Multi-stage Multimodal Deep Learning for Harmonization of Radiology Study Descriptions

Zihan Li, PhD Student, University of Washington
Paul Kinahan

Introduction

Radiology study descriptions within DICOM image headers exhibit high variability, complicating efficient data harmonization and patient cohort selection for pooled collections. This challenge is accentuated by the long-tail distribution of descriptions, where a few categories dominate, leaving others underrepresented. Existing methods, including basic NLP tools, struggle with scalability and accuracy in such contexts.

Hypothesis

A multi-stage multimodal deep learning model can effectively harmonize radiology study descriptions by addressing the challenges of long-tail data distribution and enhancing the reliability of automated categorization.

Methods

We propose MFFNet (Multi-stage Feature Fusion Network) utilizing BERT (Bidirectional Encoder Representations from Transformers) for meta data processing and TotalSegmentator for image analysis as shown in Figure 1. In Stage 1, patient exam-level features are used for coarse-grained classification. Stage 2 builds on these predictions, employing specialized models to refine the classification of LOINC Code. Stage 3 introduces image scan-level features including the image information to re-evaluate low-confidence cases flagged in earlier stages. We also adopt Focal Loss and Weighted Data Sampling to mitigate long-tail distribution issues. A confidence prediction mechanism further improves the reliability of classifications by escalating low-confidence cases for reassessment or expert review.

Results

MFFNet significantly outperformed BERT, achieving near-perfect accuracy across validation categories. The model demonstrated a high ability to handle diverse data, reducing misclassifications from 379 with BERT to one. In Stage 1, MFFNet achieved 100% accuracy in coarse-grained classification with no low-confidence cases. Stage 2 refined these predictions with 99.9% accuracy, identifying 28 low-confidence cases. In the final stage, MFFNet resolved all low-confidence cases, eliminating the need for expert intervention in most scenarios as shown in Figure 2.

Conclusion

MFFNet is a 'Helper AI' tool for radiology study description harmonization that can address the long-tail challenge, thus boosting workflow efficiency for curation and selection of image cohorts for research using large collections of radiological images.

Figure(s)

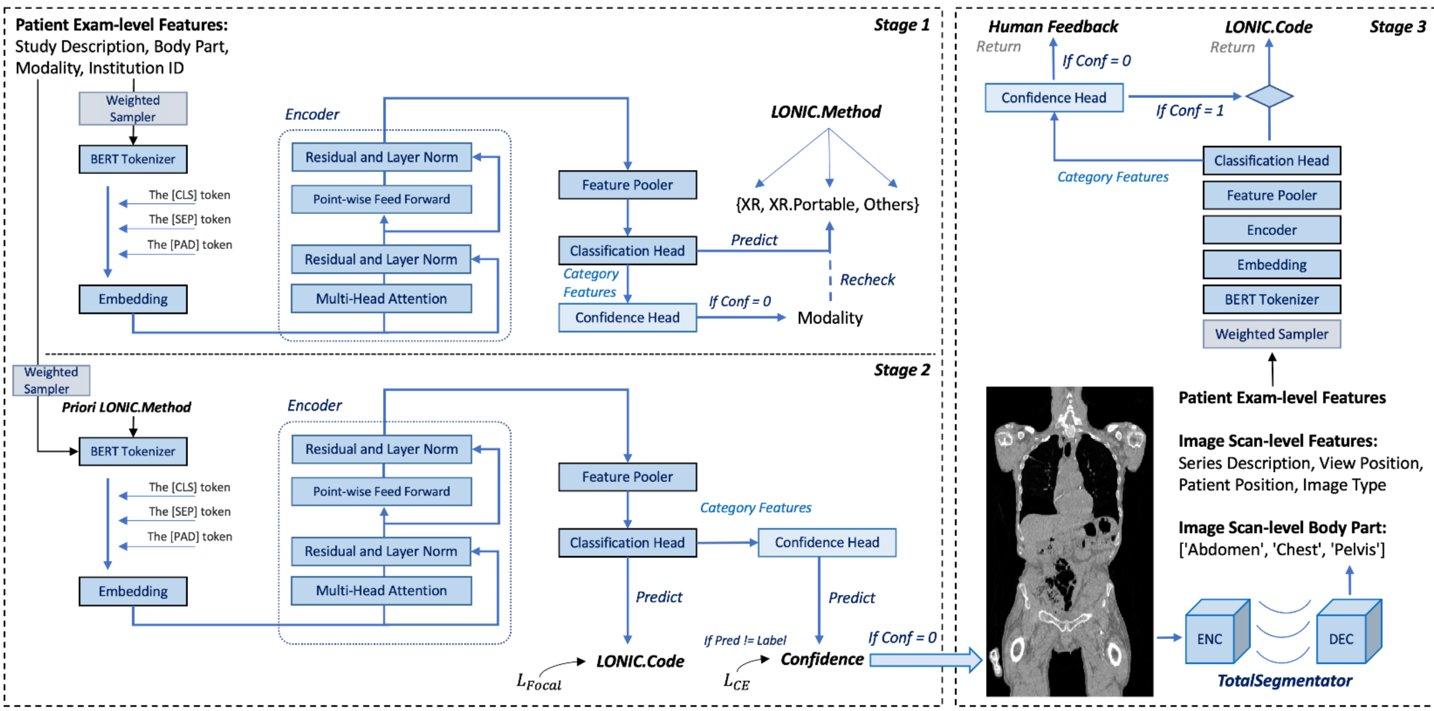
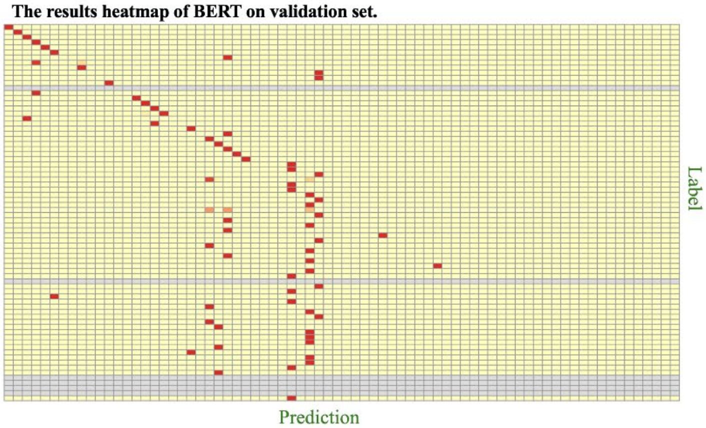
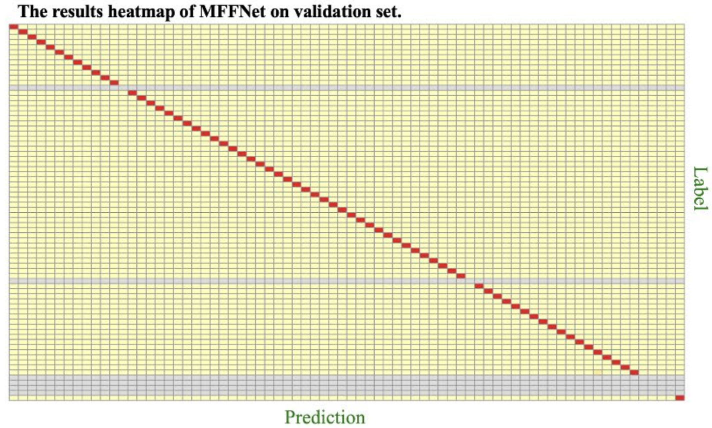


Figure 1. Overview of proposed Multi-stage Feature Fusion Network (MFFNet).

The results heatmap of BERT on validation set.



The results heatmap of MFFNet on validation set.



The results of BERT in different stages on the validation set.

BERT	Error Cases
Overall	379/30664

The results of MFFNet in different stages on the validation set.

MFFNet	Error Cases	Low Confidence
Overall	1/30664	0/30664

Stage1	Error Cases	Low Confidence
Overall	0/30664	0/30664
Stage2	**Error Cases**	**Low Confidence**
XR	11/16699	11/16699
XR portable	0/10203	0/10203
Other	18/3762	17/3762
Stage3	**Error Cases**	**Low Confidence**
XR	0/11	0/11
Other	0/17	0/17

Figure 2. Results comparison between BERT and MFFNet on the validation set, which include the results heatmap and quantitative analysis at each stage of MFFNet.

Keywords

Applications; Artificial Intelligence/Machine Learning; Emerging Technologies