RETFound+: A Multi-Modal Deep Learning Framework for Osteoporosis Screening via Retinal Imaging

Hongyu He[a], **Dianbo Liu**[a]

[a] *National University of Singapore, Singapore 119077* hongyu.h@nus.edu.sg, dianbo@nus.edu.sg

* Presenting author

1. Introduction

Early detection of age-related conditions, particularly osteoporosis, remains a significant challenge in healthcare, with current screening methods often being costly, invasive, or inaccessible to many populations. Recent advances in deep learning and medical imaging analysis suggest that retinal images may contain valuable biomarkers for various systemic conditions[1]. We hypothesized that features extracted from retinal images using the RETFound model could serve as effective biomarkers when combined with demographic information for osteoporosis screening.

2. Methods

We analyzed retinal images from the UK Biobank dataset using the RETFound deep learning model, a foundation vision transformer architecture originally designed for general retinal image analysis [1, 2]. The model was trained to extract meaningful features that could serve as potential biomarkers. Our approach integrates image features with tabular patient data to enable non-invasive early detection of osteoporosis risk.

2.1 Dataset

Using a carefully balanced dataset of retinal images with matched sampling based on age and gender, we ensured robust model training while controlling for key demographic confounders. The data set was divided into training and testing sets with an 8:2 ratio, maintaining class balance in both sets.

2.2 Model Architecture

Our framework consists of three main components: a RETFound-based visual feature extractor with partially frozen weights, a tabular data processing branch using a multi-layer perceptron (MLP), and a fusion module combining visual and tabular features for final prediction.

2.3 Evaluation

Model performance was evaluated using ROC analysis, sensitivity, specificity, and accuracy metrics. Three different thresholding methods (Youden, Distance, Default) were compared for optimal performance. Additionally, we compared our approach with the Osteoporosis Self-Assessment Tool (OST), a widely adopted screening method in clinical practice[3].

3. Results

The combined model demonstrated strong predictive performance with an AUC of 0.76. Using the Youden threshold (0.54), the model achieved a sensitivity of 0.87 and specificity of 0.57, with an overall accuracy of 0.72. The confusion matrix showed 46 true positives and 31 true negatives in the test set.

Our multi-modal approach achieved superior performance (AUC 0.76) compared to both image-only (AUC 0.62) and tabular-only baselines (OST score, AUC 0.60). Notably, the integration of demographic data with imaging biomarkers significantly improved model performance compared to using either data source alone.

External validation using PIONEER and Singapore datasets is ongoing to confirm generalizability across diverse populations.

4. Conclusions

Our study successfully identified novel retinal imaging biomarkers using the RETFound deep learning model. The strong performance metrics suggest these biomarkers, when combined with demographic information, could serve as valuable tools for assessing osteoporosis risk.

The model successfully identifies retinal biomarkers associated with osteoporosis while maintaining robustness across demographic groups. Our results suggest that deep learning-based retinal image analysis, when combined with basic patient metadata, could serve as an effective screening tool for osteoporosis, potentially enabling earlier intervention and improved patient outcomes.

Future work will focus on external validation, investigation of the biological mechanisms underlying these imaging biomarkers, and prospective clinical studies to evaluate real-world utility.

Acknowledgments

We would like to acknowledge the UK Biobank for providing the dataset used in this study.

References

[1] Yukun Zhou, Mark A. Chia, Siegfried K. Wagner, Murat S. Ayhan, Dominic J. Williamson, Robbert R. Struyven, Timing Liu, Moucheng Xu, Mateo G. Lozano, Peter Woodward-Court, Yuka Ki-

hara, Andre Altmann, Aaron Y. Lee, Eric J. Topol, Alastair K. Denniston, Daniel C. Alexander, and Pearse A. Keane. A foundation model for generalizable disease detection from retinal images. *Nature*, 622(7981):156–163, October 2023.

[2] Cathie Sudlow, John Gallacher, Naomi Allen, Valerie Beral, Paul Burton, John Danesh, Paul Downey, Paul Elliott, Jane Green, Martin Landray, Bette Liu, Paul Matthews, Giok Ong, Jill Pell, Alan Silman, Alan Young, Tim Sprosen, Tim Peakman, and Rory Collins. Uk biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. *PLoS medicine*, 12(3):e1001779, March 2015.

[3] L. K. H. Koh, W. B. Sedrine, T. P. Torralba, A. W. Kung, S. Fujiwara, S. P. Chan, Q. Huang, R. Rajatanavin, K. Tsai, H. M. Park, and J. Y. Reginster. A simple tool to identify asian women at increased risk of osteoporosis. *Osteoporosis International*, 12:699–705, 2001.

Appendix A. Model Performance Details

A detailed comparison of different models and settings is provided in Table A1.

Table A1: Performance Comparison of Different Models and Settings

Model	**Test Cohort**	**Additional Data**	**AUC**
OST Score	Balanced	—	0.54
OST Score	Unbalanced	—	0.60
RETFound	Balanced	Images only	0.62
RETFound	Unbalanced	Images only	0.62
RETFound	Unbalanced	Images + Demographics	0.76