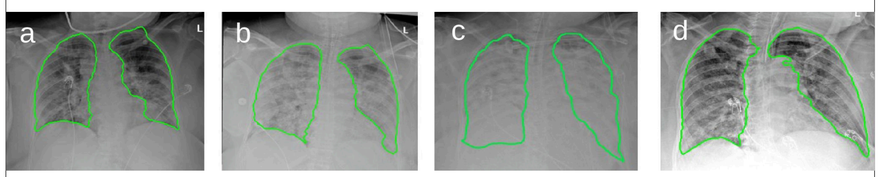
Predicting COVID-19 Lung Infiltrate Progression on Chest Radiographs Using Spatio-temporal LSTM based Encoder-Decoder Network

Aishik Konwer[1], Joseph Bae[2], Gagandeep Singh[3], Rishabh Gattu[3], Syed Ali[3], Jeremy Green[3], Tej Phatak[3], Amit Gupta[4], Chao Chen[2], Joel Saltz[2], Prateek Prasanna[2]

[1]Department of Computer Science, Stony Brook University
[2]Department of Biomedical Informatics, Stony Brook University
[3]Department of Radiology, Newark Beth Israel Medical Center
[4]University Hospitals Cleveland Medical Center

Clinical motivation

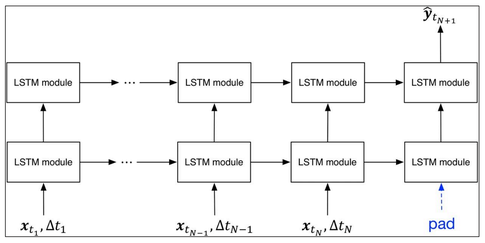


Chest radiographs of the chest (a-d) from a single patient over four days
(a) Ground glass opacities
(b) Slightly increased opacities.
(c) Extensive confluent consolidations.
(d) Similar findings as seen on day 1

- Currently no imaging models can predict the severity of disease at a later time point, based on the trajectory in the first few days of treatment.
- Radiographic analysis on sequential CXR can provide insights on treatment response.

Technical Motivation

- Existing ML based COVID-19 studies primarily utilize single timepoint radiographic scans for outcome prediction.
  -CNN features
  -Radiomic features
- LSTMs have been used in disease progression modelling to capture evolution of severity using sequential medical imaging data (ophthalmic disease [1], lung cancer)



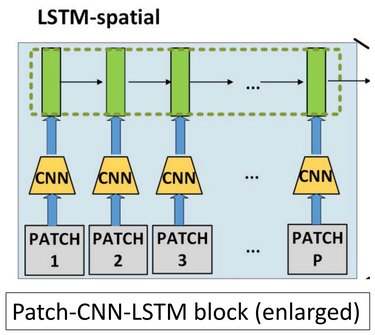
Example of a typical LSTM approach [2]

GOAL: To predict chest x ray severity scores at a future timpeoint

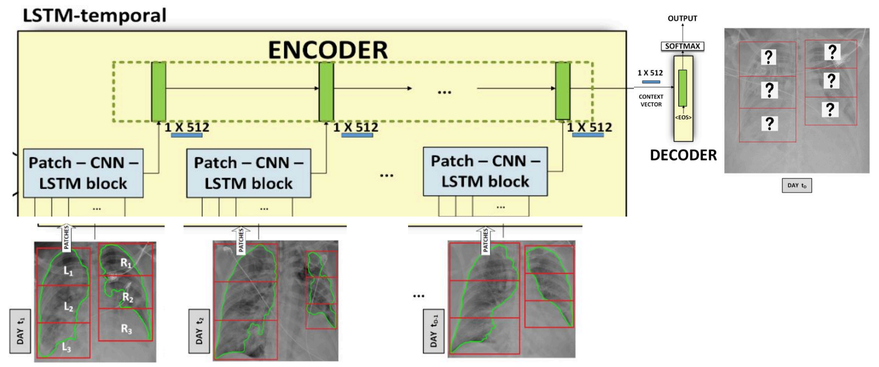
Our encoder framework learns in 2 stages:

- **LSTM-Spatial** aggregates spatial information from different lung locations
- **LSTM-Temporal** aggregate information from CXRs at different timepoints

Methodology



Patch-CNN-LSTM block (enlarged)



Step 1: 256 x 256 patches extracted from each lung zone.
- Oversampling handles severity grade imbalance.
- Neighbor patches enhance contextual information.

Step 2: CNN features from each patch fed as input to each timestep of ***LSTM-Spatial***

Step 3: 512 dimension global feature vector from LSTM-Spatial input to each timestep of ***LSTM-Temporal***, which eventually generates the context vector.

Step 4: Context vector and EOS are inputs to first timestep of decoder. Softmax applied to classify its output into severity scores.

Results

Methods	Left Lung Upper							Left Lung Middle							Left Lung Lower						
	Acc(%)	Pre			Rec			Acc(%)	Pre			Rec			Acc(%)	Pre			Rec		
		0	1	2	0	1	2		0	1	2	0	1	2		0	1	2	0	1	2
Baseline-1	60 ±4.76	0.55	0.66	0.52	0.51	0.68	0.52	64 ±5.47	0.5	0.71	0.59	0.47	0.74	0.56	58 ±4.89	0.5	0.62	0.56	0.47	0.71	0.48
Baseline-2	57 ±4.63	0.48	0.67	0.48	0.41	0.64	0.60	61 ±5.80	0.48	0.7	0.56	0.52	0.64	0.60	55 ±4.64	0.5	0.60	0.50	0.52	0.70	0.43
Variant-1	66 ±4.33	0.57	0.67	0.63	0.56	0.73	0.48	69 ±4.68	0.45	0.72	0.51	0.65	0.74	0.56	64 ±4.87	0.39	0.66	**0.74**	0.45	0.74	0.70
Variant-2	68 ±3.51	0.53	0.74	0.41	**0.67**	**0.78**	0.63	70 ±2.89	0.43	0.68	0.55	0.47	0.69	**0.77**	61 ±4.16	0.35	0.57	0.64	0.48	**0.81**	0.64
Our Approach	**71** ±3.58	**0.69**	**0.75**	**0.64**	0.62	0.77	**0.69**	**73** ±2.56	**0.72**	**0.77**	**0.6**	**0.69**	**0.83**	0.52	**69** ±3.94	**0.61**	**0.73**	0.67	0.52	0.73	**0.72**

Methods	Right Lung Upper							Right Lung Middle							Right Lung Lower						
	Acc(%)	Pre			Rec			Acc(%)	Pre			Rec			Acc(%)	Pre			Rec		
		0	1	2	0	1	2		0	1	2	0	1	2		0	1	2	0	1	2
Baseline-1	64 ±3.23	0.56	0.72	0.52	0.6	0.71	0.5	55 ±4.08	0.45	0.60	0.51	0.40	0.63	0.51	58 ±3.91	0.54	0.62	0.54	0.42	0.59	0.61
Baseline-2	67 ±3.38	0.63	0.72	0.6	0.7	0.65	0.66	52 ±3.29	0.47	0.60	0.39	0.45	0.63	0.37	56 ±4.36	0.4	0.65	0.51	0.42	0.63	0.51
Variant-1	72 ±3.09	0.65	0.69	0.52	0.67	0.73	0.62	66 ±2.62	0.42	0.56	**0.72**	0.56	0.62	**0.80**	63 ±3.85	0.50	0.62	0.57	**0.60**	0.61	0.52
Variant-2	70 ±2.81	**0.84**	0.62	**0.67**	0.70	0.64	0.57	62 ±1.76	0.59	**0.78**	0.63	0.57	0.71	0.45	64 ±3.27	**0.77**	0.43	0.66	0.40	**0.66**	0.66
Our Approach	**76** ±2.33	0.68	**0.84**	0.66	**0.73**	**0.80**	**0.66**	**67** ±2.72	**0.59**	0.71	0.65	**0.59**	**0.75**	0.58	**65** ±3.73	0.5	**0.67**	**0.67**	0.5	0.65	**0.69**

Quantitative results for the six Left and Right lung zones

Dataset: 657 temporal AP CXRs from 100 COVID-19 patients, where the duration between the CXRs are variable.Severity scores (0,1,2) assigned by radiologists to each lung zone.
Our approach **outperforms** both variants and state-of-the-art baselines in all six lung zones.

Baseline 1: Fine-tuned VGG16 [3] features
Baseline 2: Radiomic approach
Variant 1: LSTM-Spatial was removed.
Variant 2: Oversampling and Neighbor patch removed.

Conclusion

1) A novel multi-stage LSTM framework that learns both spatial and temporal information from sequential CXRs.
2) Our model can potentially inform duration and timing of clinical treatments (e.g. proning).

References:

1) Jiang et al.:Predicting the progression of ophthalmic disease based on slit-lamp images using a deep temporal sequence network. *PloS one* 13.7 (2018): e0201142.
2) Wang et al.: Predictive modeling of the progression of Alzheimer's disease with recurrent neural networks. *Scientific reports* 8.1 (2018): 1-12.
3) Simonyan et al.: Very deep convolutional networks for large-scale image recognition.