 17%: Estimated false positive rate for mammograms.
    - $417 million: Estimated societal loss per woman due to false positive results. This includes costs related to unnecessary procedures, anxiety, and potential negative psychological impacts.
- **Reduced societal loss from FN results:** (13% * 12.5%) * $575 million = **$93.75 million**
    - 13%: Represents the prevalence of breast cancer.
    - 12.5%: Estimated false negative rate for mammograms.

    - $575 million: Estimated societal loss per woman due to false negative results. This includes costs related to delayed diagnosis, more aggressive treatments, and potentially poorer health outcomes.
- **Total reduced societal loss:** $92.27 million + $93.75 million = **$186.02 million**
    - $92.27 million: Estimated reduction in societal losses due to fewer false positive results from the application.
    - $93.75 million: Estimated reduction in societal losses due to fewer false negative results from the application.

4. Dataset Description

The DDSM Mammography dataset, which we selected from Kaggle, consists of images from the DDSM and CBIS-DDSM datasets. This dataset was chosen because it is well-detailed and provides ample data for breast cancer classification. These images were pre-processed to 299x299 pixels by extracting regions of interest (ROIs). The dataset contains 55,890 training examples, with 14% positive (cancerous) cases and 86% negative. The images are divided into training, test, and validation sets, though the test and validation sets need to be combined for balanced data. Labels include binary classification for normal/abnormal and multi-class labels for benign or malignant masses and calcifications. However, we recognize that the class imbalance may present challenges, and we will need to address this issue during our analysis accordingly.

5. Modeling Approach

We plan to use YOLO (You Only Look Once) to solve this problem for several reasons, each contributing to its suitability for our specific needs and objectives:

1. **Real-Time Object Detection and Classification:** YOLO is known for its high speed and efficiency in processing images in real-time, making it ideal for applications requiring quick analysis, like ours. Additionally, YOLO has been used in previous breast cancer detection studies[11], showing promising results in identifying abnormalities quickly and accurately. This makes it a strong contender for our application.
2. **Simultaneous Detection and Classification:** Our project involves multi-class classification (negative, benign calcification, benign mass, malignant calcification, malignant mass), and YOLO is well-suited for handling multiple classes efficiently.
3. **Localization of Abnormalities:** In medical imaging, especially mammography, accurately and precisely detecting the exact location of abnormalities is crucial. YOLO's ability to provide

[11] Prinzi, Francesco, et al. "A Yolo-Based Model for Breast Cancer Detection in Mammograms." *Cognitive Computation*, vol. 16, no. 1, Aug. 2023, pp. 107–20. https://doi.org/10.1007/s12559-023-10189-6.

specific bounding boxes around detected objects helps us highlight specific regions of interest (ROIs) in the mammograms.

This model operates under **supervised learning** because it relies on labeled data to learn patterns and make predictions. The labeled data includes both labeled bounding boxes and classification labels—allowing YOLO to learn how to detect and classify objects in new images. After examining the data, we identified that the images fall under two types of labels:

1. **label_normal:** 0 for negative and 1 for positive label
2. **Full Multi-Class Labels:** 0 for negative, 1 for benign calcification, 2 for benign mass, 3 for malignant calcification, 4 for malignant mass.

Thus, this problem is a **multi-class classification problem** where the model needs to classify certain regions of the mammogram images based on the presence of abnormalities like calcifications or masses.

However, if YOLO does not meet our performance expectations, we're open to exploring other models such as CNNs like ResNet and VCG-16, which have demonstrated success in outperforming radiologists in breast cancer detection[12][13]. In any case, given our environmental constraints, we will need to strike a delicate balance between accuracy, recall, and speed to ensure we select the most effective model for our needs.

6. Thirteen-Week Project Plan

Week 1: Identify the Problem Statement and Dataset

- **Objective:** Clearly define the project's purpose and select a suitable dataset.
- **Finalize Problem Statement:** Develop a machine learning model to detect breast cancer in mammography images to assist medical professionals in early diagnosis.
- **Dataset Selection:** Use the DDSM Mammography dataset from Kaggle as the primary data source. We'll explore alternative datasets to supplement the primary dataset, specifically images that are "positive" due to imbalanced classes.
- **Designate Success Metrics:** We must define the potential key performance metrics— accuracy, precision, recall, F1 score, and AUC-ROC. Due to an imbalanced dataset, "accuracy" may not be

---

[12] Adam, Richard, et al. "Deep learning applications to breast cancer detection by magnetic resonance imaging: a literature review." *Breast Cancer Research*, vol. 25, no. 1, July 2023, https://doi.org/10.1186/s13058-023-01687-4.

[13] Shen, Li, et al. "Deep Learning to Improve Breast Cancer Detection on Screening Mammography." *Scientific Reports*, vol. 9, no. 1, Aug. 2019, https://doi.org/10.1038/s41598-019-48995-4.

the most holistic choice in determining model performance. “Recall” and “AUC-ROC” will be critical metrics because a false negative case will have severe consequences in cancer detection.

- **Project Documentation:** Create this initial documentation detailing project objectives, team roles, and timeline. Additionally, create the GitHub repository and set up the JupyterHub.

Week 2: Ingest and Explore the Dataset

- **Objective:** Familiarize ourselves with the data and understand the structure.
- **Data Ingestion:** Ingest the DDSM dataset from Kaggle. Verify the integrity of the data and check for missing or corrupted files.
- **Explore the Dataset:** Understand and summarize data attributes (e.g., image formats, labels) and generate sample mammogram images, along with their labels (benign or malignant) to gain better insights on if additional preprocessing is required. If so, we’ll clean the initial data by removing invalid images and checking label consistency.

Week 3: Perform Exploratory Data Analysis (EDA)

- **Objective: Gain a deeper understanding of the dataset and identify any potential challenges.**
- **Visualization:** Create visualizations to explore relationships between different features, and understand class distributions (e.g., benign vs malignant cases).
- **Data Quality Analysis:** Check for class imbalances and assess the severity, examine pixel intensity distributions, image resolutions, and detect anomalies or outliers in the dataset.
- **Hypothesis Generation:** Formulate a hypothesis about which image characteristics may contribute to cancer detection.

Week 4: Make Data Model-Ready

- **Objective:** Prepare the data for modeling by addressing preprocessing needs
- **Data-preprocessing:** Handle any missing data, perform noise reduction, resize images, and normalize pixel values as needed. Employ label encoding techniques and prepare annotation files for YOLO (bounding boxes, object labels).
- **Data Splitting:** Split the dataset into training, validation, and test sets, ensuring balanced class distribution. Implement techniques such as SMOTE or undersample the majority class.
- **Image Augmentation:** Improve the generalization of the model by augmenting the dataset with techniques like flipping, rotating, and scaling.

Week 5: Engineer Features

- **Objective:** Define important features such as pixel intensity, texture, and shape and create new features that will enhance the model performance. Explore different pre-trained models to extract features to improve performance and enhance efficiency.
- **Feature Extraction:** Use feature engineering to extract key visual features, such as edges, contours, and masses.
- **Data Processing and Optimization:** Implement additional data augmentations specifically tailored to mammography images, like contrast enhancement. Employ techniques like Principal Component Analysis (PCA) to reduce the complexity of feature space.
- **Feature selection:** Evaluate the importance of features and retain only those that contribute to model accuracy.

Week 6: Develop 1st Modeling Approach (Simple, the Baseline)

- **Objective:** Establish a baseline model for comparison means.
- **Model Selection:** Implement a basic YOLO model (e.g., YOLOv5).
- **Training:** Train the model on a subset of the dataset with minimal hyperparameter tuning.
- **Evaluation:** Evaluate model performance using simple metrics such as accuracy and loss curves. Then, analyze the initial results to identify issues or areas for improvement.
- **Documentation:** Document the baseline performance and any insights learned from the model's limitations.

Week 7: Develop 2nd Modeling Approach (More Complex)

- **Objective:** Improve upon the baseline model with a more advanced approach.
- **Model Selection:** Implement and fine-tune a pre-trained YOLO model (transfer learning with a pre-trained backbone).
- **Hyperparameter Tuning:** Adjust hyperparameters such as learning rate, batch size, number of layers, activation functions, and epoch number. Consider applying better-bounding box selection methods (e.g., using IoU thresholding).
- **Model Evaluation:** Evaluate model performance on the validation set using precision, recall, and F1-score. Compare the performance l to the baseline, and identify any areas of improvement.

Week 8: Develop 3rd Modeling Approach (Even More Complex)

- **Objective:** Experiment with additional YOLO variants (e.g., YOLOv7 or YOLOv8) or try ensemble methods combining YOLO with other object detection techniques.

- **Experimentation:** Try more advanced augmentation techniques such as CutMix or Mosaic augmentation. Experiment with adding additional layers to capture more fine-grained features, such as smaller abnormalities.
- **Model Evaluation:** Compare the performance of this advanced model against the previous two and assess the suitability of deployment.

Week 9: Select the Winning Model

- **Objective:** Compare all trained models on key performance metrics such as accuracy, precision, recall, F1 score, and AUC-ROC.
- **Model Selection:** choose the model that provides the best trade-off between accuracy, interpretability, and performance for real-time deployment on **mobile**.

Week 10: Data-Centric AI

- **Objective:** Improve model performance by focusing on data quality and preprocessing.
- **Refine the Data:** Re-examine the dataset to further improve preprocessing steps, such as including additional augmentation.
- **Steps to Take:** Consider more sophisticated data augmentation techniques in creating synthetic data (GANs), apply additional noise reduction techniques to improve image quality, and re-examine labels to ensure they are not mislabeled.

Week 11: Explain the Model, Analyze risk, Bias, and Ethical Considerations

- **Objective:** Ensure the model is built on a transparent, unbiased, and ethically sound foundation.
- **Risk Assessment:** Analyze potential risks such as false negatives (missed cancer, delaying treatment) or false positives (unnecessary anxiety).
- **Bias Analysis:** Identify any biases in the model, such as biases related to patient demographics (age, gender, ethnicity) or image quality.
- **Ethical Considerations:** Discuss ethical considerations, such as the app's impact on medical diagnoses and the need for human oversight.

Week 12: Save and Package Your Model for Deployment.

- **Objective:** Convert the selected model to TensorFlow Lite format for mobile deployment.
- **Deployment Testing:** Test the model on mobile devices for performance and latency checks.
- **Monitoring Plan:** Develop a model monitoring plan for ongoing performance tracking, especially after deployment (e.g., detecting model drift).

- **User Interface:** Build a simple user interface for the app that allows users to upload and analyze images.

Week 13: Bring It All Together

- **Objective:** Finalize the project and summarize findings.
- **End-to-End Testing:** Conduct end-to-end testing of the app to ensure user experience and model works together seamlessly.
- **Future Considerations:** Identify problem areas for future improvement or research, potentially expanding the dataset or refining the model further.