Bachelor Project Proposal

Gabriela Zhelyazkova (gazh@itu.dk)

Evaluating the impact of shortcuts in chest X-rays

Background

Chest X-rays are the most prevalent radiological examination worldwide, utilized for diagnosing a multitude of conditions. Deep learning models are using chest x-ray images for training and for further predicting of numerous diseases. However, these models often struggle to generalize across different patient demographics, clinical settings, and healthcare institutions. A significant issue is the model's tendency to utilize non-clinical metadata such as age, sex, and BMI as shortcuts in disease prediction, which could undermine the robustness and fairness of these diagnostic tools.

Importance of the Topic and Previous Work

As chest X-rays are integral in medical diagnostics globally, the automation of disease detection using these images holds substantial significance in both medical and technological fields. Notably, a study by Jabbour et al. (2020) on applying transfer learning to mitigate shortcut usage in CNNs underscores the critical need for developing models that perform reliably across diverse clinical environments.

Problem Statement

Current models exhibit a notable decline in accuracy when deprived of metadata shortcuts. This project aims to debias CNN models by reducing their reliance on these shortcuts and instead, enhancing their use of clinically relevant features, thereby improving their diagnostic accuracy.

Methodology

This research will employ CNNs to build and test models that prioritize clinical evidence over demographic metadata. The primary dataset for this investigation will be the CheXpert dataset, available from Stanford ML Group. This dataset will provide a robust foundation for training and validating the CNNs.

Literature (Key references for this project baseline will include)

1. Jabbour, S., Fouhey, D., Kazerooni, E., Sjoding, M. W., & Wiens, J. (2020, September 21). Deep Learning applied to chest X-Rays: Exploiting and preventing shortcuts. arXiv.org. https://arxiv.org/abs/2009.10132v1
2. Jiménez-Sánchez, A., Juodelyte, D., Chamberlain, B., & Cheplygina, V. (2022, November 8). Detecting shortcuts in medical images -- a case study in chest x-rays. arXiv.org. https://arxiv.org/abs/2211.04279
3. Banerjee, I., Bhattacharjee, K., Burns, J. L., Trivedi, H., Purkayastha, S., Seyyed-Kalantari, L., Patel, B. N., Shiradkar, R., & Gichoya, J. (2023). "Shortcuts" causing bias in radiology artificial intelligence: causes, evaluation and mitigation. ScienceDirect. https://www.sciencedirect.com/science/article/pii/S1546144023005264
4. Brown, A., Tomasev, N., Freyberg, J., Liu, Y., Karthikesalingam, A., & Schrouff, J. (2023). Detecting shortcut learning for fair medical AI using shortcut testing. Nature Communications, 14(1). https://doi.org/10.1038/s41467-023-39902-7
5. Gichoya, J. W., Banerjee, I., Bhimireddy, A. R., Burns, J. L., Celi, L. A., Chen, L. C., Correa, R., Dullerud, N., Ghassemi, M., Huang, S. C., Kuo, P. C., Lungren, M. P., Palmer, L. J., Price, B. J., Purkayastha, S., Pyrros, A. T., Oakden-Rayner, L., Okechukwu, C., Seyyed-Kalantari, L., Trivedi, H., … Zhang, H. (2022). AI recognition of patient race in medical imaging: a modelling study. The Lancet. Digital health, 4(6), e406–e414. https://doi.org/10.1016/S2589-7500(22)00063-2
6. Damgaard, C., Eriksen, T. N., Juodelyte, D., Cheplygina, V., & Jiménez-Sánchez, A. (2023, September 5). Augmenting Chest X-ray Datasets with Non-Expert Annotations. arXiv.org. https://arxiv.org/abs/2309.02244

Deliverables

1. A comprehensive project report detailing the findings, methodologies, and implications of the research.
2. A GitHub repository containing all code, datasets, and experiments conducted during the project, ensuring reproducibility and facilitating further research by the academic community.