Comparison of Supervised and Unsupervised Techniques for Computer Assisted Decision Support in Medical Imaging

Thomas JS Durant, MD[1,2]; George C. Linderman, BS[1]; Isaac G. Freedman, BPhil, MPH[2]; Harlan M. Krumholz MD, SM[1,3], Wade L Schulz, MD, PhD[1,2]

1 – Yale University School of Medicine, Department of Laboratory Medicine; 55 Park Street PS345D, New Haven, CT 06511

2 – Center for Outcomes Research and Evaluation; 1 Church St #200, New Haven, CT 06510

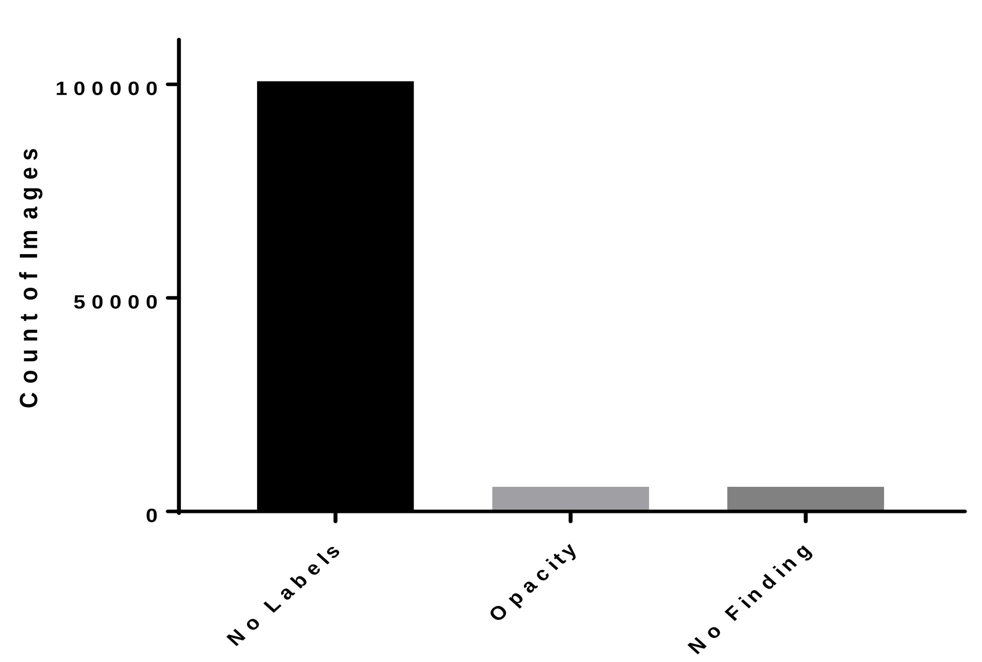
3 – Yale University School of Medicine, Department of Internal Medicine, 330 Cedar St, Boardman 110, New Haven CT, 06520

BACKGROUND

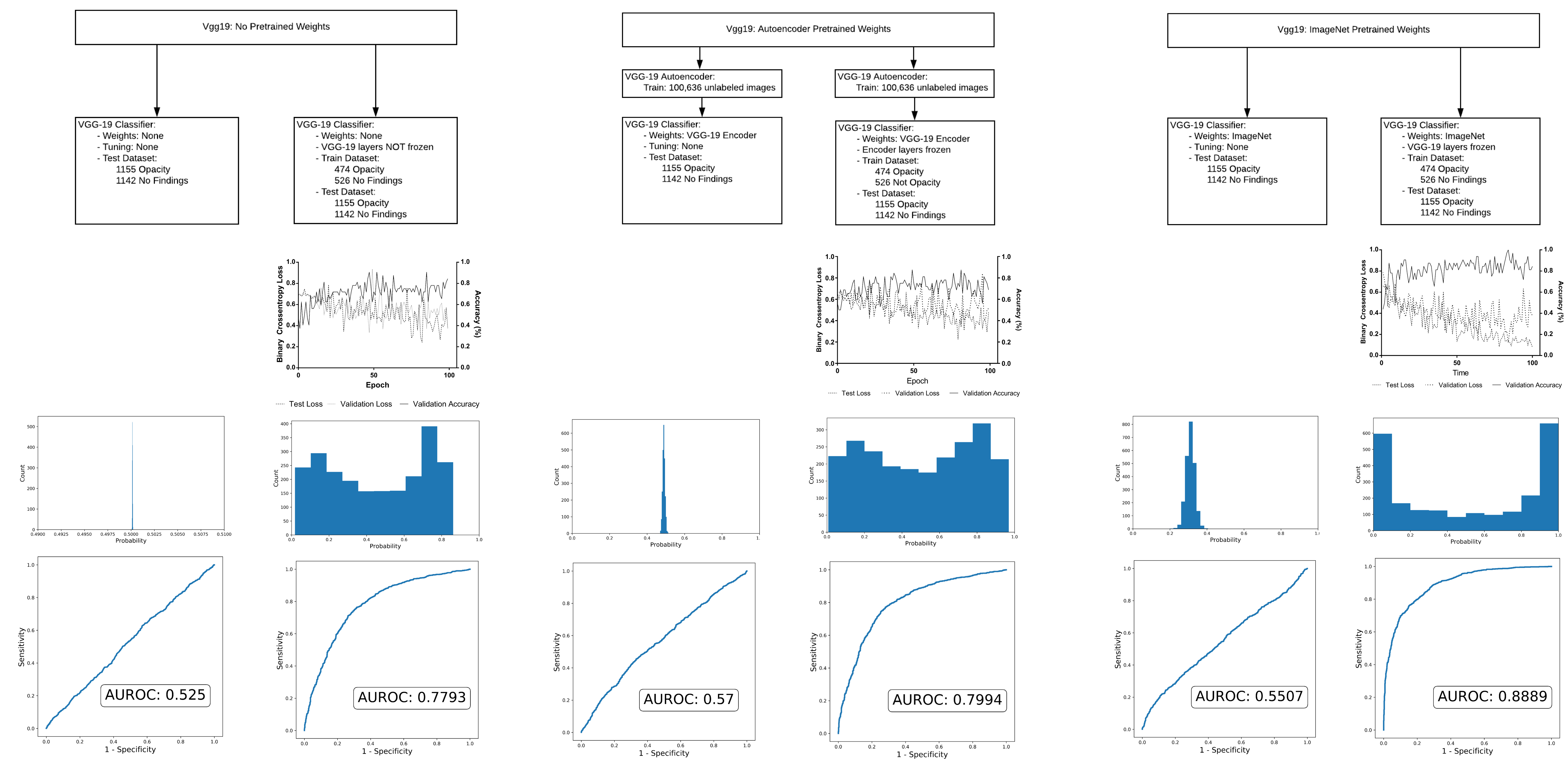
Over the past decades, digital imaging studies and imaging modalities in healthcare have expanded in both volume and complexity, with a projected growth rate which may be unsustainable for the interpretive capacity of physicians. Accordingly, three is significant interest in the development and implementation of computer-assisted decision support (CADS) tools. However, initial published efforts rely on supervised machine learning methods, which limits the development of CADS tools due to the cost and time-prohibitive nature of curating labeled data sets. Emerging unsupervised and semi-supervised machine learning technologies may offer the functionality needed to facilitate the development of performant decision support tools. To this end, we evaluate the utility of available unsupervised machine learning, where the availability of reliable, highly annotated data is often the limiting factor.

METHODS

- VGG-19 was chosen as the convolutional neural network architecture due to it's non-branching structure, compatible with denoising autoencoder (DA) implementation.
- 112,120 labeled Chest Radiographs were obtained from the publically available National Institutes of Health Chest X-Ray Dataset.
- The DA training dataset consisted of a total of 100,636 unlabeled images.
- Opacity was created as a label which combined at least two of the following labels: 'Pneumonia', 'Infiltration', 'Atelectasis', 'Consolidation' for the test dataset, which consisted of a total of 5,742 'opacity' and 'not opacity' labeled images.



RESULTS



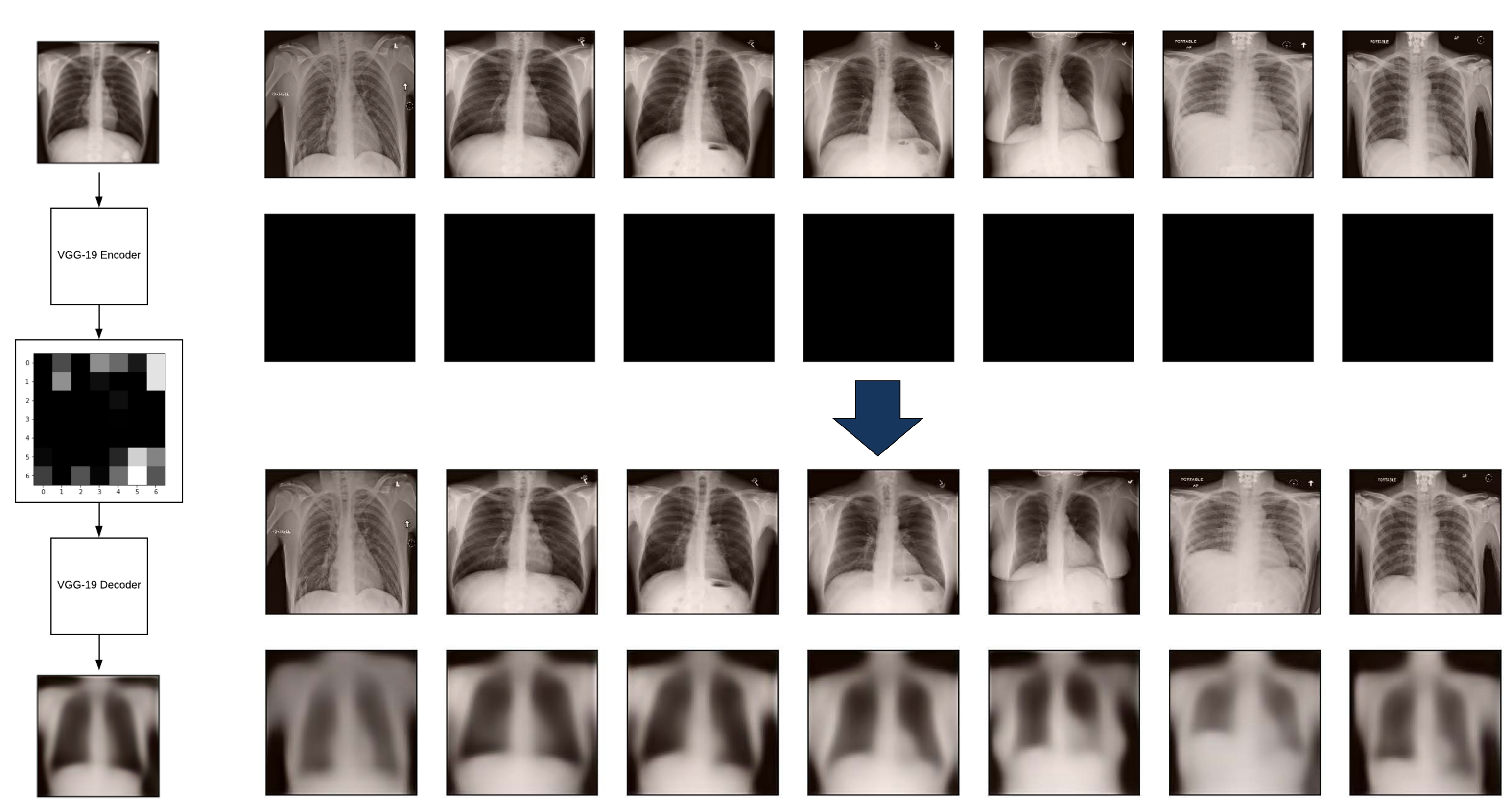


Fig: The encoder block finds latent representations of input data, which are used by the decoder block for image reconstruction. The resulting DA architecture offers a model representing features of input data derived from the relative position of pixels within the image.

CONCLUSIONS

- This work demonstrates that denoising autoencoder techniques offer potentially relevant feature representations to the classifier block, more so than random initialization, which may augment performance metrics with subsequent classification problems.
- Future work will seek to evaluate denoising autoencoder techniques with an equivalent number of ImageNet images and unlabeled CXRs.
- Further, these findings support prospective evaluation of emerging unsupervised or semi-supervised machine learning techniques in the context of medical imaging, and further elucidate the potential for context-specific performance benefits, as this may help to facilitate the development of decision support tools in clinical medicine.

REFERENCES

1. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 2017, ChestX-ray8_Hospital-Scale_Chest_CVPR_2017_paper.pdf
2. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014)