Deep Learning Mammography-based Age Prediction:
Insights into model performance and demographic disparities

Theo Dapamede, MD, PhD[1], Bardia Khosravi, MD, MPH, MHPE[2], Mohamammadreza Chavoshi, MD[1], Frank Li, PhD[1], Hari Trivedi, MD[1], Judy W. Gichoya, MD, MS[1]

[1]Emory University, [2]Mayo Clinic

Background

Age prediction from mammograms using deep learning has potential applications in identifying patients with accelerated aging or increased breast cancer risk, which could inform personalized screening strategies and risk assessment.

We developed and evaluated deep learning models for age prediction using bilateral and unilateral mammograms and investigated model performance across various demographic subgroups.

Methods

A dataset of 84,968 mammograms from the EMBED dataset (71,816 training; 13,152 testing) was used. The mean age was 57.0 years for both sets. ConvNext-base CNN and DenseNet-121 models were trained on resized (512x512) and padded images to predict age using unilateral (left, right) and bilateral mammograms.

Subgroup analyses were performed on race, marital status, tissue density, and BI-RADS.

The penultimate model features were used to predict these demographic factors to see if these are considered by the model when predicting patients' age.

Discussion

We show that deep learning models can predict age from mammograms with high accuracy, and bilateral imaging yields superior results compared to unilateral. However, the models exhibit demographic disparities, with lower performance in African-American and non-married patients.

Importantly, the penultimate features of the age prediction model can predict sensitive demographic information such as race and breast density with high AUCs, indicating that the model considers these factors during age prediction.

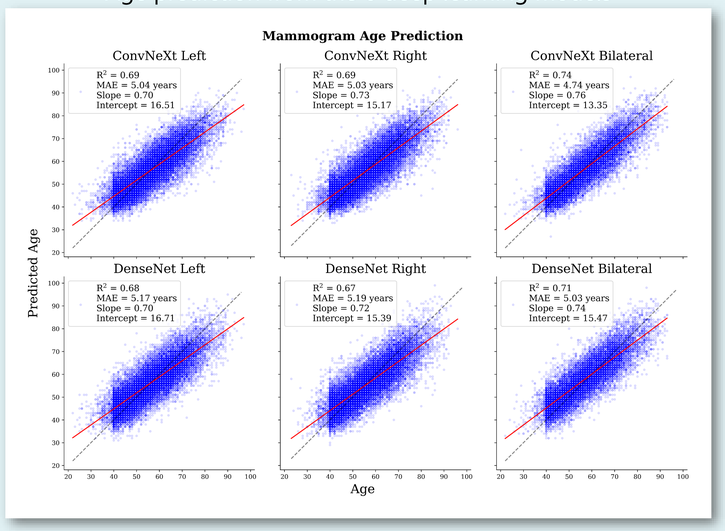
This finding is alarming as it suggests potential bias in the model's predictions. Furthermore, the presence of pathology acts as a confounder, decreasing the model's performance in age prediction.

EMORY UNIVERSITY SCHOOL OF MEDICINE
Department of Radiology and Imaging Sciences

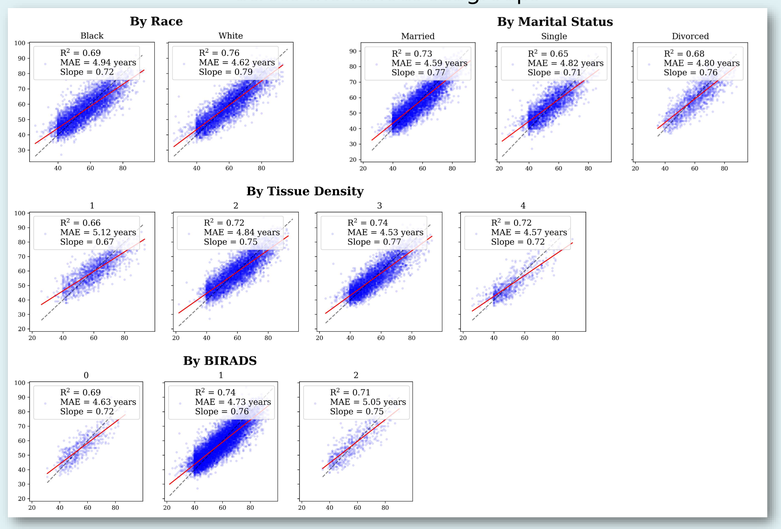
HITILAB

Demographic disparities in model performance and the model's ability to **predict sensitive patient information** raise significant **concerns about the fairness and transparency** of deep learning models.

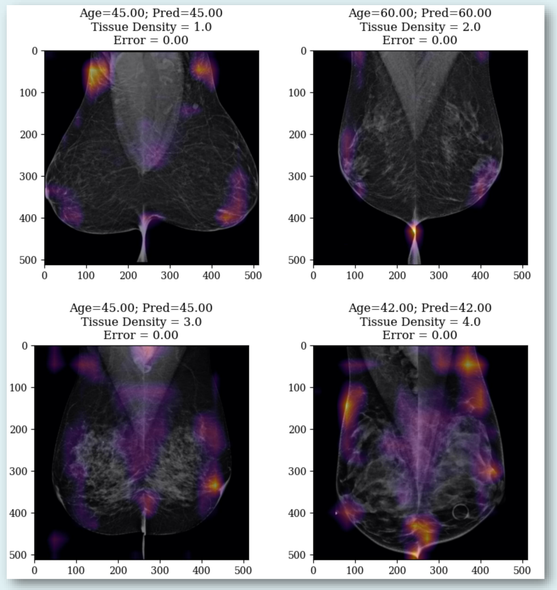
1 Age prediction from the 6 deep learning models



2 Age prediction of the ConvNext-Bilateral model across the various subgroups



3 GradCAM++ overlaid on mammograms of the correct age prediction across tissue densities



4 Penultimate feature layer-based prediction of demographic factors, breast density and BIRADS