Predicting Breast Cancer Using Sequential Mammogram Analysis

Breast cancer remains a significant global health concern, with mammography serving as a primary screening tool. Despite advancements in imaging technology, early detection of malignancies remains challenging due to the limited contrast of lesions and high rates of false negatives. Traditional assessments rely on radiologist expertise, often requiring multiple reviews, which increases cost and resource burden. Recent research explores the potential of computational approaches, such as machine learning, to enhance early detection. This study investigates the effectiveness of temporal subtraction of sequential mammograms combined with advanced classification techniques to predict near-term breast cancer occurrence, improving detection before visible abnormalities appear.

The ability to predict malignancies before they become radiologically evident can greatly improve patient outcomes. The current approach in clinical practice involves comparing prior mammograms to recent scans manually, which is both time-intensive and prone to subjective interpretation. Moreover, high breast density can obscure lesions, further complicating early detection efforts. The integration of machine learning techniques with temporal subtraction aims to address these challenges by offering an automated and objective approach to identifying changes indicative of future malignancy.

Methodology and Data Collection

The study collected three consecutive rounds of digital mammograms from 75 women aged 46 to 79. Each screening round was separated by an average interval of two years. A total of 450 images were acquired, with the most recent mammogram serving as the reference for biopsy-confirmed malignancies. The two preceding normal mammograms were processed using temporal subtraction to identify subtle changes in breast tissue. The regions of interest (ROIs) were segmented and classified using a machine-learning approach. Feature selection was performed to identify the most predictive markers and various classifiers were applied, including support vector machines, k-nearest neighbours and ensemble voting.

Must Read:* *AI Trial for Breast Cancer Detection in the NHS

A crucial aspect of the study design was ensuring the dataset's integrity and clinical relevance. Cases were selected based on strict inclusion criteria to ensure that only patients with two normal prior mammograms and a confirmed malignant mass in the most recent screening were included. Ethical approval was obtained, and biopsy confirmations were conducted following standard medical protocols. This rigorous approach ensured that the dataset was both representative and reliable for training and validating the predictive model.

Processing and Analysis of Mammograms

A crucial step in the analysis was the identification of potential malignancy locations before they became visible in the most recent mammogram. Preprocessing techniques such as normalisation, contrast-limited adaptive histogram equalisation and gamma correction were used to enhance image quality. The "future" mammogram was registered to the "current" image to pinpoint the likely location of the emerging malignancy. The subtraction of the two previous mammograms facilitated the identification of new developments over time. Machine learning models were trained to differentiate between benign and possible future malignancies using extracted features, including texture and intensity-based metrics. The highest classification accuracy was achieved using ensemble voting, with 98.8% accuracy, 93.6% sensitivity and 98.8% specificity.

The feature extraction process focused on 98 parameters, including shape-based, intensity-based and statistical measures derived from the images. Eight feature selection techniques were used to identify the most significant predictors of malignancy. The implementation of advanced classification techniques, such as leave-one-patient-out cross-validation, ensured that the model's predictive accuracy was robust and generalisable. The study also incorporated adaptive synthetic sampling to address class imbalances and enhance classification performance.

Results and Performance Evaluation

The study demonstrated that temporal subtraction significantly improved the ability to predict future breast malignancies before radiological signs became evident. The contrast ratio of detected changes doubled compared to the original mammograms. Feature selection revealed 14 optimal characteristics that best distinguished between benign and malignant regions. Among various classification models, ensemble voting provided the most reliable results. The study also compared this approach to deep learning-based prediction models and found that, despite smaller datasets, the feature-based method outperformed deep learning models trained on limited data. The proposed method exhibited a high level of robustness and generalisability across different patient cases, making it a viable approach for clinical application.

Statistical analysis confirmed the superiority of ensemble voting compared to other classifiers. The leave-one-patient-out cross-validation approach ensured that the results were not biased by patient-specific variations. The performance metrics indicated that the proposed methodology could significantly enhance the current screening process by reducing false negatives and improving early detection. Despite these promising findings, the study acknowledged the limitations of a relatively small dataset and highlighted the need for further validation using larger cohorts.

The study highlights the potential of using temporally sequential mammogram subtraction combined with machine learning to predict breast cancer before visible abnormalities emerge. The high accuracy of the model suggests that such an approach could significantly enhance early detection efforts and reduce false-negative rates. While additional validation using larger datasets is necessary, the methodology provides a promising avenue for improving breast cancer screening and patient outcomes. Future research should explore the integration of digital breast tomosynthesis and deep learning models with larger sample sizes to further refine prediction accuracy and clinical applicability.

The successful implementation of this methodology in clinical practice could revolutionise breast cancer screening by offering an automated and reliable tool for early detection. By identifying patients at risk before a mass is visible on mammograms, this approach could enable earlier interventions, potentially reducing mortality rates. The study provides a strong foundation for further research into integrating AI-driven tools with existing radiological workflows, paving the way for a more precise and efficient diagnostic process.

Source: **Journal of Imaging Informatics in Medicine**

Image Credit: **iStock**

Published on : Mon, 17 Mar 2025