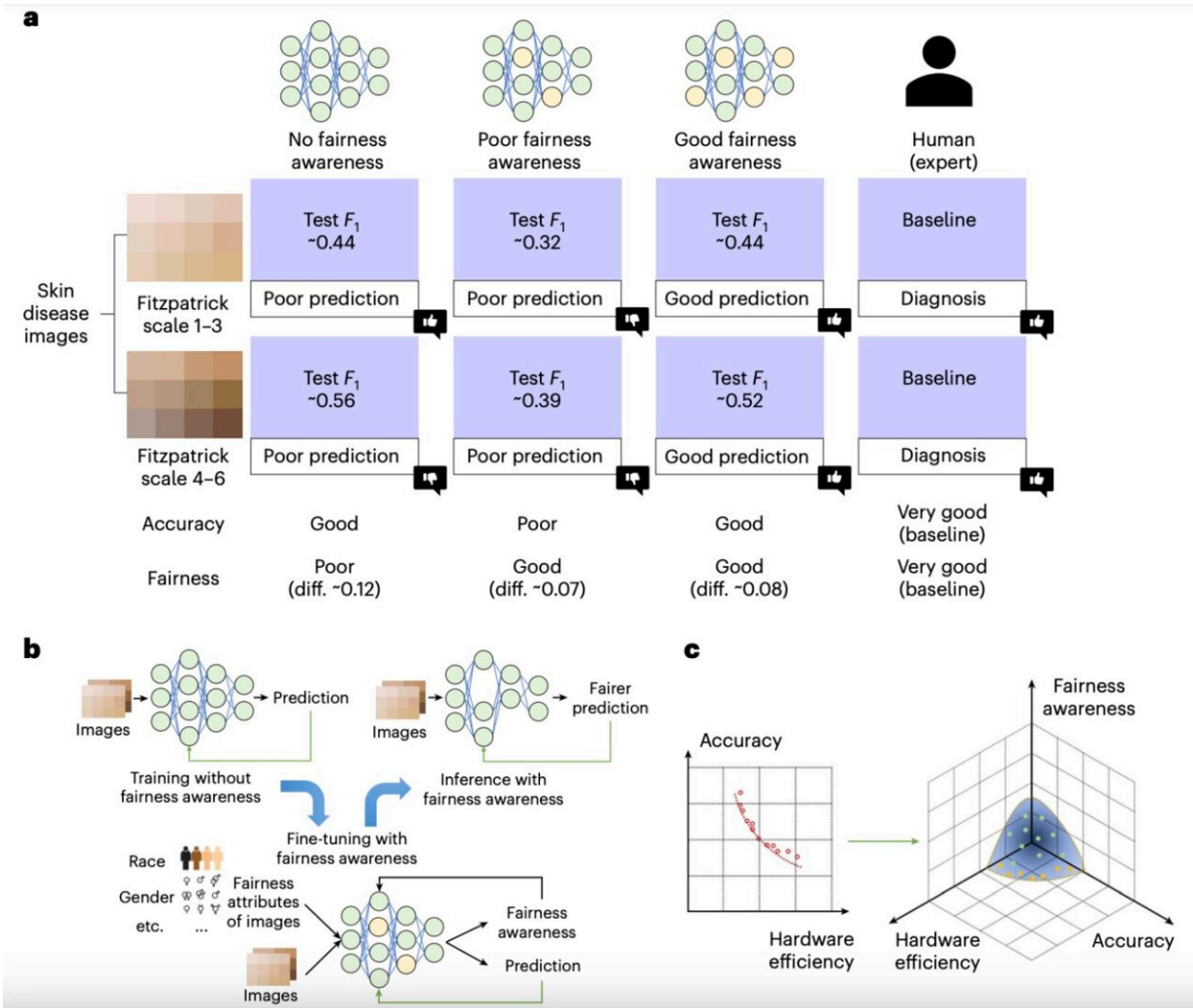


How hardware contributes to the fairness of artificial neural networks

August 24 2024, by Ingrid Fadelli



Fairness of neural networks. a. Illustrative example of neural network fairness awareness in dermatological disease detection. b. Process of model training with fairness awareness. c. New objective to be considered in system design. Fairness awareness adds a new objective goal we are looking for, extending a new

dimension of the problem. Credit: Guo et al

Over the past couple of decades, computer scientists have developed a wide range of deep neural networks (DNNs) designed to tackle various real-world tasks. While some of these models have proved to be highly effective, some studies found that they can be unfair, meaning that their performance may vary based on the data they were trained on and even the hardware platforms they were deployed on.

For instance, some studies showed that commercially available deep learning–based tools for facial recognition were significantly better at recognizing the features of fair-skinned individuals compared to dark-skinned individuals. These observed variations in the performance of AI, in great part due to disparities in the [training data](#) available, have inspired efforts aimed at improving the [fairness](#) of existing models.

Researchers at University of Notre Dame recently set out to investigate how hardware systems can contribute to the fairness of AI. Their paper, [published](#) in *Nature Electronics*, identifies ways in which emerging hardware designs, such as computing-in-memory (CiM) devices, can affect the fairness of DNNs.

"Our paper originated from an urgent need to address fairness in AI, especially in high-stakes areas like health care, where biases can lead to significant harm," Yiyu Shi, co-author of the paper, told Tech Xplore.

"While much research has focused on the fairness of algorithms, the role of hardware in influencing fairness has been largely ignored. As AI models increasingly deploy on resource-constrained devices, such as mobile and edge devices, we realized that the underlying hardware could potentially exacerbate or mitigate biases."

After reviewing past literature exploring discrepancies in AI performance, Shi and his colleagues realized that the contribution of hardware design to AI fairness had not been investigated yet. The key objective of their recent study was to fill this gap, specifically examining how new CiM hardware designs affected the fairness of DNNs.

"Our aim was to systematically explore these effects, particularly through the lens of emerging CiM architectures, and to propose solutions that could help ensure fair AI deployments across diverse hardware platforms," Shi explained. "We investigated the relationship between hardware and fairness by conducting a series of experiments using different hardware setups, particularly focusing on CiM architectures."

As part of this recent study, Shi and his colleagues carried out two main types of experiments. The first type was aimed at exploring the impact of hardware-aware neural architecture designs varying in size and structure, on the fairness of the results attained.

"Our experiments led us to several conclusions that were not limited to [device](#) selection," Shi said. "For example, we found that larger, more complex neural networks, which typically require more hardware resources, tend to exhibit greater fairness. However, these better models were also more difficult to deploy on devices with limited resources."

Building on what they observed in their experiments, the researchers proposed potential strategies that could help to increase the fairness of AI without posing significant computational challenges. One possible solution could be to compress larger models, thus retaining their performance while limiting their computational load.

Name	Number of levels	Device variability			
		L_0	L_1	L_2	L_3
RRAM ₁	1	0.0100	0.0100	0.0100	0.0100
FeFET ₂	4	0.0067	0.0135	0.0135	0.0067
RRAM ₄	4	0.0038	0.0151	0.0151	0.0038
FeFET ₆	4	0.0026	0.0155	0.0155	0.0026

For a device with several levels, the device variability for each level is L_x .

Device non-ideality modeling for different real and synthesized devices. Credit: Guo et al.

"The second type of experiments we carried out focused on certain non-idealities, such as device variability and stuck-at-fault issues coming together with CiM architectures," Shi said. "We used these hardware platforms to run various neural networks, examining how changes in hardware—such as differences in memory capacity or processing power—affected the [model](#)'s fairness.

"The results showed that various trade-offs were exhibited under different setups of device variations and that existing methods used to improve the robustness under device variations also contributed to these trade-offs."

To overcome the challenges unveiled in their second set of experiments, Shi and his colleagues suggest employing noise-aware training strategies. These strategies entail the introduction of controlled noise while training AI models, as a means of improving both their robustness and fairness without significantly increasing their computational demands.

"Our research highlights that the fairness of neural networks is not just a function of the data or algorithms but is also significantly influenced by the hardware on which they are deployed," Shi said. "One of the key findings is that larger, more resource-intensive models generally perform better in terms of fairness, but this comes at the cost of requiring more advanced hardware."

Through their experiments, the researchers also discovered that hardware-induced non-idealities, such as device variability, can lead to trade-offs between the accuracy and fairness of AI models. Their findings highlight the need to carefully consider both the design of AI model structures and the hardware platforms they will be deployed on, to reach a good balance between accuracy and fairness.

"Practically, our work suggests that when developing AI, particularly tools for sensitive applications (e.g., medical diagnostics), designers need to consider not only the software algorithms but also the hardware platforms," Shi said.

The recent work by this research team could contribute to future efforts aimed at increasing the fairness of AI, encouraging developers to focus on both hardware and software components. This could in turn facilitate the development of AI systems that are both accurate and equitable, yielding equally good results when analyzing the data of users with different physical and ethnic characteristics.

"Moving forward, our research will continue to delve into the intersection of hardware design and AI fairness," Shi said. "We plan to develop advanced cross-layer co-design frameworks that optimize neural network architectures for fairness while considering hardware constraints. This approach will involve exploring new types of [hardware platforms](#) that inherently support fairness alongside efficiency."

As part of their next studies, the researchers also plan to devise adaptive training techniques that could address the variability and limitations of different hardware systems. These techniques could ensure that AI models remain fair irrespective of the devices they are running on and the situations in which they are deployed.

"Another avenue of interest for us is to investigate how specific hardware configurations might be tuned to enhance fairness, potentially leading to new classes of devices designed with fairness as a primary objective," Shi added. "These efforts are crucial as AI systems become more ubiquitous, and the need for fair, unbiased decision-making becomes ever more critical."

More information: Yuanbo Guo et al, Hardware design and the fairness of a neural network, *Nature Electronics* (2024). [DOI: 10.1038/s41928-024-01213-0](https://doi.org/10.1038/s41928-024-01213-0)

© 2024 Science X Network

Citation: How hardware contributes to the fairness of artificial neural networks (2024, August 24) retrieved 24 August 2024 from <https://techxplore.com/news/2024-08-hardware-contributes-fairness-artificial-neural.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.