

“Self-supervised pre-training for disease detection and localization in chest X-ray images”

PI: Guan-Hua Huang (National Yang Ming Chiao Tung University)

Role: Principal investigator

National Science and Technology Council, Taiwan: 08/01/2025-07/31/2026

Chest X-ray (CXR) images are critical for diagnosing and localizing thoracic diseases, yet traditional deep learning approaches for medical image analysis rely heavily on supervised learning, which requires extensive labeled data. Annotating medical images, especially for object detection tasks involving dense bounding box labeling, is time-consuming, costly, and prone to inconsistencies. To address these challenges, this study leverages self-supervised learning (SSL) to pre-train deep learning models for disease detection and localization in CXR images. SSL exploits the intrinsic structure of unlabeled data to learn robust representations without requiring labeled datasets, enabling effective feature extraction for complex tasks.

The project incorporates state-of-the-art object detection models, including two-stage CNN-based detectors (Faster R-CNN), one-stage CNN-based detectors (YOLO), and transformer-based detectors (DETR), tailoring them to the unique challenges of medical imaging. We employ a range of SSL methods for pre-training, including SimCLR for backbone-only learning, SoCo for pre-training FPN and RoI heads, ADePT for RPN components, and DETReg for detection heads in transformer-based models. These methods are applied to general-purpose datasets such as ImageNet and MS COCO and large unlabeled medical datasets like NIH ChestX-ray and Stanford CheXpert, capturing the distinct characteristics of medical images. Weakly supervised object localization techniques are also used to generate pseudo-bounding boxes from image-level labels, enabling efficient pre-training for object detection. To further enhance performance, we integrate pre-trained weights from large CXR datasets with ground-truth bounding boxes, including VinDr-CXR and CXR-AL14. By aligning the source and target domains, this approach reduces the need for fine-tuning and improves the accuracy of disease detection and localization. This comprehensive framework aims to advance CXR image analysis by addressing data scarcity and annotation challenges, paving the way for more effective medical image object detection.

Keywords: chest X-ray image, deep learning, object detection, self-supervised learning, weakly supervised object localization