UNIVERSITY OF COPENHAGEN

Lung Segmentation from Chest X-rays using Variational Data Imputation

Presented at ICML Workshop on Learning from Missing Data (Artemiss 2020)

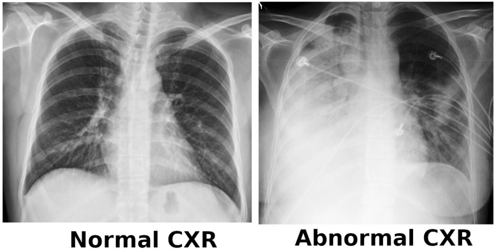
Raghavendra Selvan, Erik B. Dam, Nicki S. Detlefsen,
Sofus Rischel, Kaining Sheng, Mads Nielsen, Akshay Pai

Dept. of Computer Science,
University of Copenhagen

raghav@di.ku.dk
@raghavian

Source Code & Models: `https://github.com/raghavian/lungVAE`

COVID-19 like diseases obfuscate lungs in chest X-rays



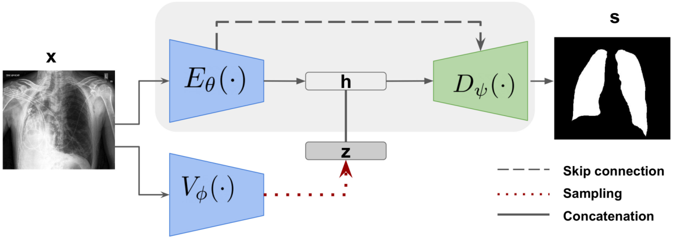
Contribution

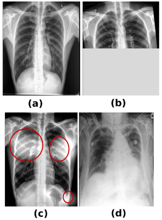
Automatic segmentation of lungs in the presence of pulmonary opacifications by posing it as a *missing data* problem

---

Caption: Normal CXR shows lungs clearly whereas abnormal CXR has high opacity where the right lung is hardly seen. Brighter regions are tissue-like as they attenuate X-rays whereas darker regions indicate presence of air, in this case inside the lungs.

Proposed segmentation model uses variational data imputation





Extensive augmentation

During training extensive data augmentation is used to simulate lung opacifications.

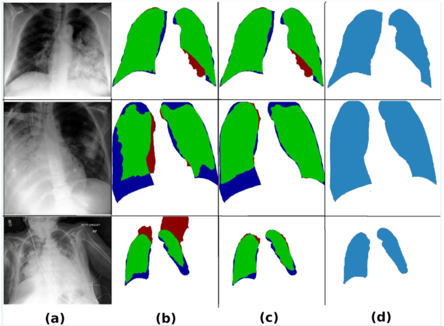
Captions: (left) The proposed model with variational encoder for data imputation, $V_\phi(\cdot)$, U-net type segmentation network with encoder $E_\theta(\cdot)$, decoder $D_\psi(\cdot)$.
(right) Chest X-rays with and without augmentation. (a) No augmentation (b) With block masking (c) With diffused noise marked with red ellipses (d) Test image with high opacity

Experiments & Results

- Public CXR datasets (Shenzhen and Montgomery hospitals)
- Training: 528 CXRs, Validation: 176 CXRs
- Test set. 30 CXRs from diverse datasets with manual annotations (in-house experts)
- U-net model for baseline comparison
- Ablation of different augmentation strategies

Table 1. Performance measures on the test set

Models	Augmentation	Dice Overlap	Accuracy
Baseline	Standard	0.7335 ± 0.17	0.8449 ± 0.09
Proposed	Standard	0.7204 ± 0.18	0.8392 ± 0.10
Baseline	Block	0.7563 ± 0.15	0.8522 ± 0.09
Proposed	Block	0.7688 ± 0.17	0.8552 ± 0.10
Baseline	Diffuse	0.7757 ± 0.15	0.8654 ± 0.10
Proposed	Diffuse	0.7965 ± 0.11	0.8652 ± 0.11
Baseline	Block+Diffuse	0.8173 ± 0.12	0.8654 ± 0.11
Proposed	Block+Diffuse	$\mathbf{0.8503 \pm 0.07}$	$\mathbf{0.8815 \pm 0.11}$



a) Three test set samples with highest and least dice accuracy for both methods (rows 1 & 2) along with an input CXR with additional variations in pose (row-3). b) baseline model predictions, c) proposed model predictions and d) the ground truth. Both predictions are for models trained with block and diffused noise.

Green:True positive, Blue: False Negative, Red: False Positive