CXR-LT: Long-tailed Multi-label Classification of Chest X-Rays

Nasir Abdulrasheed

Table of Contents

1 Description

Chest x-rays (CXRs) and medical imaging in general yield a long-tailed distribution, with a few prevalent labels and a significantly larger number of rare labels [2]. This makes traditionally trained neural networks biased towards the common 'head' classes and evaluate very poorly on 'tail' classes. This combination of being long-tailed and multi-label (label co-occurrence) makes the accurate classification of CXRs challenging. Additionally, many diagnostic features on CXRs are so rare yet important that creating models with few- and zero-shot classification capabilities is difficult, owing to the rarity or lack of samples to train on. Also, there are relatively few proven training methodologies that reliably improve the performance of models in long-tailed and multi-label settings. Overall, it can be surmised that zero-shot long-tailed multi-label recognition should be the gold standard in medical image analysis, as that would match the real-world scenario of having to classify images with very few or no training samples available for some classes. It would also more closely match what radiologists actually do in practice.

To foster research in this direction, a curated long-tailed and multi-label dataset of CXRs was created and released in 2023 by Holste et al. [2], along with a challenge with specific objectives. This project aims to contribute to the body of research in this space by applying recommendations published by the challenge authors based on the results of the first iteration of the challenge to new experiments to improve the performance of existing solutions, while also interrogating them to uncover new insights into long-tailed multi-label classification of medical images.

2 Objectives

2.1 Primary

1. Investigate the impact of using robust asymmetric loss (RAL) vs asymmetric loss (ASL) + weighted BCE in training CheXFusion (Kim [3]) in place of the asymmetric focal loss initially used [5].
2. Investigate a modified architecture by substituting simple class queries with label embeddings from PubMedBERT or UmlsBERT [1, 4].
3. Robustly investigate the zero-shot capabilities of the model from (2)

2.2 Secondary

1. Perform a comprehensive comparison of experiment results with the current state-of-the-art models on the same dataset.
2. Experiment with creating view-specific CNN backbones for the fusion model, as opposed to the single backbone used in the original CheXFusion paper [3].

2.3 Tertiary

1. Develop a web wrapper around the deployed model to test the final model in the real world, allowing users to upload CXRs and receive predictions.

3 Ethics

A full ethics application is required for the project, as pseudo-anonymised human image data is being used in the analysis. The full ethics form has been submitted, and approval has been secured.

4 Resources

Resources needed include access:

- dataset hosted on PhysionNet (see Section 6)
- GPU compute resources for training and inference
- model and web hosting platforms for deployment

5 References

[1] Yu Gu et al. "Domain-Specific Language Model Pretraining for Biomedical Natural Language Processing". In: *ACM Trans. Comput. Healthcare* 3.1 (Jan. 2022). arXiv:2007.15779 [cs], pp. 1–23. ISSN: 2691-1957, 2637-8051. DOI: 10.1145/3458754. URL: http://arxiv.org/abs/2007.15779 (visited on 06/02/2025).

[2] Gregory Holste et al. "Towards long-tailed, multi-label disease classification from chest X-ray: Overview of the CXR-LT challenge". In: *Medical Image Analysis* 97 (Oct. 2024), p. 103224. ISSN: 1361-8415. DOI: 10.1016/j.media.2024.103224. URL: https://www.sciencedirect.com/science/article/pii/S136184152400149X (visited on 01/30/2025).

[3] Dongkyun Kim. *CheXFusion: Effective Fusion of Multi-View Features using Transformers for Long-Tailed Chest X-Ray Classification.* arXiv:2308.03968 [cs]. Aug. 2023. DOI: 10.48550/arXiv.2308.03968. URL: http://arxiv.org/abs/2308.03968 (visited on 06/02/2025).

[4] George Michalopoulos et al. *UmlsBERT: Clinical Domain Knowledge Augmentation of Contextual Embeddings Using the Unified Medical Language System Metathesaurus.* arXiv:2010.10391 [cs] version: 5. June 2021. DOI: 10.48550/arXiv.2010.10391. URL: http://arxiv.org/abs/2010.10391 (visited on 06/01/2025).

[5] Wongi Park et al. *Robust Asymmetric Loss for Multi-Label Long-Tailed Learning*. arXiv:2308.05542 [cs]. Aug. 2023. DOI: 10 . 48550 / arXiv . 2308 . 05542. URL: http : //arxiv.org/abs/2308.05542 (visited on 06/02/2025).

6 Appendix A

- CXR-LT 2024: https://physionet.org/content/cxr-lt-iccv-workshop-cvamd/2.0.0/
- CXR-LT 2023: https://physionet.org/content/cxr-lt-iccv-workshop-cvamd/1.1.0/
- MIMIC-CXR-JPG V2.0.0: https://physionet.org/content/mimic-cxr-jpg/2.0.0/