nature communications

Article

https://doi.org/10.1038/s41467-025-58344-x

A clinically accessible small multimodal radiology model and evaluation metric for chest X-ray findings

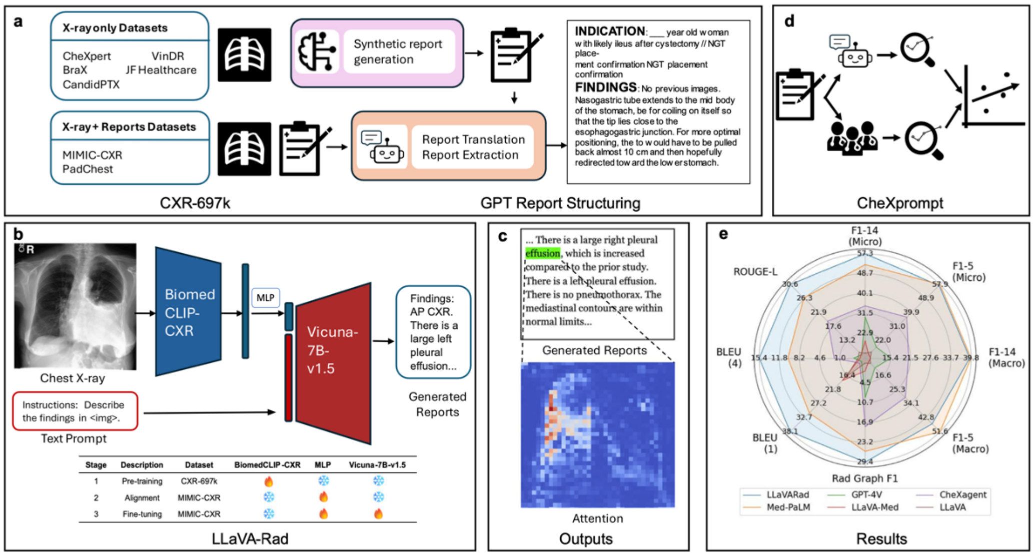
Received: 26 June 2024

Accepted: 19 March 2025

Published online: 01 April 2025

Check for updates

Juan Manuel Zambrano Chaves[1,2,7], Shih-Cheng Huang[2,7], Yanbo Xu[1,7], Hanwen Xu[3,7], Naoto Usuyama[1,7], Sheng Zhang[1,7], Fei Wang[4], Yujia Xie[1], Mahmoud Khademi[1], Ziyi Yang[1], Hany Awadalla[1], Julia Gong[1], Houdong Hu[1], Jianwei Yang[1], Chunyuan Li[1], Jianfeng Gao[1], Yu Gu[1], Cliff Wong[1], Mu Wei[1], Tristan Naumann[1], Muhao Chen[5], Matthew P. Lungren[1,2,6], Akshay Chaudhari[2], Serena Yeung-Levy[2], Curtis P. Langlotz[2], Sheng Wang[3] ✉ & Hoifung Poon[1] ✉



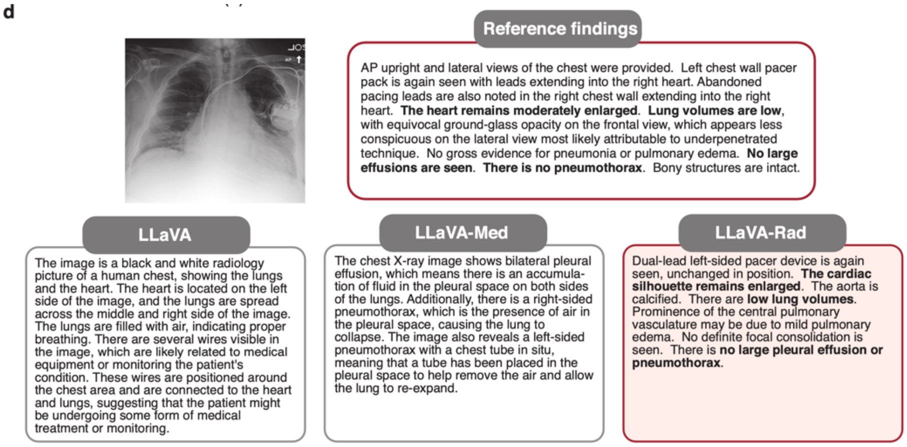
INDICATION: ___ year old w oman w ith likely ileus after cystectomy // NGT place-
ment confirmation NGT placement confirmation

FINDINGS: No previous images. Nasogastric tube extends to the mid body of the stomach, be for coiling on itself so that the tip lies close to the esophagogastric junction. For more optimal positioning, the to w ould have to be pulled back almost 10 cm and then hopefully redirected tow ard the low er stomach.

Stage	Description	Dataset	BiomedCLIP-CXR	MLP	Vicuna-7B-v1.5
1	Pre-training	CXR-697k	🔥	❄	❄
2	Alignment	MIMIC-CXR	❄	🔥	❄
3	Fine-tuning	MIMIC-CXR	❄	🔥	🔥

... There is a large right pleural effusion, which is increased compared to the prior study. There is a left pleural effusion. There is no pneumothorax. The mediastinal contours are within normal limits...

Fig. 1 | LLaVA-Rad overview. a To train LLaVA-Rad, we assemble a large dataset with over 697 thousand chest X-ray image-text pairs; GPT-4 is used to synthesize reports from labels, translate reports from Spanish, and process and structure the corresponding radiology reports. **b** We adopt a modular three-stage approach to train LLaVA-Rad, comprised of pre-training, alignment and fine-tuning. **c** A qualitative visualization of the model's attention during its generative process. **d** For evaluation, we also propose a novel factual error scoring approach using GPT-4 and demonstrate its parity with expert evaluation. **e** LLaVA-Rad outperforms much larger generalist and specialized models like GPT-4V and Med-PaLM M on prior standard report evaluation metrics. MLP multi-layer perceptron. The example chest X-ray image in **b** is obtained from ref. 27 with permission for reproduction from the authors.



Reference findings

AP upright and lateral views of the chest were provided. Left chest wall pacer pack is again seen with leads extending into the right heart. Abandoned pacing leads are also noted in the right chest wall extending into the right heart. **The heart remains moderately enlarged**. **Lung volumes are low**, with equivocal ground-glass opacity on the frontal view, which appears less conspicuous on the lateral view most likely attributable to underpenetrated technique. No gross evidence for pneumonia or pulmonary edema. **No large effusions are seen**. **There is no pneumothorax**. Bony structures are intact.

LLaVA

The image is a black and white radiology picture of a human chest, showing the lungs and the heart. The heart is located on the left side of the image, and the lungs are spread across the middle and right side of the image. The lungs are filled with air, indicating proper breathing. There are several wires visible in the image, which are likely related to medical equipment or monitoring the patient's condition. These wires are positioned around the chest area and are connected to the heart and lungs, suggesting that the patient might be undergoing some form of medical treatment or monitoring.

LLaVA-Med

The chest X-ray image shows bilateral pleural effusion, which means there is an accumulation of fluid in the pleural space on both sides of the lungs. Additionally, there is a right-sided pneumothorax, which is the presence of air in the pleural space, causing the lung to collapse. The image also reveals a left-sided pneumothorax with a chest tube in situ, meaning that a tube has been placed in the pleural space to help remove the air and allow the lung to re-expand.

LLaVA-Rad

Dual-lead left-sided pacer device is again seen, unchanged in position. **The cardiac silhouette remains enlarged**. The aorta is calcified. There are **low lung volumes**. Prominence of the central pulmonary vasculature may be due to mild pulmonary edema. No definite focal consolidation is seen. There is **no large pleural effusion or pneumothorax**.