MICHIGAN STATE UNIVERSITY

Deep Learning Algorithm for Pneumoconiosis Staging on Chest Radiographs

Michal Borek*, Meiqi Liu, Zenas Huang, Ian Loveless, Kenneth Rosenman, Adam Alessio, Ling Wang
Computational Mathematics, Science, and Engineering | Biomedical Engineering
Michigan State University, East Lansing, MI

INTRODUCTION

Pneumoconiosis is an occupational lung disease caused by inhaling mineral dust, and chest radiography remains the key screening tool. Although standardization efforts by the ILO and NIOSH—such as the B Reader Certification Program—have improved consistency, challenges like reader variability, limited certified readers, and potential conflicts of interest persist. This study leverages artificial intelligence to objectively classify pneumoconiosis severity on a 4-point scale (0–3) using posterior-anterior chest radiographs from the NIOSH repository. A ResNet framework employing various loss functions (cross-entropy, corn, coral, focal staging, hierarchical, and hierarchical cross-entropy) is explored to enhance diagnostic reliability.

METHODS

- A ResNet-based framework was implemented for classifying the severity of pneumoconiosis. Various loss functions were evaluated, including cross-entropy, Corn Loss, Coral Loss, Focal Staging Loss, Hierarchical loss and Hierarchical Cross-Entropy (HCE) loss.
- PA chest radiographs and labels were obtained from the NIOSH lung image repository. Consensus diagnostic labels were determined by the median readings of at least two certified B-readers. The dataset includes 1306 abnormal radiographs (260 with profusion grade 1, 191 with grade 2, and 84 with grade 3) and 771 normal images.
  - The images were split into 80% training, 10% validation, and 10% testing sets.
- To ensure robust model selection, we employed an extensive set of performance metrics—Accuracy (with SD), Sensitivity (SD), Specificity (SD), AUC (SD), IMCP (SD), MAMAE (SD), Cost-Sensitive Balanced Accuracy (SD), Prevalence-Weighted Cost-Sensitive Accuracy (SD), and Cost-Sensitive Multi-Accuracy. This comprehensive evaluation framework allowed us to capture both overall predictive performance and cost-sensitive considerations.

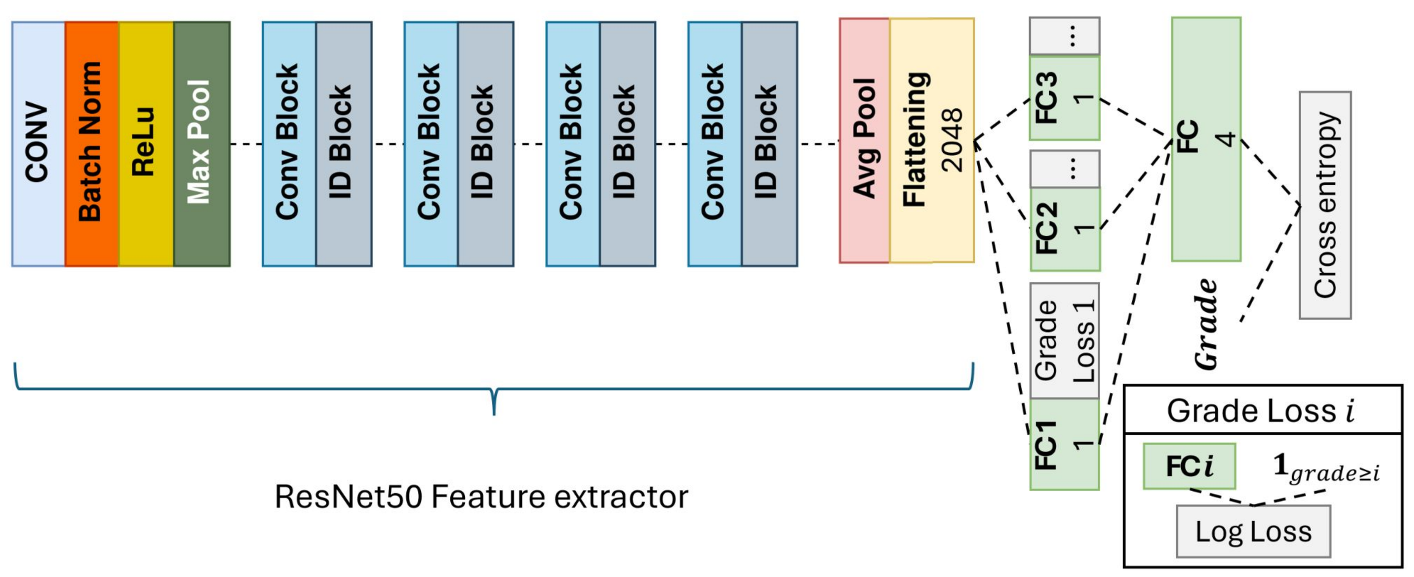

Fig 1. Adjusted ResNet architecture (image adapted from Gorlapraveen123).

This preliminary study is among the first to apply deep learning for multi-classification of occupational lung disease, demonstrating that different loss functions can effectively grade pneumoconiosis severity.

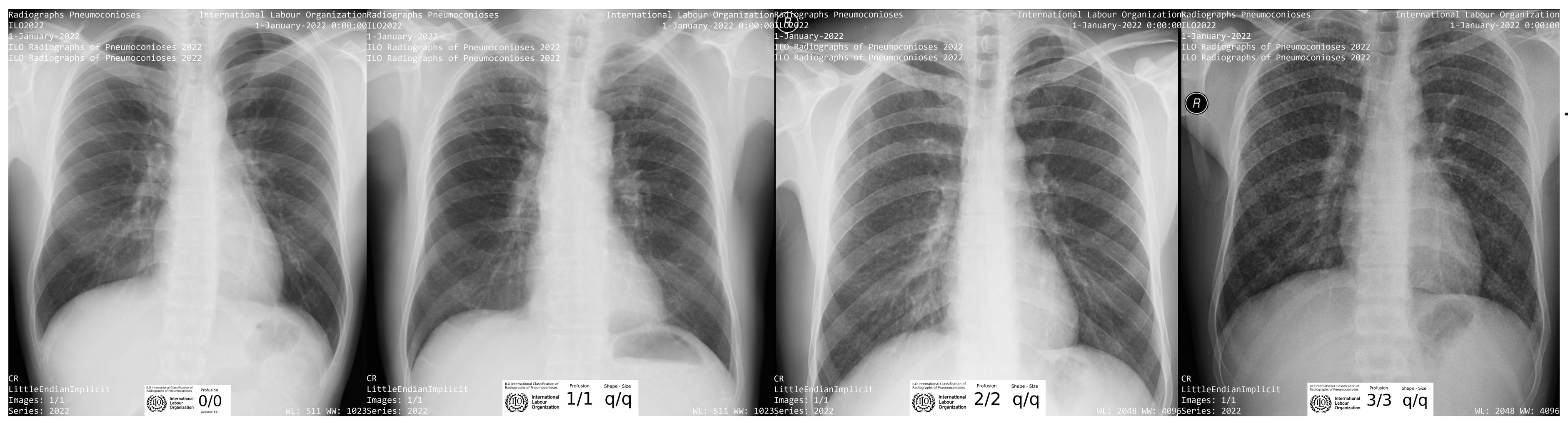

Fig 2. Example radiographs with ILO Profusion Grade 0, 1, 2 and 3.

All loss functions perform very well, around 70%

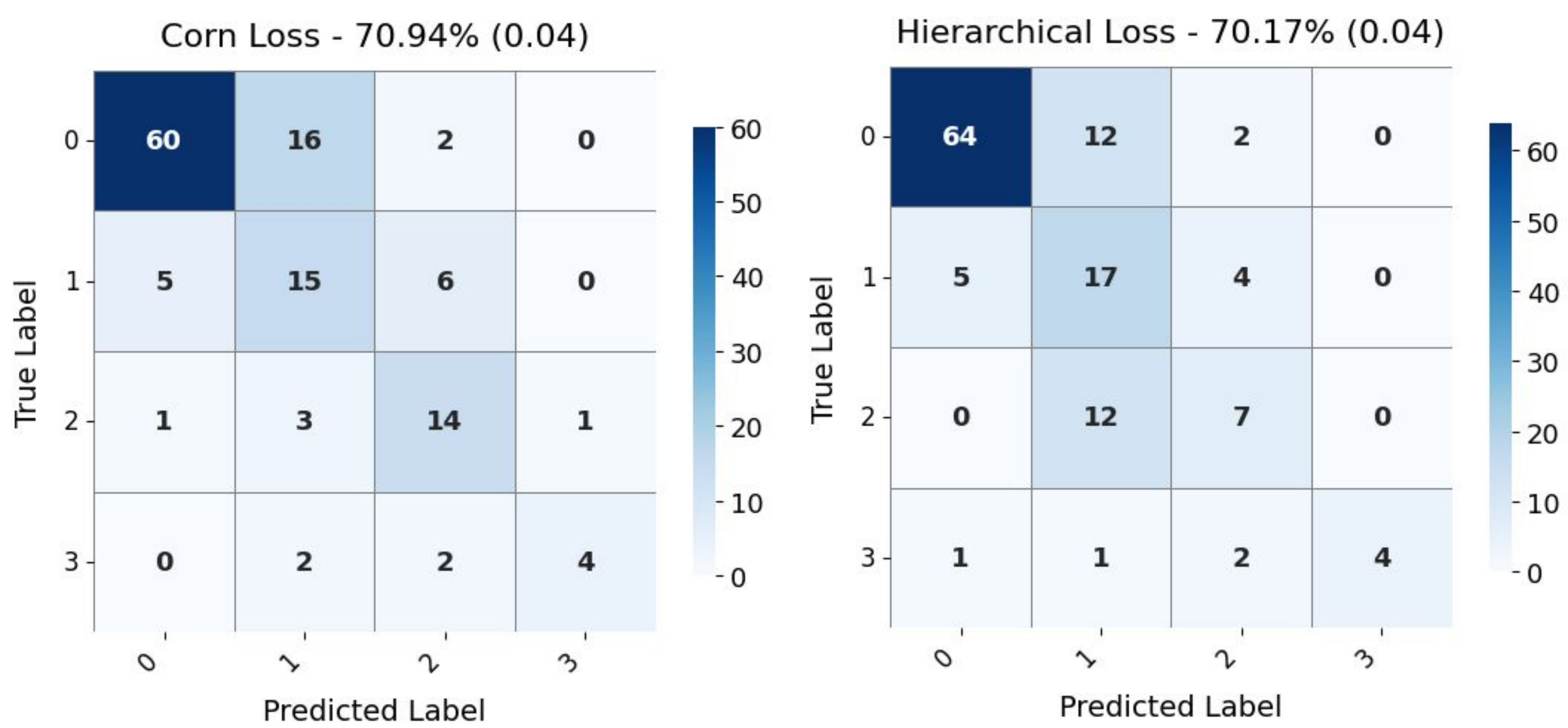

Fig 3. Accuracy comparisons and their confusion matrices for Corn Loss and Hierarchical Loss

Corn Loss and Hierarchical Loss seem to be the best performing loss functions

actual/ predicted class	Class 0	Class 1	Class 2	Class 3
Class 0	0	1	2	2
Class 1	4	0	1	2
Class 2	4	1	0	1
Class 3	4	2	1	0

Table 1. Cost-sensitive table for the imbalance dataset

* Contact: Michal Borek, Undergraduate Student, borekmi1@msu.edu

RESULTS

Table 2. Comparison of accuracy, Sensitivity, Specificity, and Cost Sensitive Multi-class Accuracy (CSMA) between cross-entropy loss, focal staging loss, coral loss, corn loss, hierarchical loss, hierarchical cross-entropy (hierarchical combined).

Models	Accuracy (SD)	Sensitivity (SD)	Specificity (SD)	CSMA (SD)
In-Distribution Test Set:				
ResNet-18 (Cross-Entropy)	0.694 (0.04)	0.601 (0.04)	0.890 (0.01)	0.820 (0.03)
ResNet-18 (Focal Staging with λ)	0.695 (0.04)	0.572 (0.04)	0.883 (0.02)	0.800 (0.03)
ResNet-18 (CORAL)	0.703 (0.04)	0.577 (0.04)	0.853 (0.02)	0.685 (0.04)
ResNet-18 (CORN)	**0.709**(0.04)	**0.646**(0.04)	**0.898**(0.01)	**0.838**(0.03)
ResNet-18 (Hierarchical)	0.703 (0.04)	0.586 (0.04)	0.895 (0.01)	0.837 (0.03)
ResNet-18 (Combined Hierarchical)	0.700 (0.04)	0.541 (0.04)	0.877 (0.02)	0.786 (0.03)

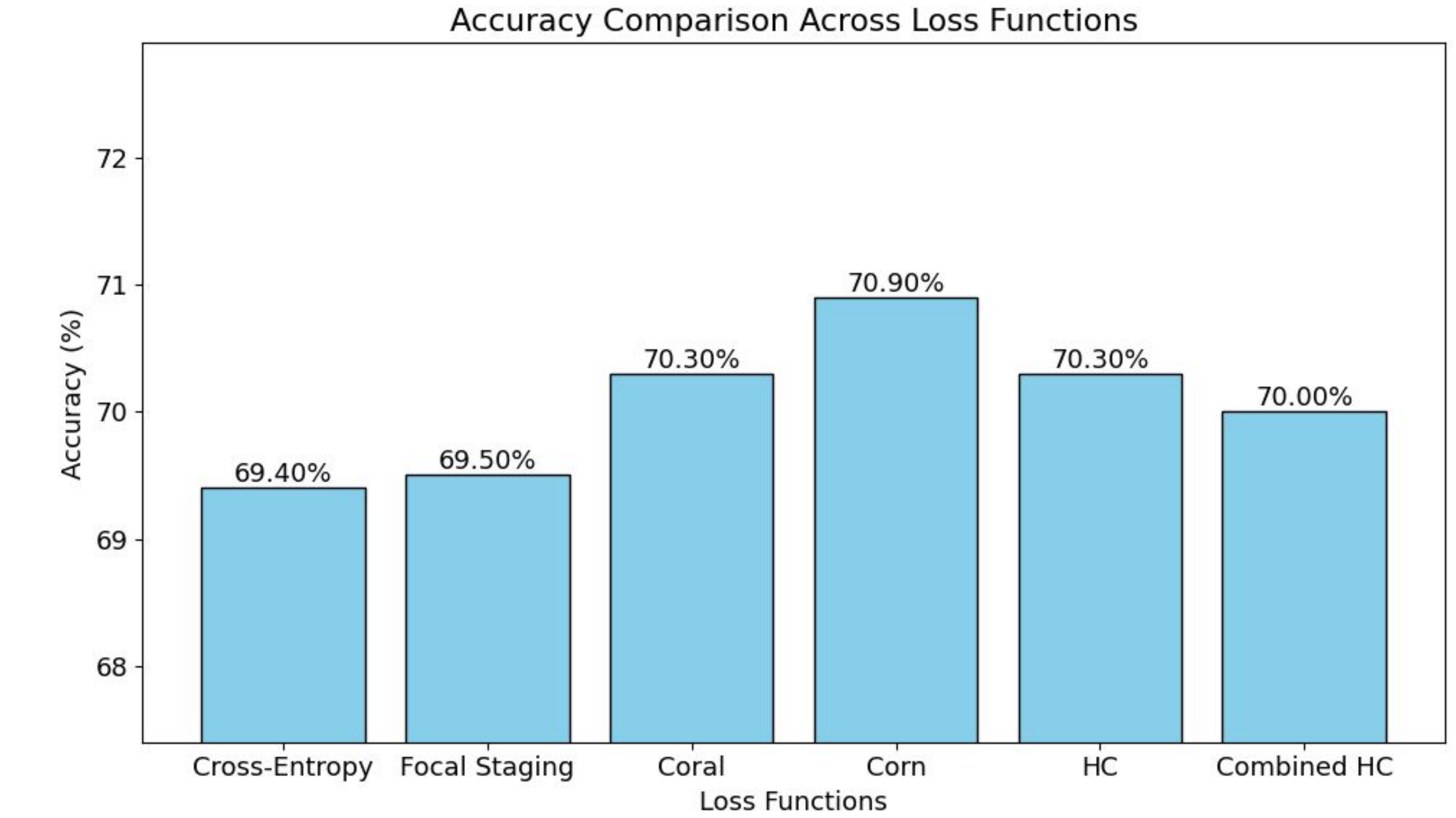

Fig 4. Hist Plots comparing accuracy for a different loss functions.

DISCUSSION

- All models perform similarly, but the CORN loss model slightly outperforms others with an accuracy of 0.709, sensitivity of 0.646, specificity of 0.898, and CSMA of 0.838. Further hyperparameter tuning for focal staging and hierarchical losses may improve their performance.
- At this time it seems that Corn Loss performs the best out of all tested loss functions. Further fine-tuning hyperparameters for Focal Staging Loss or Hierarchical Loss might increase performance.

REFERENCES

[1] Castranova, V. and Vallyathan, V., “Silicosis and coal workers’ pneumoconiosis.,” Environmental health perspectives 108(suppl 4), 675–684 (2000).

[2] Wagner, G. R., Attfield, M. D., Kennedy, R. D., and Parker, J. E., “The NIOSH B reader certification program: an update report,” Journal of Occupational Medicine , 879–884 (1992).

[3] Cohen, J. P., Viviano, J. D., Bertin, P., Morrison, P., Torabian, P., Guarrera, M., Lungren, M. P., Chaudhari,
A., Brooks, R., Hashir, M., et al., “TorchXRayVision: A library of chest x-ray datasets and models,” in [International Conference on Medical Imaging with Deep Learning], 231–249, PMLR (2022).

[4] Sun W, Wu D, Luo Y, Liu L, Zhang H, Wu S, Zhang Y, Wang C, Zheng H, Shen J, Luo C. A Fully Deep Learning Paradigm for Pneumoconiosis Staging on Chest Radiographs. IEEE J Biomed Health Inform. 2022 Oct;26(10):5154-5164. doi: 10.1109/JBHI.2022.3190923. Epub 2022 Oct 4. PMID: 35834466.

[5] Wenzhi Cao, Vahid Mirjalili, and Sebastian Raschka (2020). Rank Consistent Ordinal Regression for Neural Networks with Application to Age Estimation. Pattern Recognition Letters 140, pp. 325-331; https://doi.org/10.1016/j.patrec.2020.11.008.

[6] Xintong Shi, Wenzhi Cao, and Sebastian Raschka (2021). Deep Neural Networks for Rank-Consistent Ordinal Regression Based On Conditional Probabilities. Arxiv preprint; https://arxiv.org/abs/2111.08851

[7] Villar, V. A., de Soto, K., & Gagliano, A. (2023). Hierarchical Cross Entropy for Classification of Astrophysical Transients. Arxiv preprints; https://arxiv.org/abs/2312.02266

MIDI Lab www.midilab.org