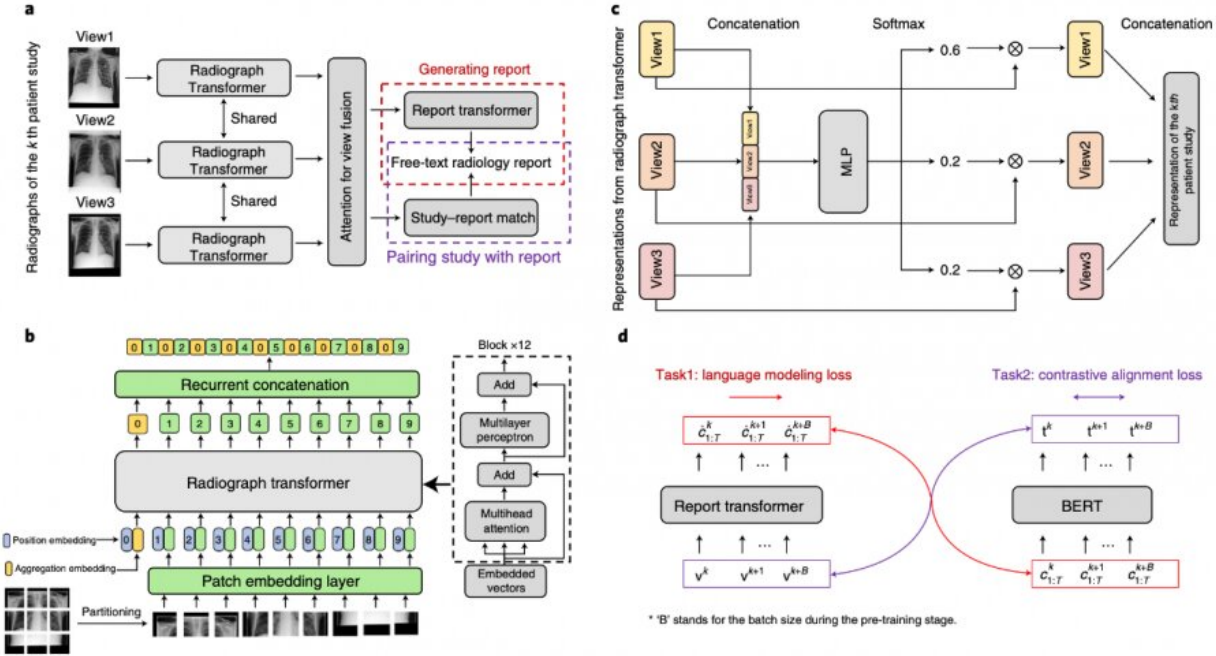# New AI algorithms for cost-effective medical image diagnostics

April 13 2022



REFERS workflow. Researchers forward radiographs of the k-th patient study through the radiograph transformer, fuse representations of different views using an attention mechanism, and use report generation and study–report representation consistency reinforcement to exploit the information in radiology reports. Graph a, an overview of the whole pipeline. Graph b, the architecture of the radiograph transformer. Graph c, attention for view fusion is elaborated. MLP stands for a multi-layer perceptron. Graph d, two supervision tasks are shown, report generation and study–report representation consistency reinforcement. Credit: The University of Hong Kong

Medical imaging is an important part of modern healthcare, enhancing both the precision, reliability and development of treatment for various diseases. Artificial intelligence has also been widely used to further enhance the process.

However, conventional medical image diagnosis employing AI algorithms require large amounts of annotations as supervision signals for model training. To acquire accurate labels for the AI algorithms—radiologists, as part of the clinical routine, prepare radiology reports for each of their patients, followed by annotation staff extracting and confirming structured labels from those reports using human-defined rules and existing natural language processing (NLP) tools. The ultimate accuracy of extracted labels hinges on the quality of human work and various NLP tools. The method comes at a heavy price, being both labor intensive and time consuming.

An engineering team at the University of Hong Kong (HKU) has developed a new approach "REFERS" (Reviewing Free-text Reports for Supervision), which can cut human cost down by 90%, by enabling the automatic acquisition of supervision signals from hundreds of thousands of radiology reports at the same time. It attains a high accuracy in predictions, surpassing its counterpart of conventional medical image diagnosis employing AI algorithms.

The innovative approach marks a solid step towards realizing generalized medical artificial intelligence. The breakthrough was published in *Nature Machine Intelligence* in the paper titled "Generalized radiograph representation learning via cross-supervision between images and free-text radiology reports."

"AI-enabled medical image diagnosis has the potential to support medical specialists in reducing their workload and improving the diagnostic efficiency and accuracy, including but not limited to reducing

the diagnosis time and detecting subtle disease patterns," said Professor Yu Yizhou, leader of the team from HKU's Department of Computer Science under the Faculty of Engineering.
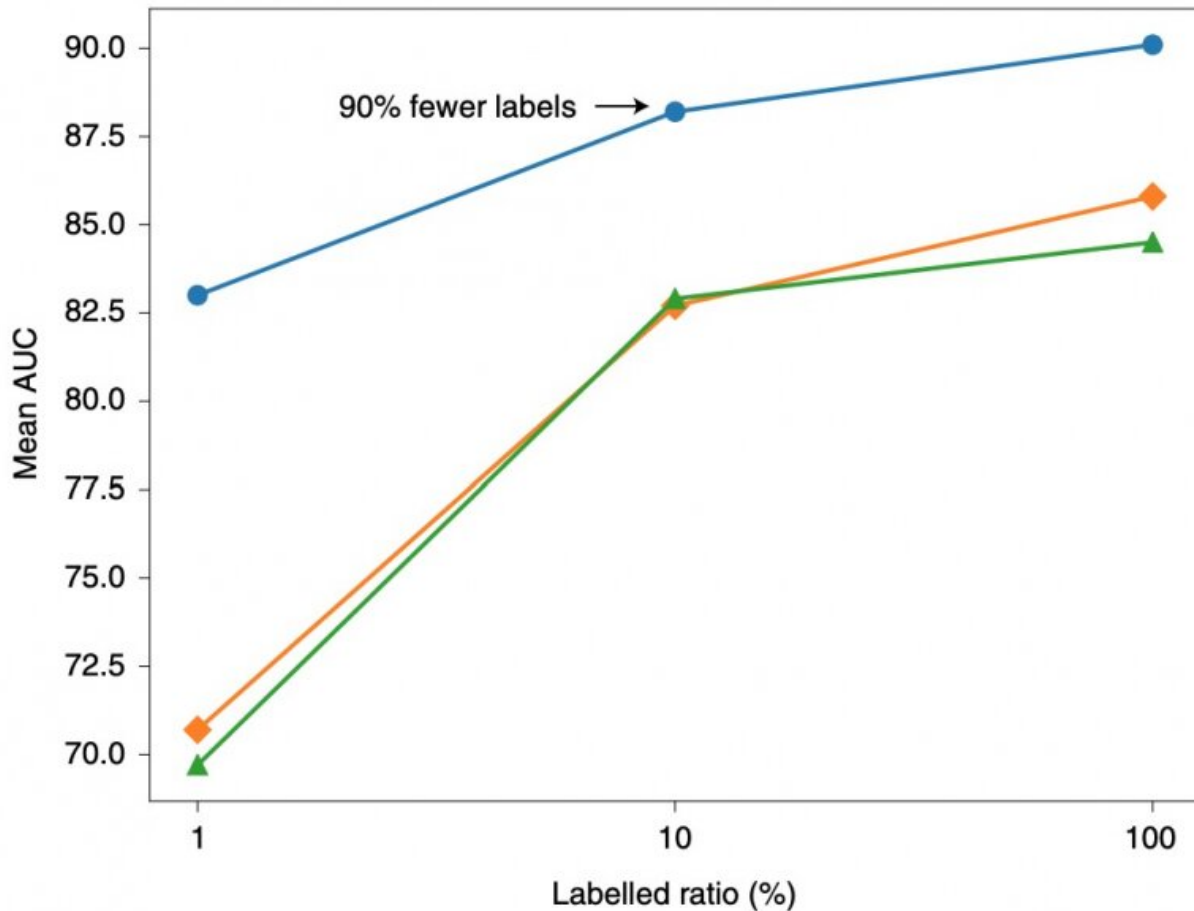


Performance of the new approach under various label ratios on NIH ChestX-ray. Researchers denote the percentage of annotated training data in the target domain that REFERS needs to achieve comparable results with those of Model Genesis and ImageNet pre-training. Note that all three methods share the same transformer-based backbone. Credit: The University of Hong Kong

"We believe abstract and complex logical reasoning sentences in radiology reports provide sufficient information for learning easily transferable visual features. With appropriate training, REFERS directly learns radiograph representations from free-text reports without the need to involve manpower in labeling." Professor Yu remarked.

For training REFERS, the research team uses a public database with 370,000 X-Ray images, and associated radiology reports, on 14 common chest diseases including atelectasis, cardiomegaly, pleural effusion, pneumonia and pneumothorax. The researchers managed to build a radiograph recognition model using 100 radiographs only, and attains 83% accuracy in predictions. When the number was increased to 1,000, their model exhibits amazing performance with an accuracy of 88.2%, which surpasses its counterpart trained with 10,000 radiologist annotations (accuracy at 87.6%). When 10,000 radiographs were used, the accuracy is at 90.1%. In general, an accuracy above 85% in predictions is useful in real-world clinical applications.

REFERS achieves the goal by accomplishing two report-related tasks, i.e., report generation and radiograph–report matching. In the first task, REFERS translates radiographs into text reports by first encoding radiographs into an intermediate representation, which is then used to predict text reports via a decoder network. A cost function is defined to measure the similarity between predicted and real report texts, based on which gradient-based optimization is employed to train the neural network and update its weights.

As for the second task, REFERS first encodes both radiographs and free-text reports into the same semantic space, where representations of each report and its associated radiographs are aligned via contrastive learning.

"Compared to conventional methods that heavily rely on human annotations, REFERS has the ability to acquire supervision from each

Provided by The University of Hong Kong