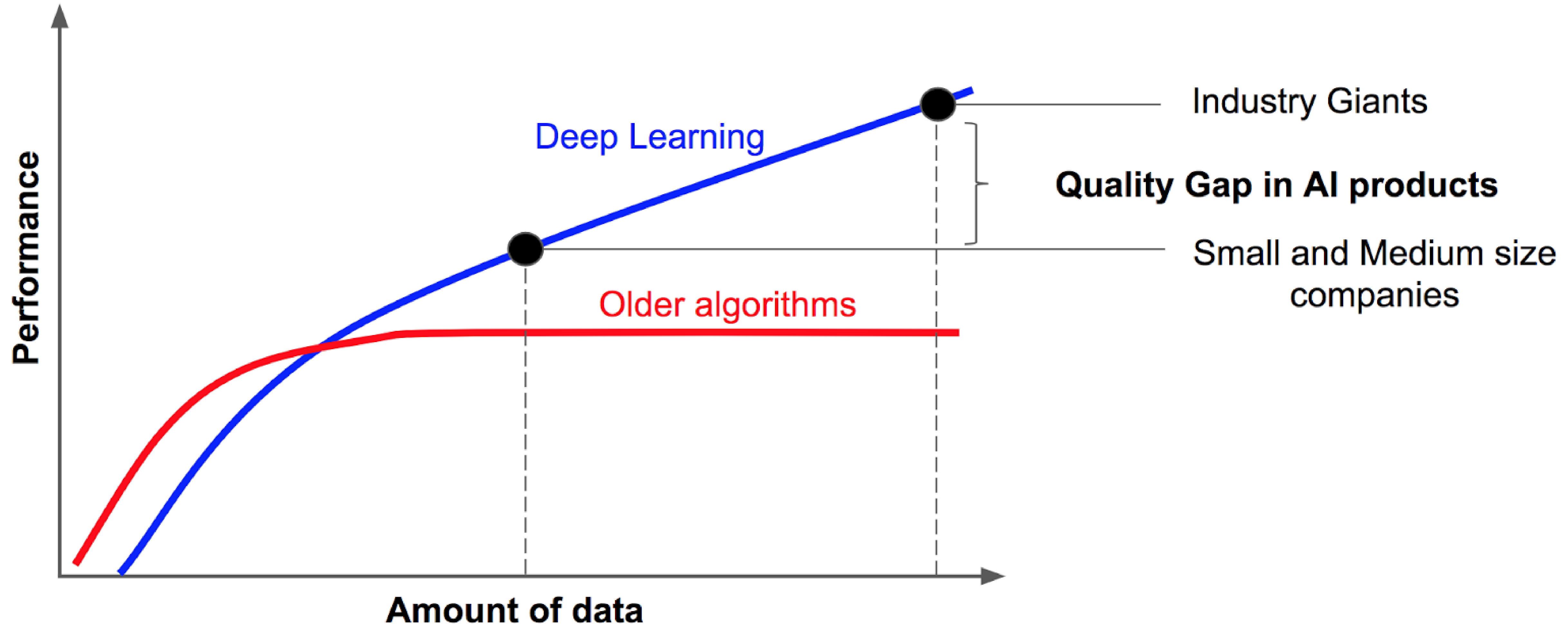
Lister Hill National Center for Biomedical Communications

NIH U.S. National Library of Medicine
Lister Hill National Center for Biomedical Communications
USA.gov

Deep neural ensembles for improved pulmonary abnormality detection in chest radiographs

Dr. Sivaramakrishnan Rajaraman
National Library of Medicine

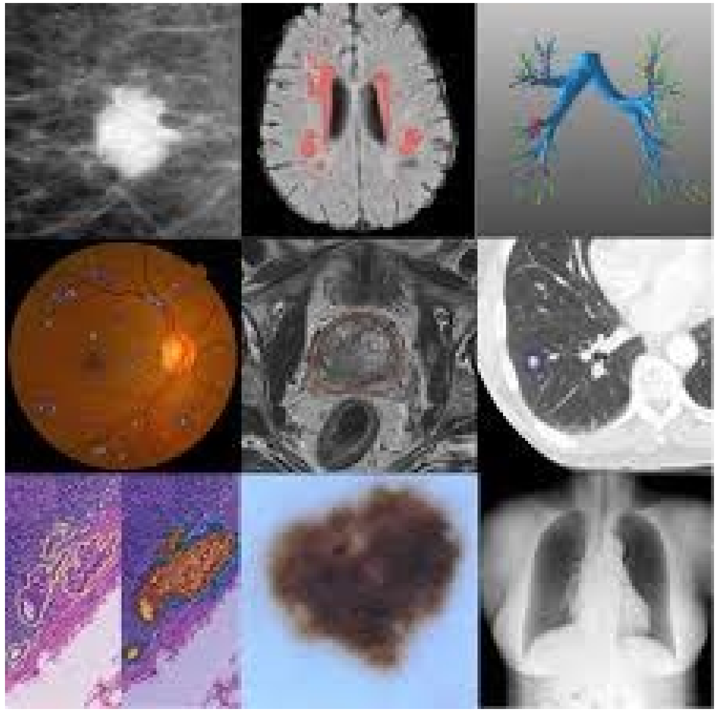
Deep Learning – The real need



Challenges in applying Deep Learning to Medical Image Analyses

a) Limited availability of medical imaging data: Development of massive training dataset with expert annotations is a laborious time consuming task

b) Lack of standard models: No dedicated models available for medical computer vision.



Continued …

c) Data Interoperability and Standards: Nature of data differ from hardware to hardware; Huge variation in images due to sensors and other factors. Need to combine several datasets for better algorithms learning and accuracy.

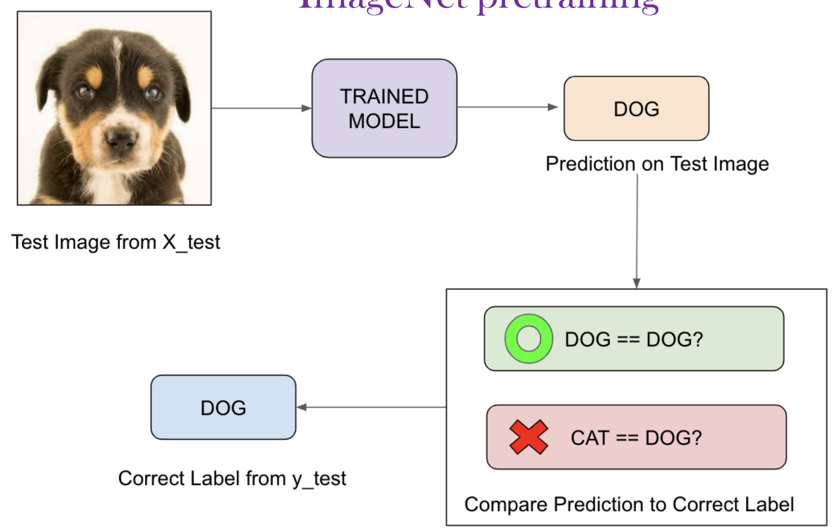
d) Privacy and Legal Issues: Sharing of medical data is severely complex and difficult.

e) Uninterpretable Black Box Model: Weight matrices created with increase in layer depth makes the model uninterpretable. Need to explain predictions to faithfully supplement clinical decision making.

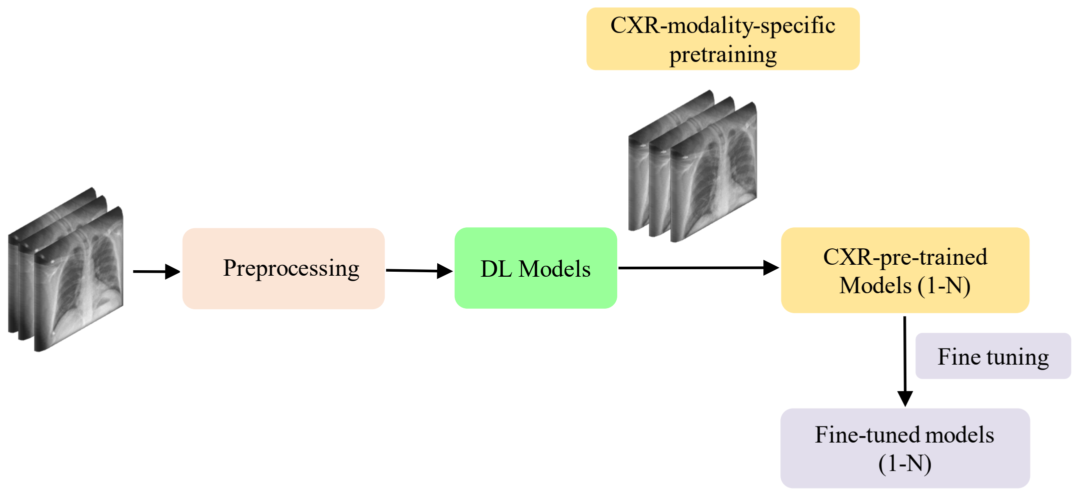
Modality-specific Learning

a) Visual characteristics of medical images are **different** than in natural images.
b) Example -- CXRs
   a) have highly localized ROI
   b) exhibit high inter-class similarity and intra-class variance.
c) Datasets too small for conventional transfer learning to be reliable.
d) Pretraining on modality-specific data improves generalization and performance for related target tasks that use smaller datasets.

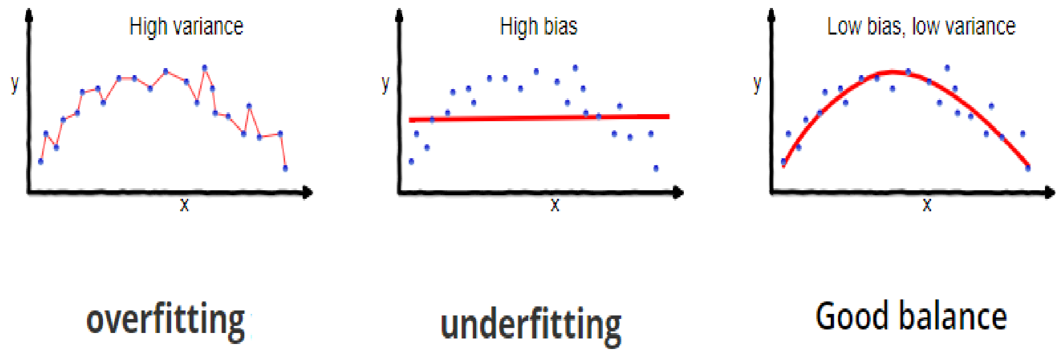
ImageNet pretraining

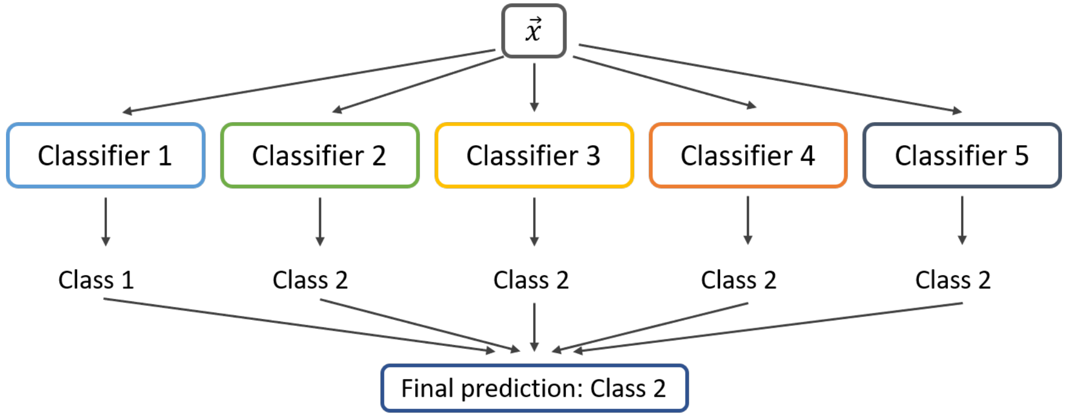


Modality-specific pretraining



Ensemble Learning





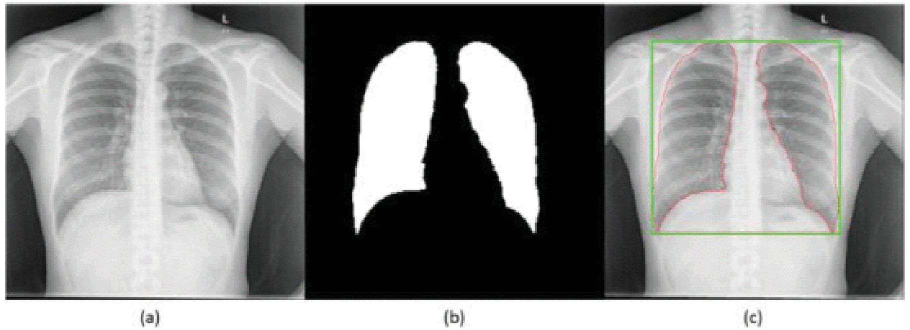
a) To combine ACCURATE and DIVERSE classifiers and improve predictions.

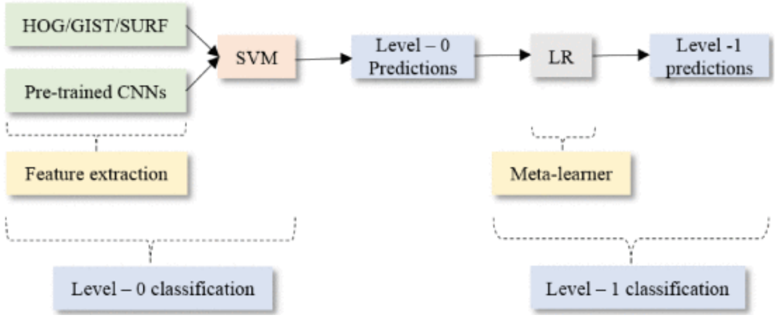
b) Resolve **BIAS-VARIANCE** trade-off.

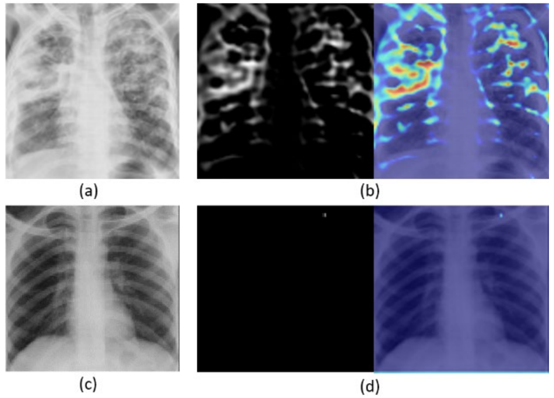
A Current Research Interest

- Study benefits of combining modality-specific model training and ensemble learning for improving task-related performance and generalization through:
- (a) transfer modality-specific knowledge to improve performance in a related task;
- (b) reduce prediction variance, sensitivity to the training data, and model overfitting.

Study 1: A novel stacked generalization of models for improved TB detection in chest radiographs





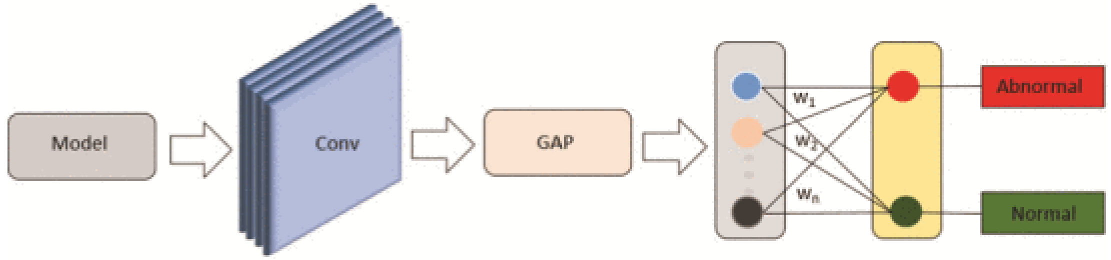


a) Trained an SVM using hand-crafted (HOG/GIST/SURF) and DL features.
b) Ensemble method: Stacking
c) Stacked ensemble learning delivered superior performance.

		Literature				Proposed approaches	
		[5]	*[16]*	*[13]*	*[8]*	*E[P1]*	*E[P1P2]*
S	**Acc**	0.840	0.837	0.847	-	**0.934**	**0.934**
	AUC	0.900	0.926	0.926	-	0.955	**0.991**
M	**Acc**	0.783	0.674	0.826	-	**0.875**	**0.875**
	AUC	0.869	0.884	0.926	-	0.875	**0.962**
K	**Acc**	-	-	-	-	0.733	**0.776**
	AUC	-	-	-	-	0.825	**0.826**
I	**Acc**	-	-	-	0.943	**0.960**	**0.960**
	AUC	-	-	-	0.960	0.960	**0.965**

S. Rajaraman et al., "A novel stacked generalization of models for improved TB detection in chest radiographs," 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, 2018, pp. 718-721, doi: 10.1109/EMBC.2018.8512337.

Study 2: Assessment of an Ensemble of Machine Learning Models Toward Abnormality Detection in Chest Radiographs



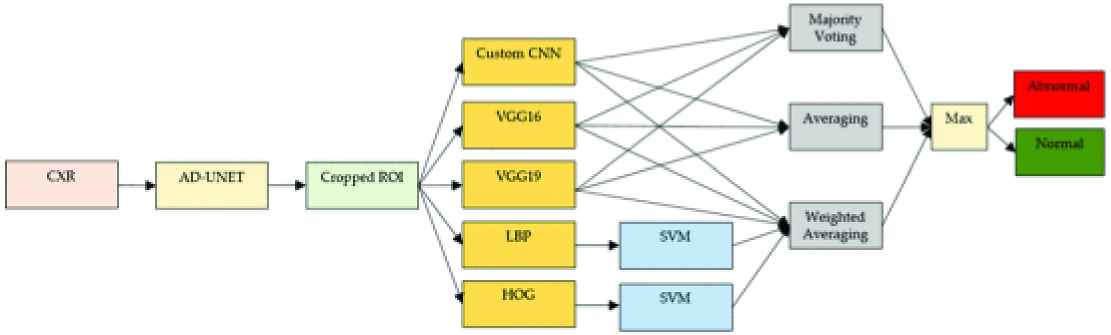
Fold	*Weighted averaging*			
	Acc	***AUC***	***F***	***MCC***
1	98.7	100.0	99.1	97.0
2	98.6	100.0	99.1	96.8
3	98.6	100.0	99.1	96.7
4	98.5	99.9	99.0	96.5
5	98.6	100.0	99.0	96.8
Mean	98.7	100.0	99.1	96.8
SD	0.78	0.02	0.05	0.18

a) Ensemble predictions of handcrafted feature descriptors/classifiers and DL models.

b) Handcrafted: used LBP/HOG feature descriptors to train SVM classifier; DL models: Custom, VGG-16, VGG-19.

c) Ensemble methods: Majority voting, averaging, and weighted averaging

d) Weighted averaging resulted in superior performance.



S. Rajaraman et al., Assessment of an ensemble of machine learning models toward abnormality detection in chest radiographs. Conf Proc IEEE Eng Med Biol Soc. 2019;2019:3689-3692. doi:10.1109/EMBC.2019.8856715

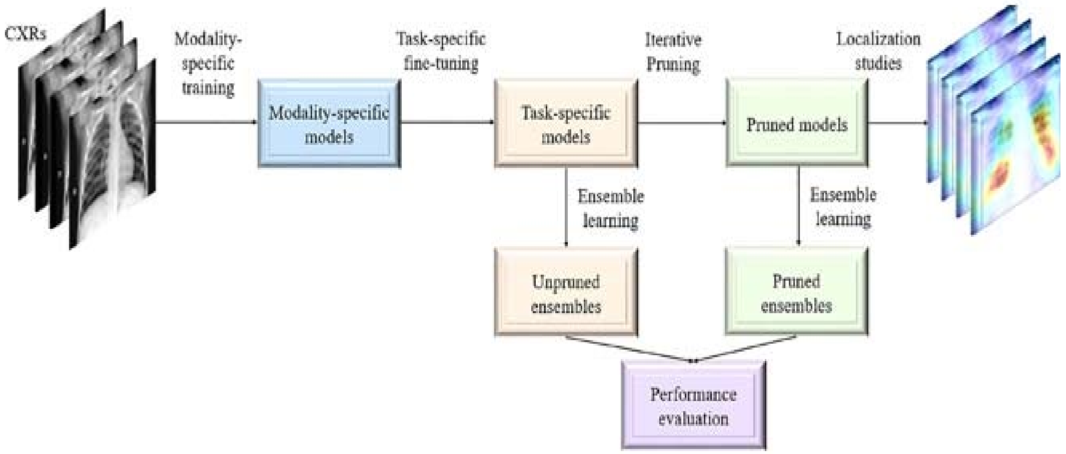
Study 3: Iteratively Pruned Deep Learning Ensembles for COVID-19 Detection in Chest X-Rays

a) Ensemble of iteratively-pruned, modality-specific CNNs to classify CXRs as normal or as showing bacterial pneumonia or COVID-19 pneumonia.

b) Iteratively pruned to remove filters with the highest average percentage of zeros.

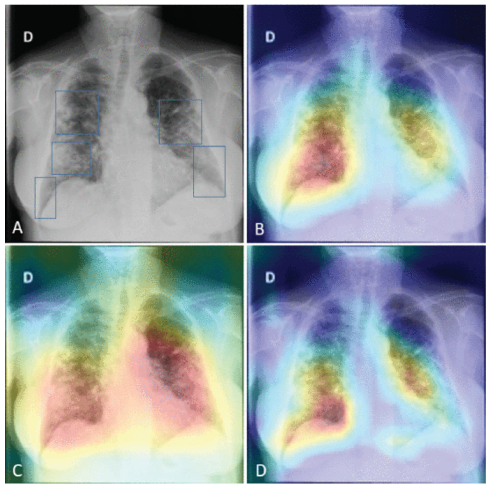
c) Ensemble methods: Majority voting, averaging, and weighted averaging

d) Weighted averaging resulted in superior classification and localization performance.



Method	Method	Acc.	AUC	Sens.	Prec.	F	MCC
Majority Voting	Unpruned	0.9742	0.9807 [0.9686 0.9928]	0.9742	0.9748	0.9742	0.9537
	Pruned	0.9821	0.9866 [0.9765 0.9967]	0.9821	0.9822	0.9821	0.9676
Averaging	Unpruned	0.9782	0.9969 [0.992 1.0]	0.9782	0.9786	0.9782	0.9607
	Pruned	0.9821	0.9969 [0.992 1.0]	0.9821	0.9823	0.9821	0.9677
Weighted Averaging	Unpruned	0.9762	0.9968 [0.9918 1.0]	0.9762	0.9767	0.9762	0.9572
	Pruned	**0.9901**	**0.9972 [0.9925 1.0]**	**0.9901**	**0.9901**	**0.9901**	**0.9820**
Stacking	Unpruned	0.9663	0.9865 [0.9764 0.9966]	0.9663	0.968	0.9662	0.9402
	Pruned	0.9712	0.9876 [0.9779 0.9973]	0.9712	0.9711	0.9712	0.9473

*Bold values stand for the model with a statistically significant better performance than the other models.



S. Rajaraman et al., "Iteratively Pruned Deep Learning Ensembles for COVID-19 Detection in Chest X-Rays," in *IEEE Access*, vol. 8, pp. 115041-115050, 2020, doi: 10.1109/ACCESS.2020.3003810.

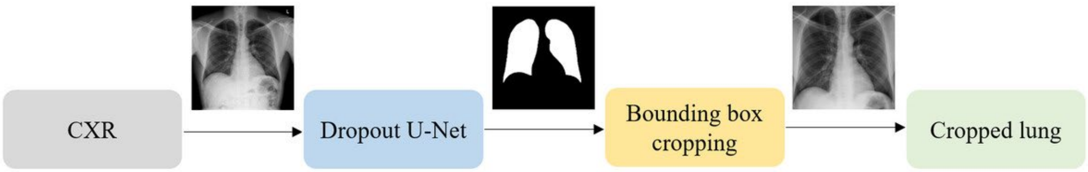
Study 4: Detection and visualization of abnormality in chest radiographs using modality-specific convolutional neural network ensembles

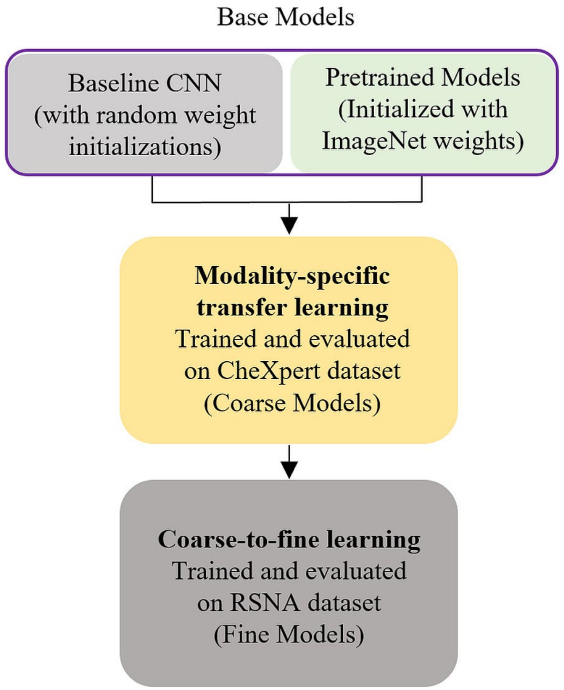
a) Modality-specific training of custom and ImageNet-pretrained models.

b) Modality-specific models fine-tuned to detect abnormalities in a smaller dataset.

c) Ensemble methods: Majority voting, simple averaging, weighted averaging, and stacking

d) Ensemble averaging resulted in superior classification and localization performance.

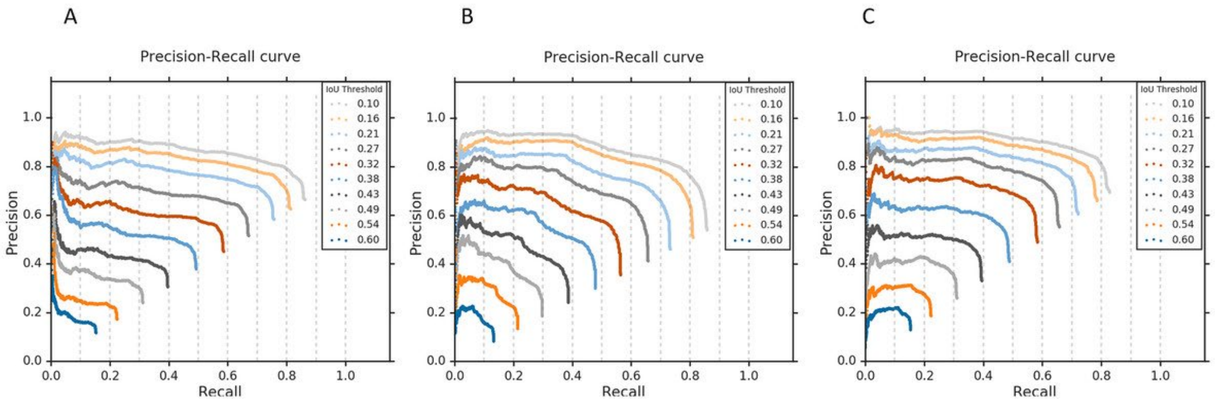




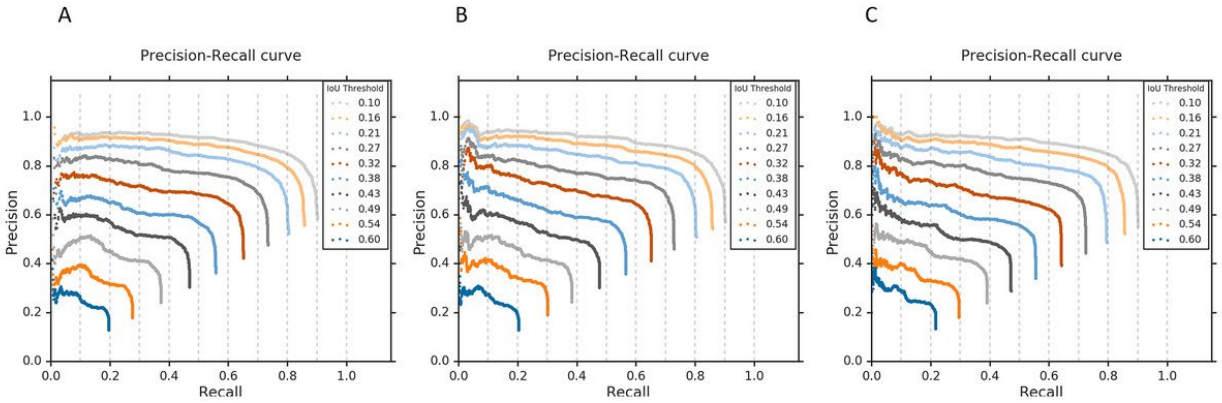
S. Rajaraman et al., Detection and visualization of abnormality in chest radiographs using modality-specific convolutional neural network ensembles. *PeerJ* 8:e8693 https://doi.org/10.7717/peerj.8693

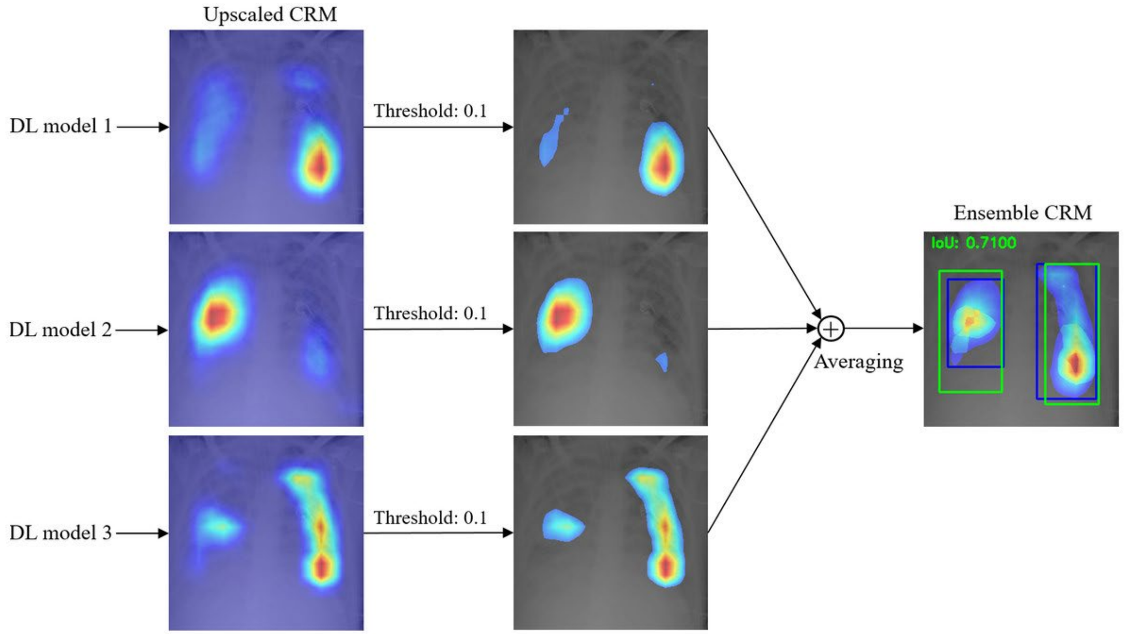
Kim, I.; Rajaraman, S.; Antani, S. Visual Interpretation of Convolutional Neural Network Predictions in Classifying Medical Image Modalities. Diagnostics 2019, 9, 38.

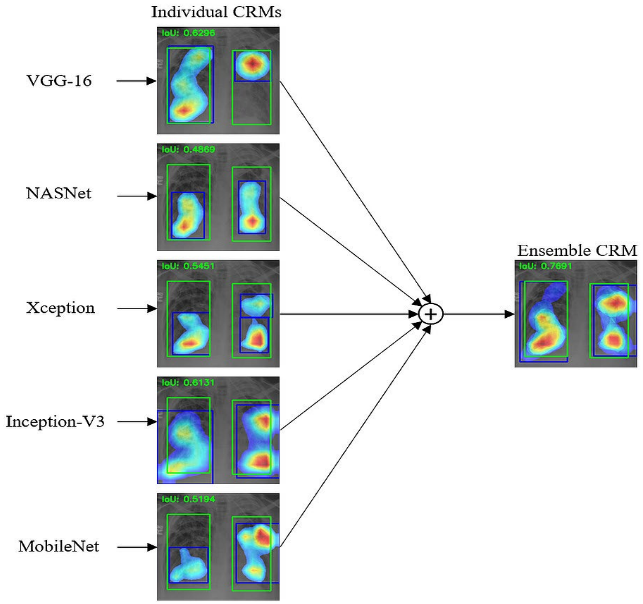
Metrics	VGG-16	VGG-19	Xception	Inception-V3	MobileNet	NASNet-mobile	DenseNet-121
IoU	0.383	0.357	0.377	0.351	0.368	0.375	0.355
mAP@[0.1 0.6]	0.377	0.341	0.388	0.348	0.352	0.382	0.317



Metrics	Ensemble-3	Ensemble-5	Ensemble-7
IoU	0.430	**0.433**	0.432
mAP@[0.1 0.6]	0.420	**0.447**	0.434







References

- S. Rajaraman, J. Siegelman, P. O. Alderson, L. S. Folio, L. R. Folio and S. K. Antani, "Iteratively Pruned Deep Learning Ensembles for COVID-19 Detection in Chest X-Rays," in IEEE Access, vol. 8, pp. 115041-115050, 2020, doi: 10.1109/ACCESS.2020.3003810.
- S. Rajaraman, S. Candemir, I. Kim, G. Thoma, S. Antani, "Visualization and Interpretation of Convolutional Neural Network Predictions in Detecting Pneumonia in Pediatric Chest Radiographs", Appl. Sci. 2018, 8, 1715.
- Rajaraman S, Kim I, Antani SK. Detection and visualization of abnormality in chest radiographs using modality-specific convolutional neural network ensembles. PeerJ 8:e8693 https://doi.org/10.7717/peerj.8693
- Rajaraman S, Antani SK. Modality-Specific Deep Learning Model Ensembles Toward Improving TB Detection in Chest Radiographs. IEEE Access, vol. 8, pp. 27318-27326, 2020.
- Rajaraman S, Candemir S, Xue Z, Alderson P, Thoma G, Antani SK. A Novel Stacked Model Ensemble for Improved TB Detection in Chest Radiographs. In Santosh KC et al. (Eds.). Medical Imaging: Artificial Intelligence, Image Recognition, and Machine Learning Techniques. (pp. 1-26). New York, NY: CRC Press, Taylor & Francis Group.
- Rajaraman S, Sornapudi S, Kohli M, Antani SK. Assessment of an ensemble of machine learning models toward abnormality detection in chest radiographs. Proc. IEEE Engineering in Medicine and Biology Conference (EMBC), Berlin, Germany, 23 – 27 July 2019.

https://github.com/sivaramakrishnan-rajaraman

rajaramans2@mail.nih.gov

Thank you!