Case Study: Risk Assessment Model for Cancer Diagnosis

Background

Breast cancer is one of the most common cancers worldwide, affecting millions of women each year. It arises from the breast tissue, often presenting as a lump or abnormality. Early diagnosis is critical for improving survival rates and treatment outcomes. Despite advances in screening, there is a need for more precise and personalized risk assessment models to enhance early detection and management.

Risk assessment models play a pivotal role in breast cancer care by identifying individuals at higher risk based on various factors.

Traditionally, breast cancer diagnosis involves imaging techniques such as mammography, ultrasound, and magnetic resonance imaging (MRI), often followed by a biopsy to check if the cancer is present.

Early models like the Gail model estimated breast cancer risk based on factors such as age, family history, and reproductive history. Although foundational, these models had limitations in accuracy and did not account for genetic factors.

Recent advancements include models that integrate genetic information, such as the Tyrer-Cuzick model, which uses genetic markers alongside clinical data.

To better individualize breast cancer screening and more accurately predict the risk of breast cancer development, researchers from Massachusetts General Hospital (MGH) and Massachusetts Institute of Technology (MIT) teamed together to develop a new precision medicine tool. This tool was designed to identify the specific risk of each patient based on their mammogram findings.

In 2019, a team of scientists from MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) and Jameel Clinic demonstrated a deep learning system to predict cancer risk using just a patient's mammogram. Constance Lehman, MD, PhD, Chief of Breast Imaging at Massachusetts General Hospital, collaborated with colleagues from MIT's Computer Science and Artificial Intelligence Laboratory.

Researchers at Mass General used artificial intelligence (AI)-based technology to analyze patients' mammogram images against a database of over 200,000 other mammograms to identify patients who had higher risk of developing breast cancer. The AI component of this technology allowed for a more complex analysis than a human could perform while simultaneously allowing each mammogram to be almost instantly compared to far more images than would ever be practical using human evaluation.

Also used were other personalized data, such as age and hormonal factors, as part of the algorithm, but the researchers made these other data optional to allow the AI technology to be used in the widest variety of clinical sites. Researchers also optimized the AI tool to work with a variety of imaging machines and different environments so that it could be more widely applied.

While two years of research has gone into developing this AI tool, dubbed "Mirai", the technology has only recently been integrated into actual clinical practice, with MGH being the first adopter of this new technology.

Implementation

Mirai was trained on a large dataset of over 200,000 mammograms from Massachusetts General Hospital (MGH) in the United States and tested on held-out test sets from MGH, Karolinska University Hospital in Sweden, and Chang Gung Memorial Hospital (CGMH) in Taiwan.

It takes as input all standard views of a mammogram: left craniocaudal (L CC), left mediolateral-oblique (L MLO), right craniocaudal (R CC), and right mediolateral-oblique (R MLO). Mirai consists of four modules:

1. **Image Encoder**: Uses ResNet-18 to encode each mammogram view into a 512-dimensional vector.
2. **Image Aggregator:** Collects and processes mammographic images from different views to create a detailed breast representation using Transformer network.
3. **Data Aggregation:** Combines mammogram data with patient-specific factors, such as age and family history, to enhance the model's prediction capabilities.

4. **Risk-Factor Prediction:** Estimates risk factors associated with breast cancer based on the aggregated data.
5. **Additive-Hazard Layer:** Synthesizes all data inputs to predict the likelihood of developing breast cancer over the next five years.

In the study, mammograms were captured using Hologic Selenia or Selenia Dimensions devices. The images were converted from DICOM to PNG16 using the DCMTK library. Python libraries torchvision and Pillow used for image preprocessing and data augmentations.

Mirai was trained in two phases. First, the image encoder, risk factor predictor, and additive hazard layer were trained to predict breast cancer independently without adversarial training. The image encoder was initialized with ImageNet weights, and the training set was augmented with image flips and rotations. In the second phase, the image encoder was frozen, and the rest of the model was trained using conditional adversarial training. Hyperparameter searches were conducted, and the model with the highest C-index on the development set was chosen.

The impact of predicted risk factors on Mirai's performance was assessed by comparing two scenarios: "Mirai with risk factors" (using electronic health record-based risk factors) and "Mirai without risk factors" (using predicted risk factors). Mirai was evaluated against three other models: Hybrid DL (using both mammograms and risk factors), Image-Only DL (using only mammograms), and TCv8 (a traditional risk model). Discrimination was assessed using a matched area under the curve (AUC) statistic.

Mirai outperformed other algorithms in risk assessment (AUC 0.64–0.74 for large cancers) and detection (AUC 0.81–0.89). It excelled in identifying larger cancers (30 mm+), with detection AUCs of 0.88–0.92, and showed relatively strong performance for smaller cancers (0 to < 10 mm) with a risk AUC of 0.64, compared to 0.54–0.56 for other models. Mirai's ability to assess risk and detect cancer sizes made it the most effective among the tested algorithms.

Results

The AI models developed at Mass General significantly outperform traditional risk models, especially for women of color, who face a 40% higher mortality rate from breast cancer. Traditional models, developed primarily on European Caucasian populations, performed worse than chance for women of color. Mirai showed no racial bias and was effective across diverse patient groups, enabling personalized screening strategies. It helped identify women most in need of immediate screening, including those with a history of breast cancer.

The Mirai model achieved an accuracy rate of 89.3% in identifying future breast cancer cases, surpassing traditional models like the Tyrer-Cuzick model.

The model reduced false negatives by 63%, ensuring that fewer cases of breast cancer were missed.

Impact

The impact of the Mirai model on breast cancer detection and overall patient care has been substantial:

- Early Detection: The model's ability to predict breast cancer risk up to five years in advance has increased early detection rates by 35%, allowing for earlier interventions and improving patient outcomes.
- Reduction in Unnecessary Biopsies: By providing more accurate risk assessments, Mirai has helped reduce unnecessary biopsies by 27%, minimizing patient stress and reducing healthcare costs.
- Efficiency in Radiology: Mirai's high accuracy has reduced the workload for radiologists by 20%, enabling them to focus on high-risk cases and improving overall efficiency in breast cancer screening programs.