

# Highlights:

- A two-pass approach to localize and characterize radiologic markers is fast and reliable for image anonymization.
- Finetuning the localizer network increases de-identification performance to 99.6% accuracy (p-value = 0.04).
- Selective retention of markers enables granular control over image de-identification.

# Anonymizing Radiographs and Making Them “AI Ready” Using a Deep Learning Algorithm

Bardia Khosravi, MD MPH, MHPE<sup>1,2</sup>; John Mickley, MD<sup>2</sup>; Pouria Rouzrokh, MD, MPH, MHPE<sup>1,2</sup>; Michael Taunton, MD<sup>2</sup>; A. Noelle Larson, MD<sup>2</sup>; Bradley J. Erickson, MD, PhD<sup>1</sup>; Cody Wyles, MD<sup>2</sup>

1. Department of Radiology, Mayo Clinic, Rochester, MN

2. Department of Orthopedic Surgery, Mayo Clinic, Rochester, MN

## INTRODUCTION

Radiographs play a crucial role in medical care by providing valuable information about a patient's anatomy and potential pathologies. Most radiographs contain radio-opaque markers that display information such as the side, positioning, radiographer initials, and often patient identifier. In this study, we develop a deep learning (DL) model that detects and selectively removes these markers from images to enable de-identified data sharing and improve the suitability of images for training robust deep learning models.

## METHODS

To create our ground-truth dataset, we annotated 2000 pelvic radiographs collected between 2000 and 2021. Each marker area was annotated with a bounding box. The data was then split at the patient level into 60%/20%/20% for training, validation, and testing, respectively. Images were resampled to 512x512 pixels, and a YOLOv5-x model was trained to localize markers. Mean average precision at a 50% threshold (mAP-50) was reported as a measure of model performance on the test set. While removing identification markers is important, laterality markers contain valuable information that helps guide the pre-processing pipeline of deep learning models.

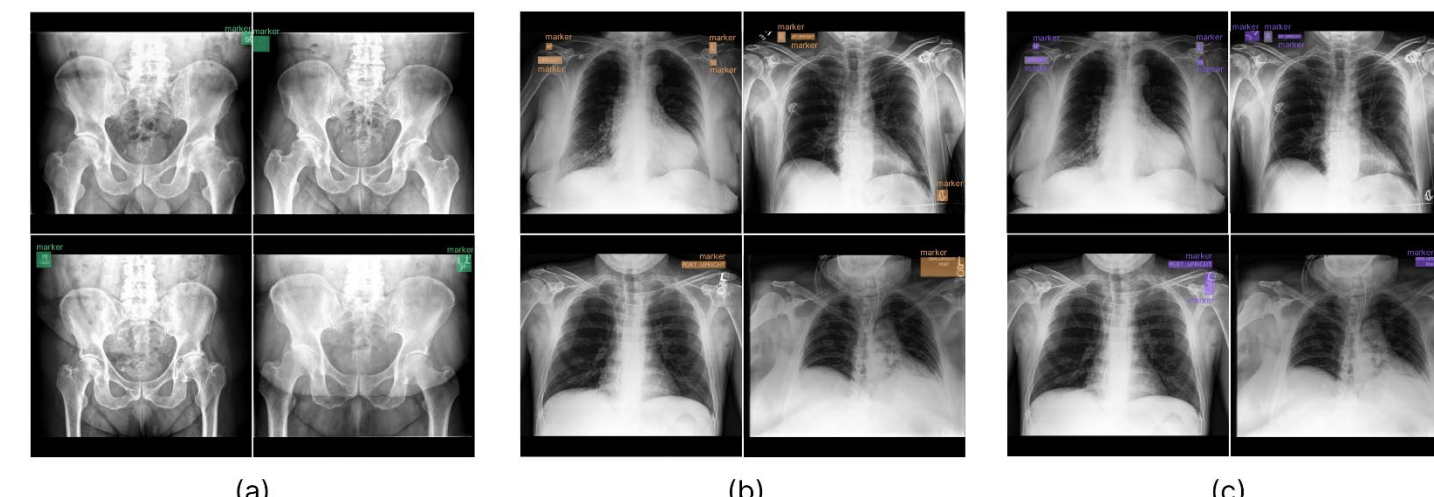
We developed an image processing algorithm that examines each marker area and sends it to an OCR algorithm for characterization. In our case, laterality markers were retained if "R" and "L" characters were detected in the image. To externally validate the model, we applied it to chest radiographs from the validation set of the CheXpert dataset to assess its performance on data from different organization.

## RESULTS

The model reached a mAP-50 of 0.99 and 0.98 on the validation and internal test sets, respectively. It had a de-identification accuracy of 100% (400/400) on the test set, with a false-positive rate of 1% (8/632). The retention accuracy was 93% (359/386). On the external set of chest radiographs, the model successfully removed markers from 96% of radiographs (221/231). After fine-tuning on 20 CheXpert radiographs, the model was able to remove markers from 99.6% (230/231) of chest radiographs (p-value=0.04).

## CONCLUSION

This work develops a deep learning model that can robustly and selectively detect and remove radiographic markers from radiographs of different anatomical regions, yielding images that are anonymized and AI ready.



**Figure.** Marker detection on pelvis radiographs (a). Results of the localizer model's performance on out-of-domain chest x-rays before (b) and after (c) finetuning. As shown, the finetuned model draws tighter bounding boxes and also has more specificity compared to the initial model.

## REFERENCES

- Barry K, Kumar S, Linke R, Dawes E. A clinical audit of anatomical side marker use in a paediatric medical imaging department. J Med Radiat Sci [Internet]. 2016 Sep;63(3):148–54.
- Edwards A, Hollin I, Barry J, Kachnowski S. Barriers to cross-institutional health information exchange: a literature review. J Healthc Inf Manag [Internet]. 2010 Summer;24(3):22–34.

## CONTACT US

Bardia Khosravi, MD: [khosravi.bardia@mayo.edu](mailto:khosravi.bardia@mayo.edu)

Bradley J. Erickson, MD PhD: [bje@mayo.edu](mailto:bje@mayo.edu)

Cody Wyles, MD: [wyles.cody@mayo.edu](mailto:wyles.cody@mayo.edu)