Predicting COVID-19 Severity on Chest X-ray with Deep Learning

Presented by: Chandresh Kumar Maurya
CSE, IIT Indore

Outline of the Talk

1. Introduction
2. Motivation
3. The problem
4. Solution Approach
5. Results
6. Discussion
7. Future direction

Introduction

- Recent corona outbreak changed the way we live and work.
- Due to non-availability of proper diagnostic tools in place, it is quite difficult to distinguish the symptoms of sars-covid-19 virus vs symptoms of other diseases such as pneomonia.
- Many studies have tried to predict the severity of sars-covid-19 infection through chest X-ray (CXR).
- We will look at studies using deep learning to predict sars-covid-19 severity.

Studies done Till date

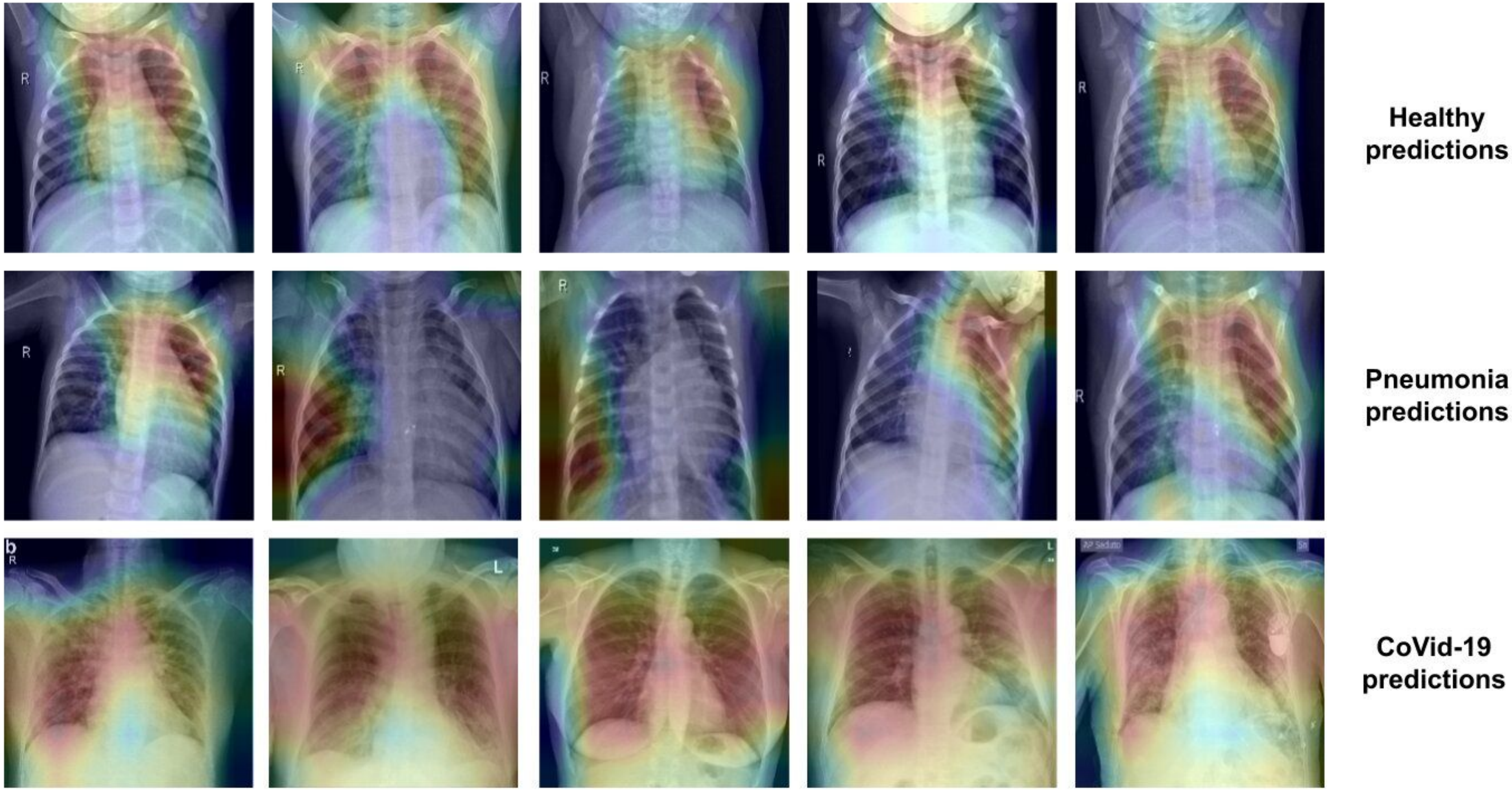
1. Shoeibi, Afshin, et al. "Automated detection and forecasting of covid-19 using deep learning techniques: A review." *arXiv preprint arXiv:2007.10785* (2020).
2.

Detection of Covid-19

- The detection of COVID-19 is significant and vital in its early stages.
- Various methods have been proposed to diagnose COVID-19, containing a variety of medical imaging techniques, blood tests (CBCs), and PCR.
- According to the WHO, all diagnoses of corona disease must be confirmed by reverse-transcription polymerase chain reaction (RT-PCR).
- However, testing with RT-PCR is highly time-consuming, and this issue is risky for people with COVID-19.
- Hence, first, medical imaging is carried out for the primary detection of COVID-19, then the RT-PCR test is performed to aid the physicians in making final accurate detection.
- Two medical imaging techniques, X-ray and CT-scan, are employed to diagnose COVID-19

Imaging Techniques: CXR

- X-ray modality is the first procedure to diagnose COVID19, which has the advantage of being inexpensive and low-risk from radiation hazards to human health.
- In the X-ray method, detecting COVID-19 is a relatively complicated task.
- In these images, the radiologist must attentively recognize the white spots that contain water and pus, which is very prolonged and problematic.
- A radiologist or specialist doctor may also mistakenly diagnose other diseases, such as pulmonary tuberculosis, as COVID-19 [1].

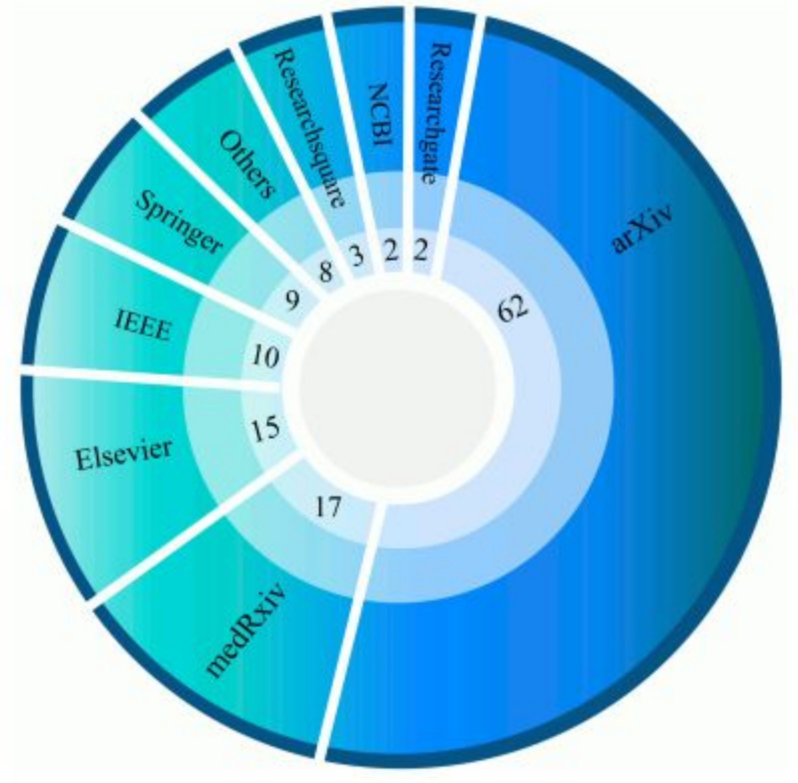


Source: https://medicalxpress.com/news/2020-05-distinguish-pneumonia-covid-chest-x-rays.html

CT Scan

- The X-ray procedure has a high error rate; hence CT images can be used for more precise detection.
- Nevertheless, these CT images are far more expensive than X-rays for patients.
- At the time of CT-scan recording, several slices are provided from each person suspected of COVID-19.
- The large volume of CT-Scan images calls for a high workload on physicians and radiologists to diagnose COVID-19.

Number of papers published on COVID-19 using DL techniques.



Computer Aided Diagnostic System (CADS)

CADS for covid

- classification
  - CNN
  - RNN
  - CNN-RNN
  - Deep-belief network
- Segmentation
  - U-net
  - mask-RCNN

Block diagram for COVID-19 detection using DL technique.

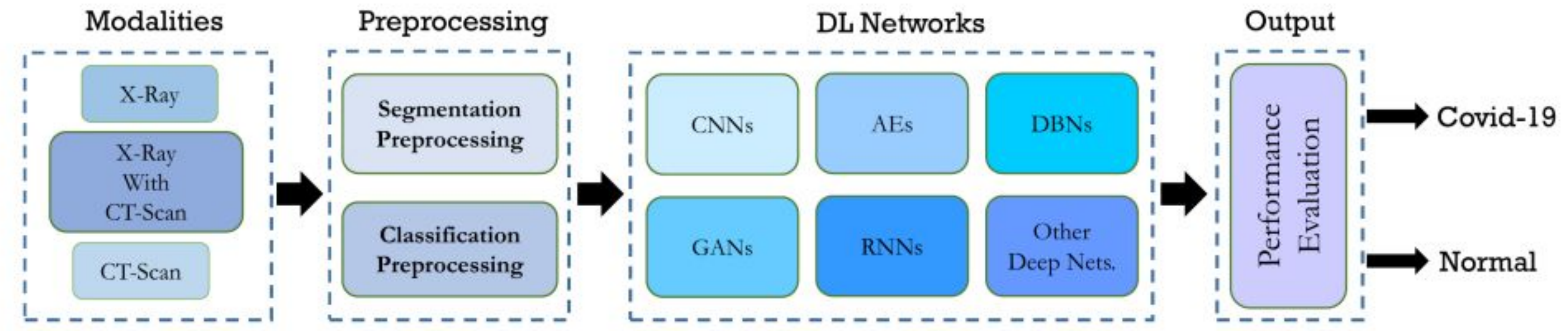


Illustration of various DL methods used for COVID-19 detection.

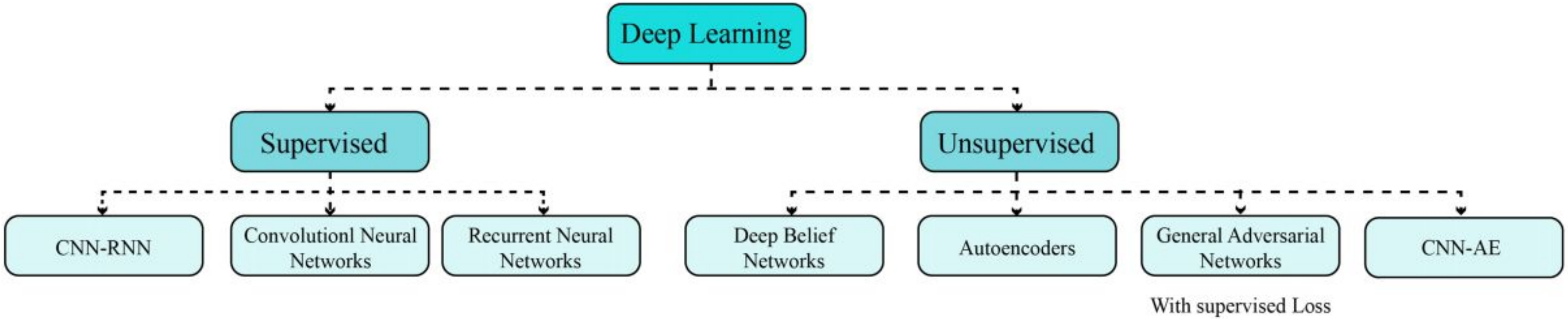


TABLE I: Public databases used for COVID-19 detection.

Dataset	Modality	Link
J. P. Cohens GitHub [29]	X-ray and CT	https://github.com/ieee8023/covid-chestxray-dataset
European Society of Radiology	X-ray and CT	https://www.eurorad.org/advanced-search?search=COVID
SIRM	X-ray and CT	https://www.sirm.org/category/senza-categoria/covid-19
BSTI	X-ray and CT	https://www.bsti.org.uk/covid-19-resources
UCSD-AI4H [30]	CT	https://github.com/UCSD-AI4H/COVID-CT
MedSeg	CT	http://medicalsegmentation.com/covid19
Kaggle	X-ray and CT	https://www.kaggle.com/datasets?search=covid
Point-of-Care Ultrasound (POCUS) [31]	Lung Ultrasound Images and Videos	https://github.com/jannisborn/covid19_pocus_ultrasound
Actualmed COVID-19 Chest X-ray Dataset Initiative	X-ray	https://github.com/agchung/Actualmed-COVID-chestxray-dataset
COVID-19 Chest X-ray Dataset Initiative	X-ray	https://github.com/agchung/Figure1-COVID-chestxray-dataset
Georgia State Universitys Panacea Lab [32]	Twitter Chatter Dataset	https://github.com/thepanacealab/covid19_twitter
Twitter COVID19 CXR dataset	X-ray	https://twitter.com/ChestImaging
COVID-19 [33]	CT	https://github.com/KevinHuRunWen/COVID-19
COVIDx [34]	X-ray	https://github.com/lindawangg/COVID-Net

Predicting Covid-19 Severity using CXR images

- **The data:**
  - 94 images of CXR
  - All patients were reported COVID-19 positive and sourced from many hospitals around the world from December 2019 to March 2020.
  - The images were de-identified prior to our use and there was no missing data.
  - The ratio between male/female was 44/36 with an average age of 56±14.8 (55±15.6 for male and 57±13.9 for female).

Predicting Covid-19 Severity using CXR images

- **The Label:**
  - Radiological scoring was done using 3 blinded experts
  - Staged severity based on the extent of lung involvement and lung opacity.
  - The extent of lung involvement by ground glass opacity or consolidation for each lung (right lung and left lung separately) was scored as: 0 = no involvement; 1 = <25% involvement; 2 = 25-50% involvement; 3 = 50- 75% involvement; 4 = >75% involvement. The total extent score ranged from 0 to 8 (right lung and left lung together).
  - The degree of opacity for each lung (right lung and left lung separately) was scored as: 0 = no opacity; 1 = ground glass opacity; 2 = consolidation; 3 = white-out. The total opacity score ranged from 0 to 6 (right lung and left lung together).

Non-Covid Pre-training Data

- RSNA Pneumonia Challenge [Shih et al., 2019].
- CheXpert - Stanford University[Irvin et al., 2019].
- ChestX-ray8 - National Institutes of Health (NIH) [Wang et al., 2017],
- ChestX-ray8 - NIH with labels from Google [Majkowska et al., 2019].
- MIMIC-CXR - MIT [Johnson et al., 2019].
- PadChest - University of Alicante [Bustos et al., 2019].
- OpenI [Demner-Fushman et al., 2016]

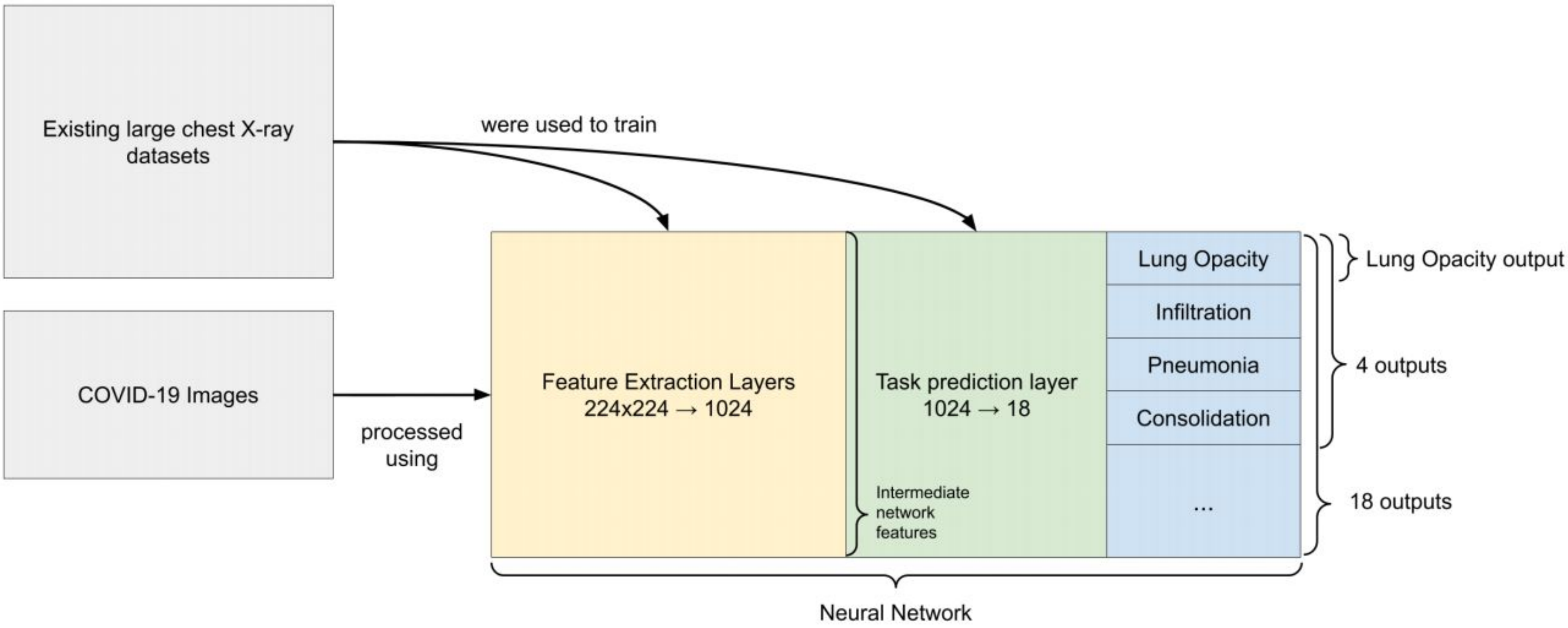
Radiological Tasks

- atelectasis, consolidation, infiltration, pneumothorax, edema, emphysema, fibrosis, fibrosis, effusion, pneumonia, pleural thickening, cardiomegaly, nodule, mass, hernia, lung lesion, fracture, lung opacity, and enlarged cardiomediastinum.
- In total, 88,079 non-COVID-19 images were used to train the model on these tasks.

Model, Preprocessing, and Pre-Training

1. **The model:** DenseNet [Huang et al., 2017].
2. **The preprocessing:** Images were resized to 224 224 pixels, utilizing a center crop if the aspect ratio was uneven, and the pixel values were scaled to [-1024, 1024] for the training.
3. A pretraining step was performed using the seven datasets to train feature extraction layers and a task prediction layer.

Training on Covid-19 images



Severity Prediction

- Linear regression was performed to predict the aforementioned scores (extent of lung involvement and opacity) using these different sets of features in place of the image itself:
    - Intermediate network features - the result of the convolutional layers applied to the image resulting in a 1024 dimensional vector which is passed to the task prediction layer;
    - 18 outputs - each image was represented by the 18 outputs (pre-sigmoid) from the pre-trained model;

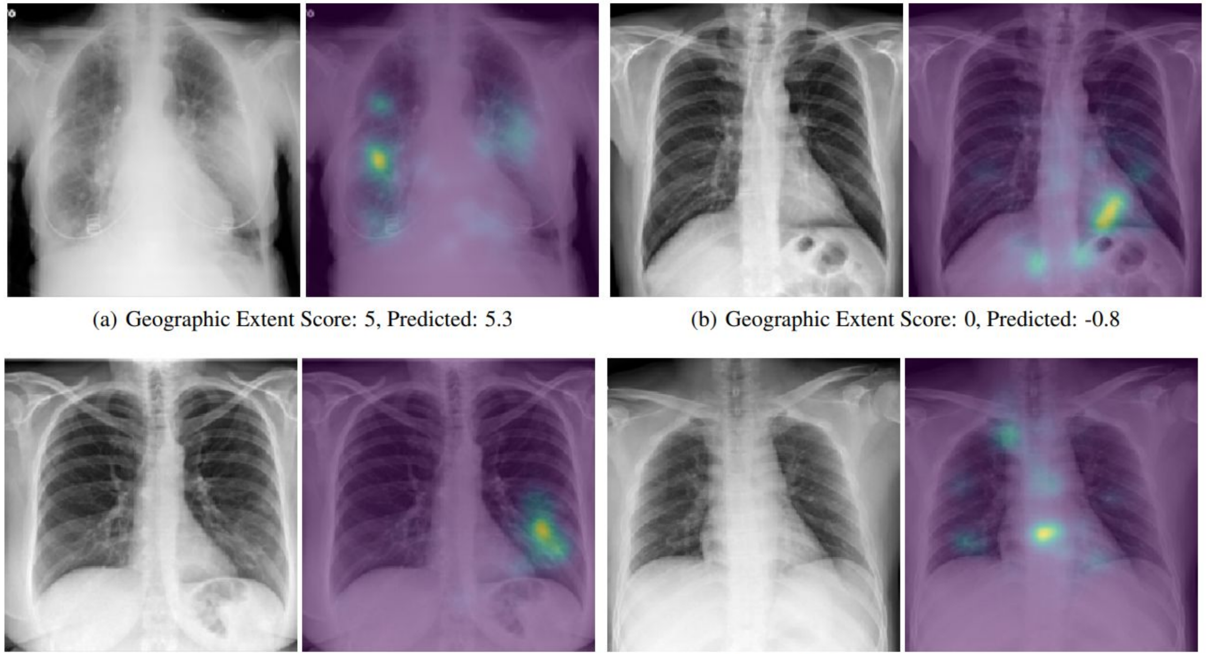
Severity Prediction

    - 4 outputs - a hand picked subset of outputs (presigmoid) were used containing radiological findings more frequent in pneumonia (lung opacity, pneumonia, infiltration, and consolidation);
    - Lung opacity output - the single output (pre-sigmoid) for lung opacity was used because it was task related. This is different from the predicted opacity score.
- MSE was used as a loss function for training the model on 50/50 split.
- **Saliency Map:**
    - In order to ensure that the models are looking at reasonable aspects of the images, a saliency map is computed by computing the gradient of the output prediction with respect to the input image (if a pixel is changed how much will it change the prediction).
    - In order to smooth out the saliency map, it is blurred using a 5x5 Gaussian kernel.

Results

Task	Using features:	# parameters (fewer is better)	Pearson Correlation	R^2	MAE	MSE
Opacity Score	”lung opacity” output	1+1	**0.78±0.04**	**0.58±0.09**	**0.78±0.05**	**0.86±0.11**
	4 outputs	4+1	0.78±0.04	0.58±0.09	0.76±0.05	0.87±0.12
	18 outputs	18+1	0.73±0.09	0.44±0.16	0.86±0.11	1.15±0.33
	Intermediate network features	1024+1	0.66±0.08	0.25±0.21	1.01±0.09	1.54±0.28
	No data	0+1	0.00±0.00	-0.08±0.10	1.24±0.10	2.26±0.36
Geographic Extent	”lung opacity” output	1+1	**0.80±0.05**	**0.60±0.09**	**1.14±0.11**	**2.06±0.34**
	4 outputs	4+1	0.79±0.05	0.57±0.10	1.19±0.11	2.17±0.37
	18 outputs	18+1	0.76±0.08	0.47±0.16	1.32±0.17	2.73±0.89
	Intermediate network features	1024+1	0.74±0.08	0.43±0.16	1.36±0.13	2.88±0.58
	No data	0+1	0.00±0.00	-0.08±0.10	2.00±0.17	5.60±0.95

Results



(a) Geographic Extent Score: 5, Predicted: 5.3

(b) Geographic Extent Score: 0, Predicted: -0.8

(c) Geographic Extent Score: 2, Predicted: 0.62

(d) Geographic Extent Score: 0, Predicted: 1.05

Figure 4. Examples of correct (a,b) and incorrect (c,d) predictions by the model are shown with a saliency map generated by computing the gradient of the output prediction with respect to the input image and then blurred using a 5x5 Gaussian kernel. The assigned and predicted scores for Geographic Extent are shown to the right.

Challenges

- Data scarcity
- DL network architecture
- Hardware resources

References

1. Orioli, Laura, et al. "COVID-19 in diabetic patients: related risks and specifics of management." *Annales D'endocrinologie*. Elsevier Masson, 2020.