Attention-Guided Fair AI Modelling for Skin Cancer Diagnosis

Mingcheng Zhu[1], Mingxuan Liu[1], Han Yuan[1], Nan Liu[1].

Center for Quantitative Medicine, Duke-NUS Medical School, Singapore

Background

- AI experts for skin cancer diagnosis are increasingly being integrated into clinical practice. However, AI might show gender-based biases due to pathophysiological differences.
- Existing works focus on ethnical biases, while the gender-based biases of AI experts for skin cancer diagnosis has not been explored.

Objectives

- Examine and mitigate the gender-based biases of AI experts for skin cancer diagnosis.
- Understand why there is gender-based biases of AI for skin cancer diagnosis.

Materials & Method

Table 1. *Characteristics of datasets*

	Average Age	Positive Rate	Count
HAM[1]			
Male	54.545	0.227	5,406
Female	48.712	0.160	4,552
Total	51.864	0.196	10,015
BCN[2]			
Male	59.121	0.586	6,499
Female	54.137	0.428	5,840
Total	56.762	0.511	12,411

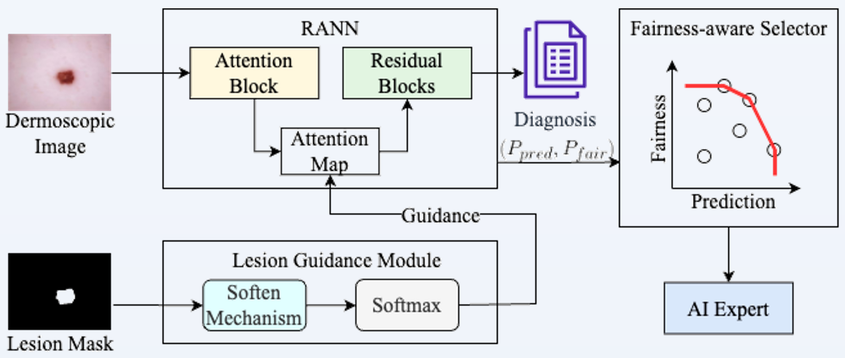

Figure 1. *The workflow of the proposed algorithm* ***Lesion Attn****. Residual Attention Neural Network (RANN) serves as the backbone. P_{pred} refers to the prediction performance and P_{fair} refers to the fairness performance. The* ***Soften Attention Guidance Module*** *guides the attention map of lesions during training. The* ***Fairness-aware Selector*** *is responsible for selecting the AI with balanced fairness and prediction performance.*

Results

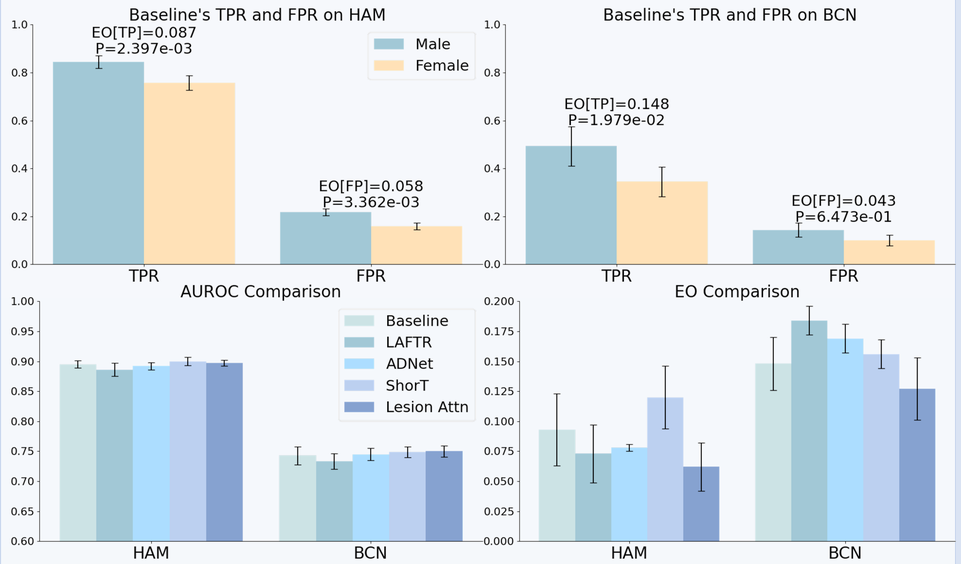

Figure 2. *gender-based biases identification and performance comparison. We took the most widely accepted algorithm as the baseline[3]. LAFTR, ADNet, and ShorT are fairness algorithms. Prediction performance was assessed using AUROC and bias was measured by Equalized Odds (EO).*

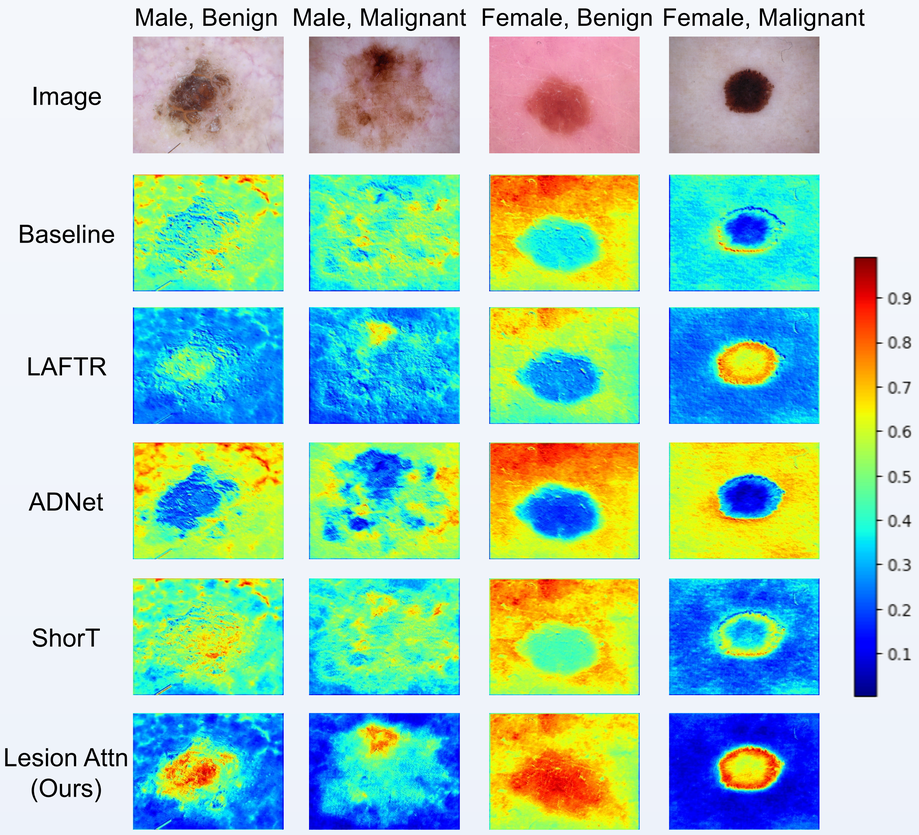

Figure 3. *Attention maps of RANNs trained with different algorithms. Lesion Attn is the only one that consistently focuses on lesions.*

Discussions

- **AI expert display higher performance for male patients**, with significant higher TPR ($p = 1.979\times10^{-2}$) but non-significantly higher FPR ($p = 6.473\times10^{-1}$) compared with female patients (Figure 2) .
- Lesion Attn outperforms other fairness algorithms in mitigating gender-based biases (Figure 2).
- **Lesion Attn is robust in mitigating disparity** (14.2% EO drop at external validation) **and simultaneously improving prediction performance** (0.9% AUROC increase at external validation) according to external validation (Figure 2).
- **Lesion Attn enables AI to diagnose skin cancer based on lesion region**--the diagnosis evidence of dermatologists. In contrast, the baseline and other fairness algorithms prioritize skin regions for prediction (Figure 3).

Conclusions

- This study identified biases of AI experts for skin cancer diagnosis towards female patients and mitigated such biases with Lesion Attn.
- AI experts' excessive attention on skin region may be associated with gender-based biases.

Reference

[1] Tschandl P, Rosendahl C, Kittler H. The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions[J]. Scientific data, 2018, 5(1): 1-9.

[2] Combalia M, Codella N C F, Rotemberg V, et al. Bcn20000: Dermoscopic lesions in the wild[J]. arXiv preprint arXiv:1908.02288, 2019.

[3] Esteva A, Kuprel B, Novoa R A, et al. Dermatologist-level classification of skin cancer with deep neural networks[J]. nature, 2017, 542(7639): 115-118.

Contacts

- Mr. Mingcheng Zhu: m.zhu23@nus.edu.sg
- Prof. Nan Liu: liu.nan@duke-nus.edu.sg