Convolutional Neural Networks and Explainable AI for Medical Imaging

Student

Honey Lane Ante

Introduction: Medical imaging is vital for diagnosing health conditions. Machine learning, more specifically deep learning, is now a key part of this field, enhancing image analysis and streamlining radiology departments. Convolutional Neural Networks (CNNs) have improved the classification of medical images, a task demonstrated by their use with the CheXpert dataset. However, the complexity of CNNs creates a challenge: their decisions are often not transparent, leading to a gap in trust and understanding.

This project employs CNNs to identify chest pathologies from the CheXpert dataset, aspiring to match the results of prior research while making the decision-making process clearer through explainable AI (XAI) methods.

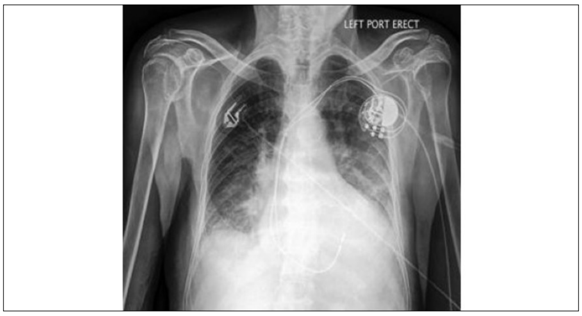
Approach: For this project, I developed different models each addressing the 5 uncertainty approaches: U-Ignore, U-Ones, U Zeroes, U-SelfTrained, and U-MultiClass. In constructing the models, a combination of strategies was employed including augmentation, effective pipeline preparation, hyperparameter tuning, transfer learning and distributed training. I tested various model variations, focusing on different combinations of augmentation and hyperparameter tuning.

Model evaluation involved running predictions on the validation set for each built model, followed by selecting the top 20 checkpoints that yielded the highest AUROC scores across five competition pathologies namely Atelectasis, Cardiomegaly, Consolidation, Edema and Pleural Effusion. The model associated with the highest-performing checkpoint, was subsequently utilized for performance assessment on the test set. Additionally, to bolster the interpretability of these models, explainability tools LIME and Grad CAM were employed, providing insights into the decision-making processes and enhancing the alignment of model predictions with medical expertise.

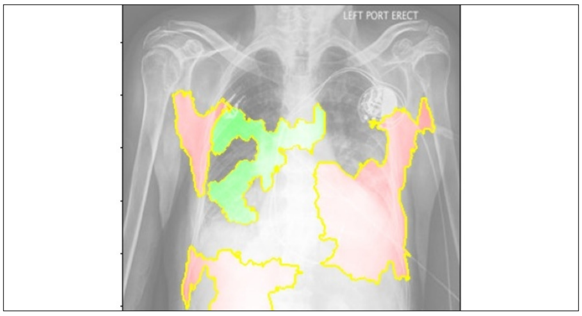
Result: In this study, a CNN was successfully implemented to classify 14 pathologies in the CheXpert dataset. The use of transfer learning was vital for efficient feature extraction and enhanced performance. However, the effectiveness of augmentation proved inconsistent. Addressing the dataset's size, efficient pipeline management and distributed training were key, while hyperparameter tuning played a crucial role in optimizing results. The best outcomes for U-Ignore, U-Ones, and U-Self Trained approaches were achieved through hyperparameter tuning without augmentation, a contrast to the variable results observed in U-Zeroes and U-MultiClass approaches. Though my results didn't entirely match CheXpert study's performance, they remained competitive, with the most significant AUROC deviation being a mere 0.06.

Explainable AI, particularly Grad CAM, was instrumental in offering transparent model explanations, effectively pinpointing relevant areas within the chest. The quality of these explanations was directly related to the prediction accuracy. This study highlights the significance of XAI in medical imaging, enhancing the trust and understanding of AI-driven diagnostics.

Test image.
I. Jeremy et al., CheXpert, 2019.



LIME result on the test image.
Own presentment



Grad CAM result on the test image.
Own presentment

Advisor
Hannes Badertscher

Subject Area
Data Science