Quantifying the Value of Lateral Views in Deep Learning for Chest X-rays

Mohammad Hashir[12], Hadrien Bertrand[1], and Joseph Paul Cohen[12]

[1] Mila, Quebec AI Institute [2] University of Montreal

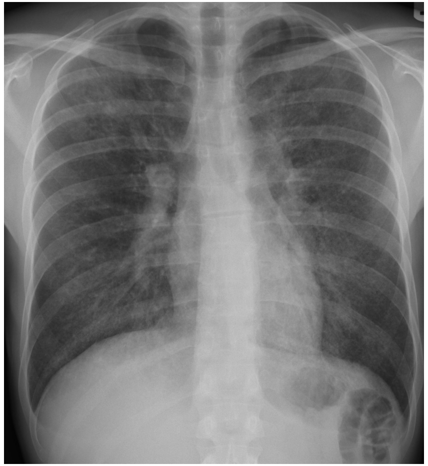
https://arxiv.org/abs/2002.02582

MIDL
Montréal 2020

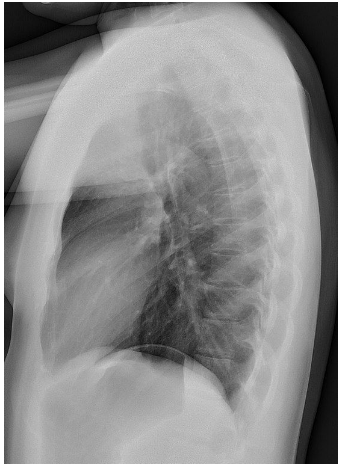
The lateral view

The L view contains information missing in the PA view that is relevant for diagnosis [1].

Most chest X-ray datasets have only the PA view, but some recent ones have also the L view.



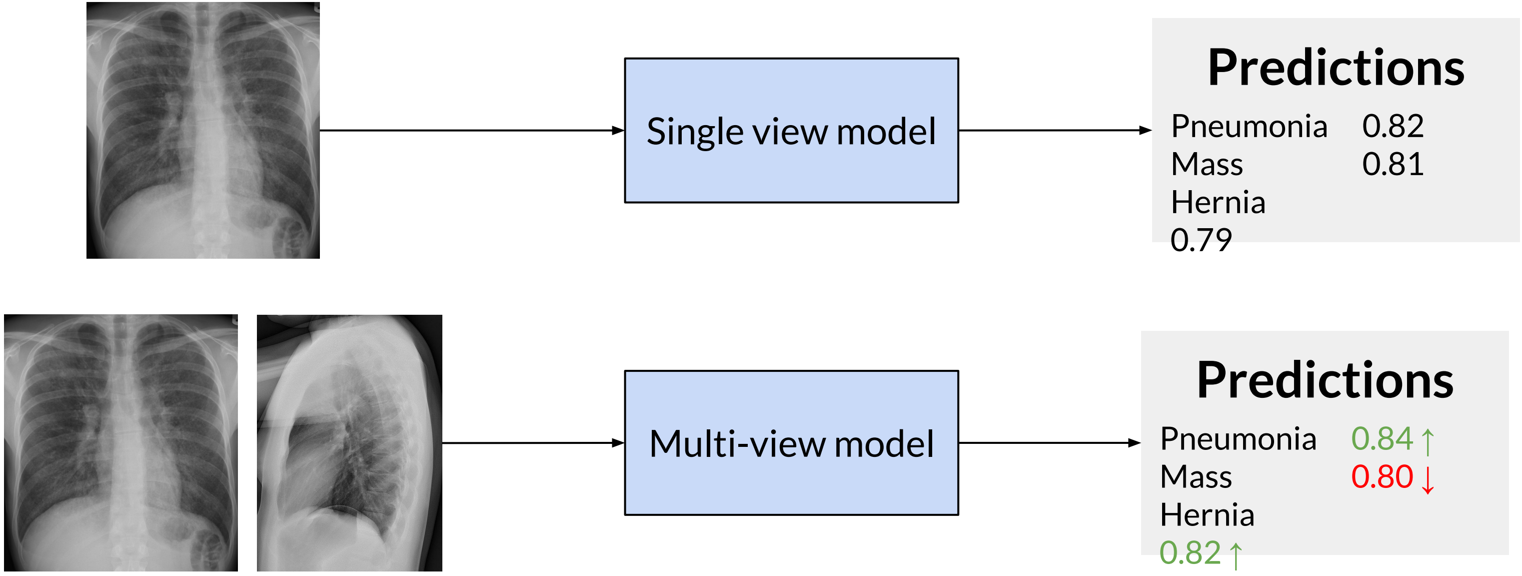
Postero-anterior (PA)



Lateral (L)

Task

Evaluate the contribution of a paired lateral view in chest X-ray prediction and find the best multi-view model



Our work

We explore the two questions

- Does a paired lateral view help in prediction? If so, for which labels specifically?

- Instead of having a paired lateral view, is it better to increase training set with more PA samples and use a single view model?

Materials and methods

Dataset and preprocessing

PadChest [2]

160k images from 67k Spanish patients.

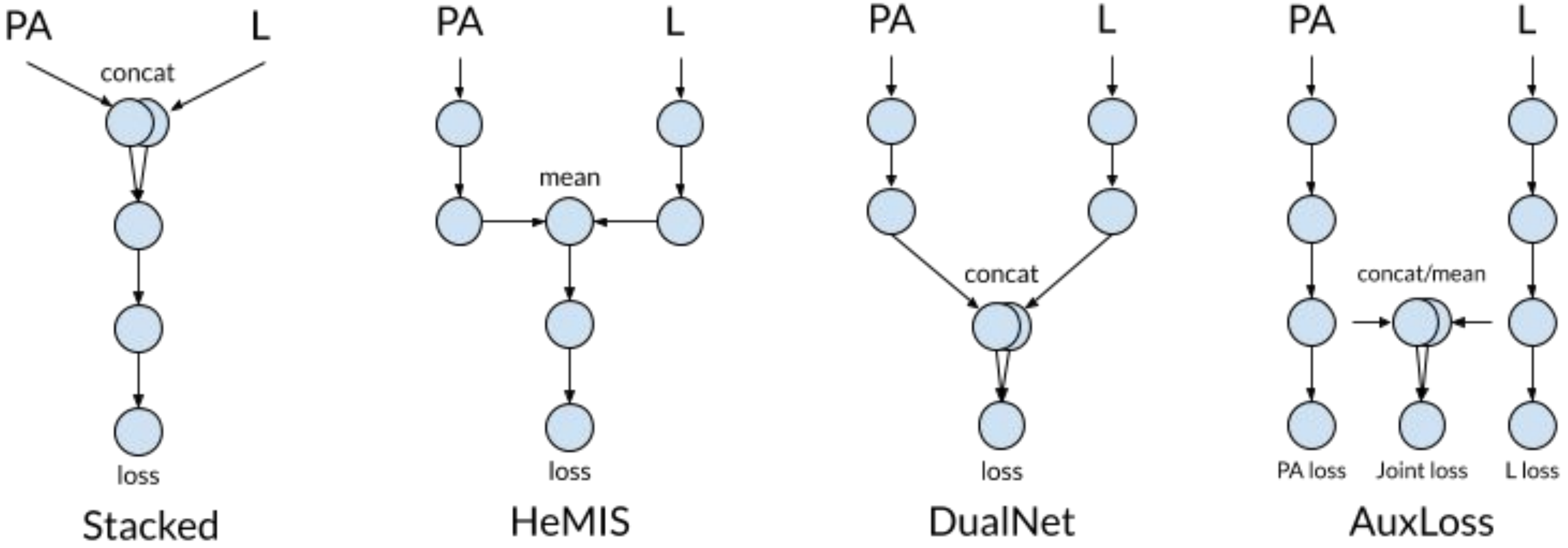
Multiple labels per image from total 194.

Preprocessing

- Keep patients with paired PA and L views: **total 31k**
- Keep labels affecting 50+ patients: **total 64**.
- Images resized to 224x224 and pixels rescaled to [-1, 1]

Models

Based on DenseNet blocks [3]. Baseline is single view DenseNet-121



Havaei et al., 2016 [4]

Rubin et al., 2018 [5]

Our contribution

Experiments and results

Performance of multiview models

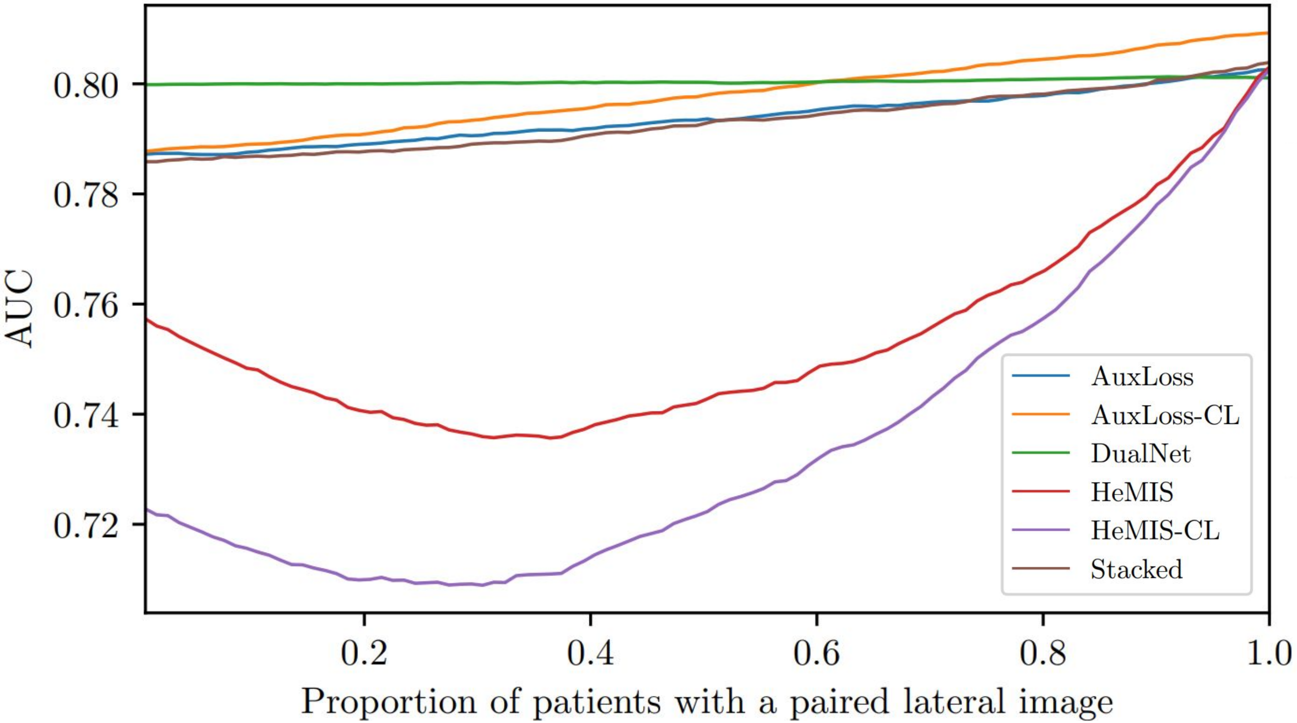
All joint view models perform better than single view models.

Table 1: Test AUC achieved by the different models, averaged over five runs with standard deviation also reported. DenseNet is trained on a single view denoted by the suffix, rest are trained on both views. The symbol in the superscript indicates the difference between a pair of models in the same column is statistically significant.

Model	**Test AUC** Both	PA	L
DenseNet-L	—	—	0.780 ± 0.004
DenseNet-PA	—	0.793 ± 0.007	—
Stacked	$0.804 \pm 0.003^{*}$	$0.786 \pm 0.009^{\ddagger}$	0.595 ± 0.046
DualNet	$0.801 \pm 0.003^{\dagger}$	$\mathbf{0.800 \pm 0.004}^{\dagger*\ddagger}$	0.539 ± 0.018
HeMIS	0.803 ± 0.006	0.758 ± 0.014	0.603 ± 0.044
HeMIS-CL	0.803 ± 0.007	0.723 ± 0.017	0.627 ± 0.036
AuxLoss	0.803 ± 0.006	$0.787 \pm 0.005^{*}$	0.753 ± 0.002
AuxLoss-CL	$\mathbf{0.809 \pm 0.003}^{*\dagger}$	$0.788 \pm 0.005^{\dagger}$	$\mathbf{0.771 \pm 0.003}$

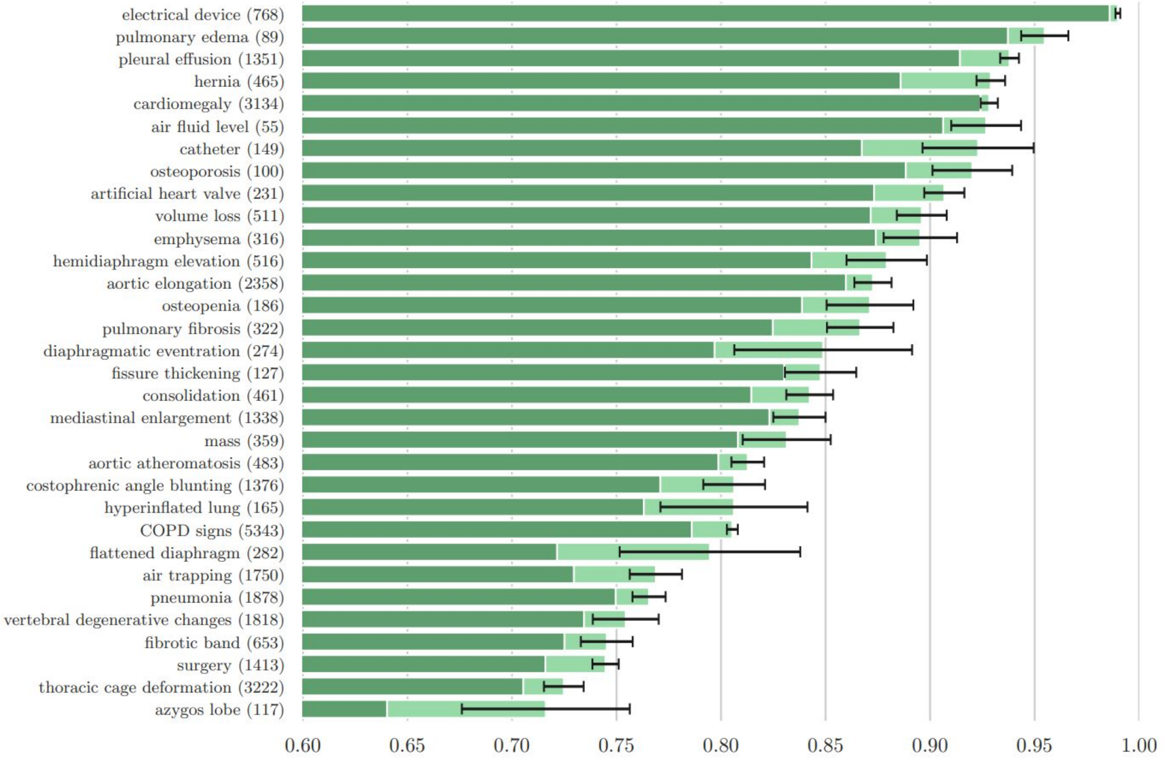
Utilization of the lateral view

Change in AUC as proportion of patients with paired lateral views increase



Label-wise increase with L view

32/64 labels see an improvement in AUC with AuxLoss



More PA samples

We add 18k patients to the training set that have a PA view but no L view.

Table 2: AUC and standard deviation of the DenseNet-PA and AuxLoss models trained on the main and extended dataset but evaluated on the same main test set. The size of the training set for the extended dataset is double that of the main. The Main AUC column copies the values from Table 1 for easier reference.

		Train Data (AUC on test reported)	
Model	**Test Data**	**Main**	**Extended**
DenseNet-PA	PA	0.793 ± 0.007	0.813 ± 0.005
AuxLoss-CL	PA	0.788 ± 0.005	0.812 ± 0.006
AuxLoss-CL	Both	0.809 ± 0.003	0.772 ± 0.018

Conclusion

Takeaways

- Multi-view models significantly better than single view overall
- 32 labels improve with multi-view model
- Doubling PA images in training set -> change in AUC not significant

Thank you

arxiv.org/abs/2002.02582

References

[1] Raoof, Suhail, et al. "Interpretation of plain chest roentgenogram." Chest 141.2 (2012): 545-558.

[2] Bustos, Aurelia, et al. "Padchest: A large chest x-ray image dataset with multi-label annotated reports." arXiv preprint arXiv:1901.07441 (2019).

[3] Huang, Gao, et al. "Densely connected convolutional networks." Proceedings of the IEEE conference on computer vision and pattern recognition. 2017.

[4] Havaei, Mohammad, et al. "Hemis: Hetero-modal image segmentation." International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer, Cham, 2016.

[5] Rubin, Jonathan, et al. "Large scale automated reading of frontal and lateral chest x-rays using dual convolutional neural networks." arXiv preprint arXiv:1804.07839 (2018).

Appendix

Why AuxLoss

Multiview models at test time perform similarly when given both views but diverge significantly when given only one view

Advantages of AuxLoss

- Uses both views productively
- Robust to missing views
- Lowest variance across multiview models
- Less sensitive to hyperparameter changes

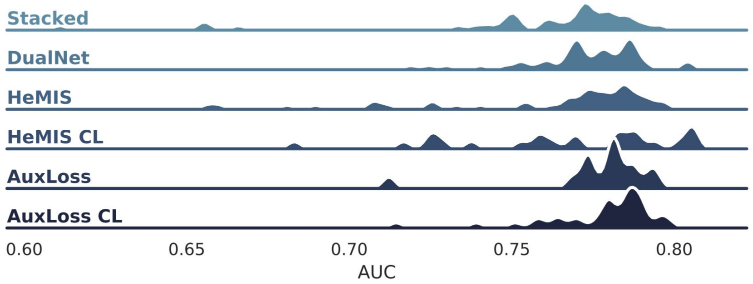


Figure 4: Distributions of AUC for a 40 combination hyperparameter search for each model. Some models are much more robust to hyperparameter changes than others.

Training details

Hyperparameters found through extensive search

- 40 epochs, batch size of 8 and Adam optimizer
- Early stopping on validation AUC
- Loss weighted by class frequency (clamped at 5.0 max)
- Learning rate scaled by 0.1 every 10 epochs but initial LR different for every model
- Curriculum learning: views dropped randomly for Hemis and AuxLoss
- Dropout of 0.1-0.2

Label-wise increase with more PA samples

32 labels

22 overlap with AuxLoss