MoCo-CXR: MoCo Pretraining Improves Representation and Transferability of Chest X-ray Models

Hari Sowrirajan, Jingbo Yang*, Andrew Y. Ng, Pranav Rajpurkar*

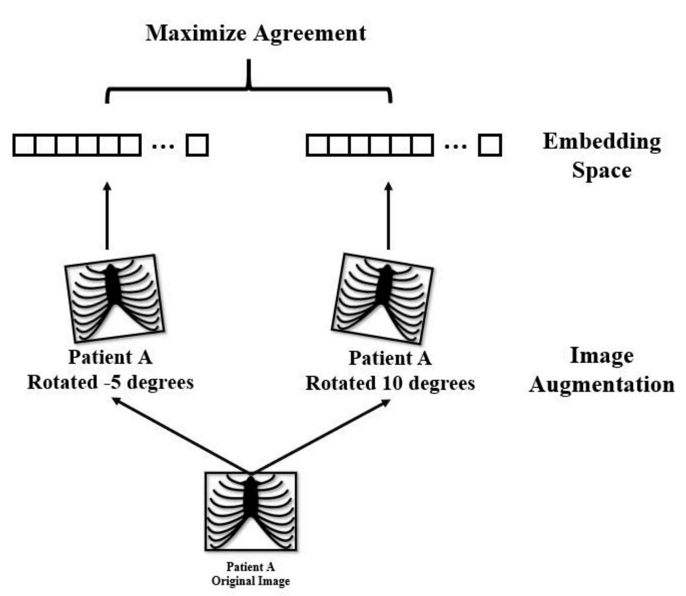
Department of Computer Science, Stanford University

Stanford
Computer Science

Background

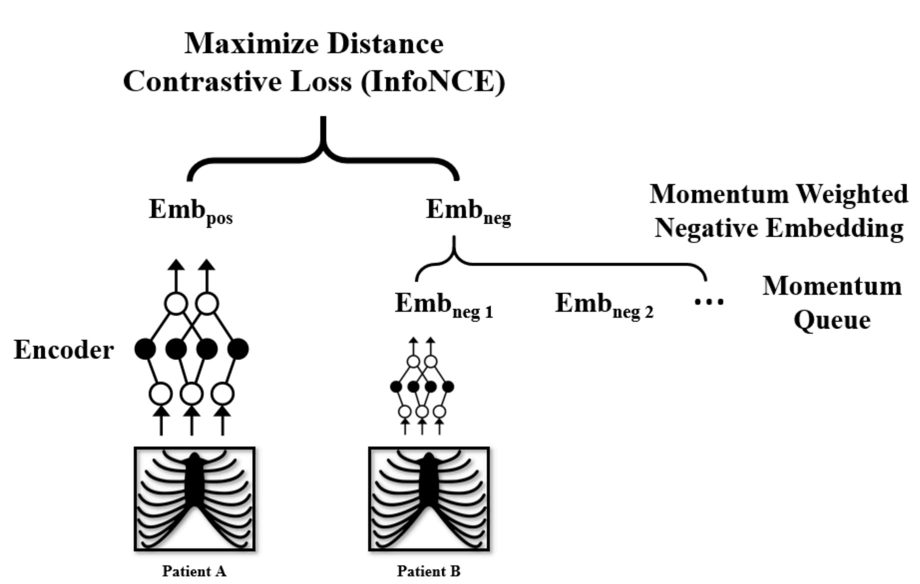
Contrastive Learning

- A form of self-supervision capable of unsupervised visual representation learning
- Learns representation by contrasting augmented versions of an image against negative pairs



Momentum Contrast (MoCo)

- A frontrunner for contrastive learning
- Utilizes a queue to maintain dictionary of negative samples – keeps **batch size small** yet maintains a **diverse set of negatives**



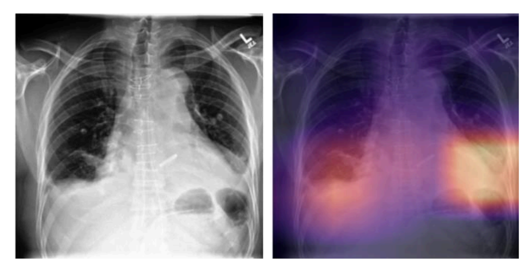
Chest X-ray Interpretation

- Critical for screening, diagnosis, and management of diseases
- Labeled datasets like **CheXpert** enabled deep learning models to classify chest X-rays
- Contrastive learning has **never been applied** to leverage unlabeled X-rays for such models

Datasets

CheXpert

- 224,316 chest radiographs of 65,240 patients
- Labeled for presence/absence of several diseases
- Pleural Effusion has high prevalence (45.63% of all images have this label



Frontal Radiograph (left) of Pleural Effusion and heatmap localizations of the effusions (right)

Shenzhen

- 662 X-ray images: 336 contain Tuberculosis

MoCo-CXR Adaptation

Model Initialization

- Initialized with ImageNet-pretrained weights – shown to have convergence benefits

Augmentations

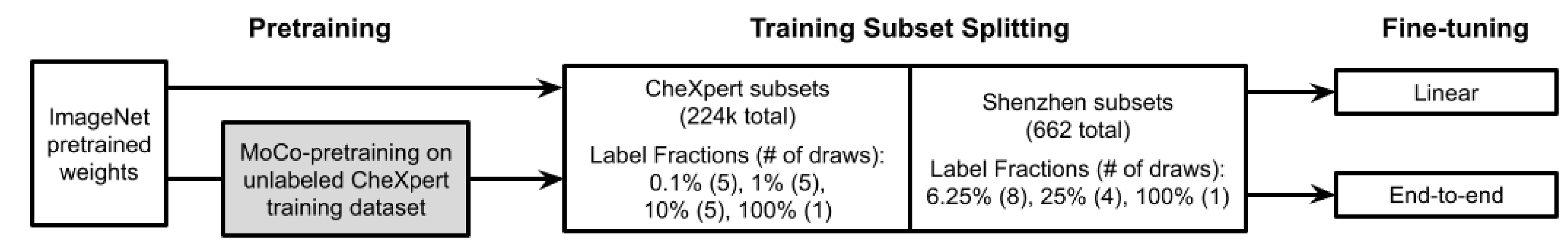
- Chose clinically appropriate image augmentations

Original | Random Rotation | Horizontal Flip

Learning Rate / Batch Size

- Adapted optimal learning rate w/ milestone scheduling and batch size for chest X-rays
- Optimal LR/Batch size much smaller than for natural images in original MoCo experiments

End-to-End MoCo-CXR Pipeline



- Performed MoCo-CXR pretraining on CheXpert unlabeled dataset
- Prepared counterpart ImageNet-pretrained models for comparison

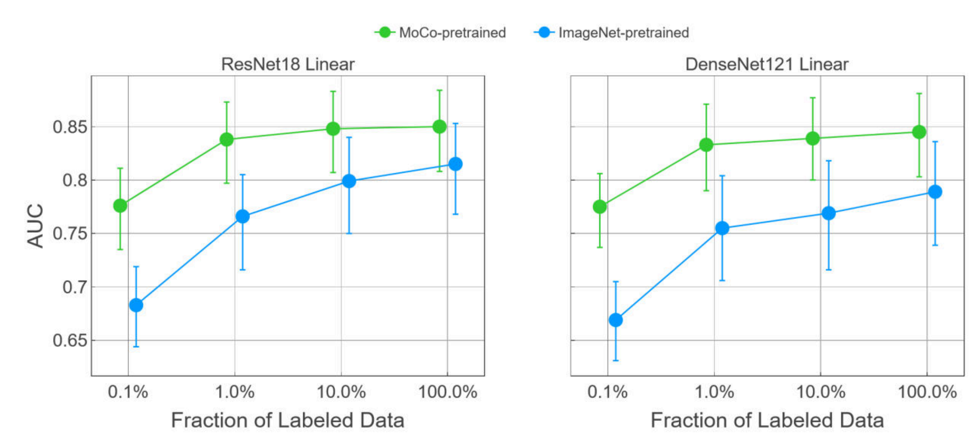
Fine-tuning with label fractions

- Fine-tuned models with different fractions of labeled training data (CheXpert / Shenzhen)
- Trained linear and End-to-end models to assess representation and initialization quality respectively
- Evaluation of models conducted on CheXpert / Shenzhen test sets

Experiments

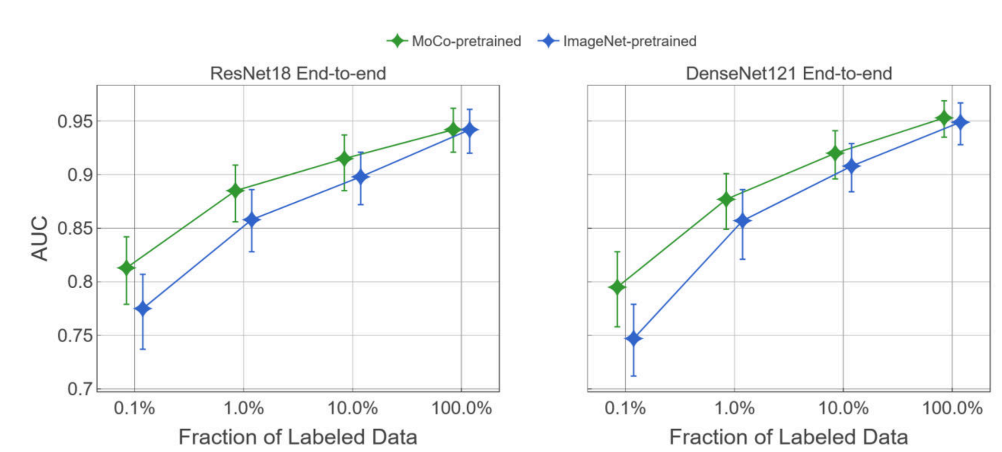
Transfer Performance of Linear Classifier

- Linear classifier trained on a frozen base model
- Clear but diminishing improvement with higher fractions of labeled data
- AUC improvement on ResNet 18 is 0.096 (95% CI 0.061, 0.130) with 0.1% labelled data and 0.034 (95% CI -0.009, 0.080) at 100% labeled data.



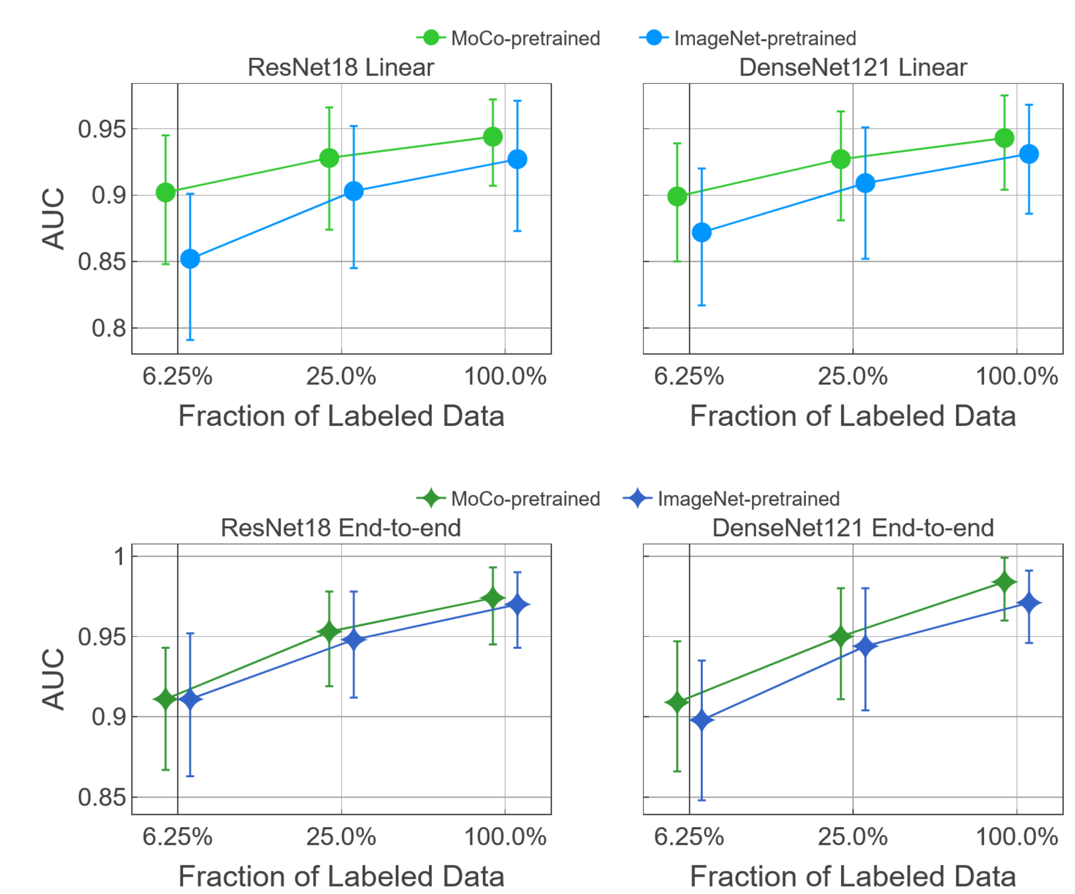
Benefits of End-to-End Training

- ResNet 18 achieved small but statistically significant improvement of 0.037 (95% CI 0.015 0.062) at 0.1% label fraction
- AUC converges to 0.942 at 100% label fraction for both backbones with and without MoCo-CXR pretraining.



Transfer Performance to an External Dataset

- Verify whether representations learned from CheXpert transfer to Shenzhen tuberculosis task.
- Fine-tuning last layer of ResNet 18 backbone achieved AUC improvement of 0.054 (95% CI 0.024 0.086) at 6.25% label fraction, representing an approximately 25 images in total.
- Much less AUC improvement for models trained end-to-end, potentially due to limited learning potential with few images.



Conclusion

- MoCo-CXR-pretraining provides high-quality representations and transferable initializations for chest X-ray interpretation.
- MoCo-CXR displayed consistent benefit across multiple training label fractions and transferred to an external, target dataset.
- Successfully adapted MoCo despite numerous data and task differences between natural image classification and chest X-ray interpretation.
- Only requires a single NVIDIA GTX 1070.
- Suggests application to other medical imaging tasks with scarce labeled data but abundant unlabeled data.

References

[1] Jeremy Irvin, Pranav Rajpurkar, Michael Ko, Yifan Yu, Silviana Ciurea-Ilcus, Chris Chute, HenrikMarklund, Behzad Haghgoo, Robyn Ball, Katie Shpanskaya, et al. Chexpert: A large chestradiograph dataset with uncertainty labels and expert comparison. InProceedings of the AAAIConference on Artificial Intelligence, volume 33, pages 590–597, 2019.
[2] Xinlei Chen, Haoqi Fan, Ross Girshick, and Kaiming He. Improved baselines with momentumcontrastive learning.arXiv preprint arXiv:2003.04297, 2020c.
[3] Simon Kornblith, Jonathon Shlens, and Quoc V. Le. Do better imagenet models transfer better? InProceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),June 2019.
[4] Stefan Jaeger, Sema Candemir, Sameer Antani, Y'ı-Xi'ang J W'ang, Pu-Xuan Lu, and George Thoma.Two public chest x-ray datasets for computer-aided screening of pulmonary diseases.Quantita-tive imaging in medicine and surgery, 4(6):475, 2014.

*Contributed equally to this work