CheXseg: Combining Expert Annotations with DNN-generated Saliency Maps for X-ray segmentation

Soham Gadgil, Mark Endo, Emily Wen, Andrew Ng, Pranav Rajpurkar
{sgadgil, markendo, emilywen, ang, pranavsr}@cs.stanford.edu

MIDL (Medical Imaging with Deep Learning) 2021 Poster

Introduction

- The black-box nature of neural networks represents a barrier to physicians' trust and model adoption in a clinical setting
- Saliency Maps are a popular set of explanation methods but they are untrustworthy for for medical image interpretation
- Segmentation models can produce accurate pixel-level maps, but expert annotations are limited
- We introduce CheXseg, a semi-supervised method that leverages both expert annotations and saliency maps
- CheXseg reduces the overall gap to radiologist localization performance (mIoU) by 57.2% compared to solely using DNN-generated saliency maps.

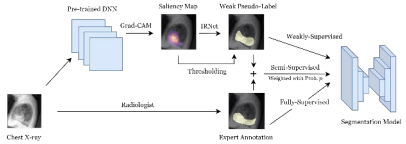
Setup

- Multi-label semantic segmentation task
- Classify each pixel of a chest X-ray image into 0 or more of 10 possible pathologies

Dataset

- We utilize CheXpert, a large dataset with ~220k chest x-rays of ~65k patients
- Validation set of 200 + Test set of 500 radiologist-annotated chest X-rays

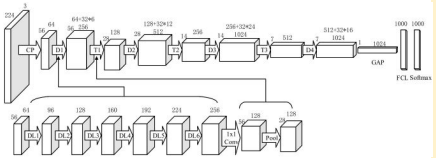
Workflow



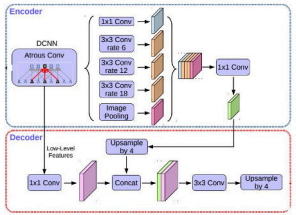
- We compare and contrast semi-, fully-, and weakly-supervised methods

Methods

- Grad-CAM is used to obtain saliency maps from the chest x-rays
- DenseNet-121 is used as the pre-trained classification model

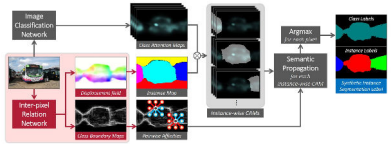


- DeepLabv3+ is used as the core semantic segmentation model



Inter-Pixel Relation Network (IRNet)

- IRNet is used to obtain per-pixel segmentation masks called weak pseudo-labels from the saliency maps
- It improves CAMs by training two output branches - a displacement vector field and a class boundary map



Notable References

For full list please checkout https://arxiv.org/pdf/2102.10484.pdf

1. Irvin, et al. Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison. In Proceedings of the AAAI Conference on Artificial Intelligence, volume 33, pages 590-597, 2019.
2. Ahn et al. Weakly supervised learning of instance segmentation with inter-pixel relations. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2209-2218, 2019.
3. Selvaraju et al. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE international conference on computer vision, pages 618-626, 2017.
4. Chen et al. Encoder-decoder with atrous separable convolution for semantic image segmentation. In Proceedings of the European conference on computer vision (ECCV), pages 801-818, 2018.

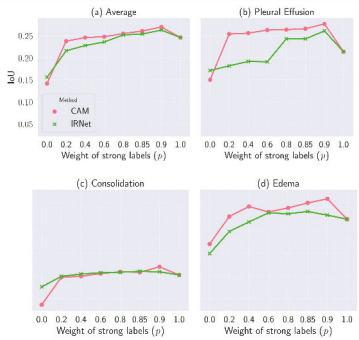
CheXseg: Combining Expert Annotations with DNN-generated Saliency Maps for X-ray segmentation

Soham Gadgil, Mark Endo, Emily Wen, Andrew Ng, Pranav Rajpurkar
{sgadgil, markendo, emilywen, ang, pranavsr}@cs.stanford.edu

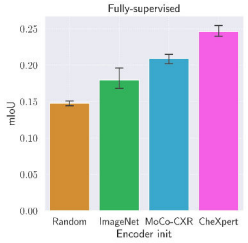
MIDL (Medical Imaging with Deep Learning) 2021 Poster

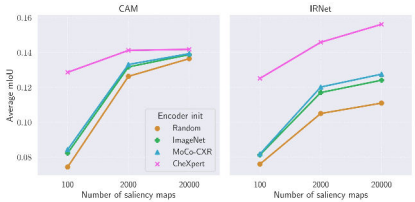
Experiments

- Combining DNN-generated saliency maps and expert annotations
- We vary the probability ρ of selecting an expert annotation in a single batch
- ρ = 0.9 (CheXseg) gives the best mIoU performance of 0.27
- CheXseg beats both the weakly supervised case (mIoU of 0.156) and the fully supervised case (mIoU of 0.246)



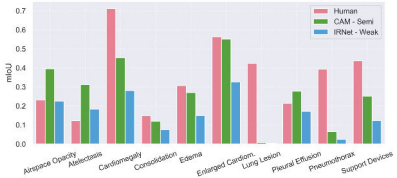
- We investigate the impact of using various encoder initializations on segmentation performance
- For all methods, the best models are initialized with CheXpert encoder weights





Comparison to Radiologists

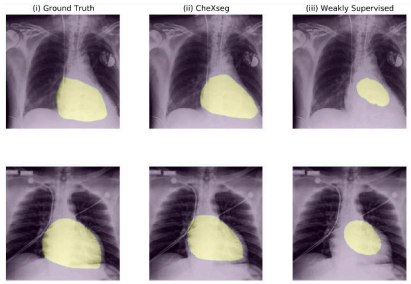
- CheXseg reduces the overall performance (mIoU) gap with radiologists by 57.2% compared to the best weakly-supervised method
- CheXseg outperforms radiologists on Atelectasis, Airspace Opacity, and Pleural Effusion



Qualitative Results

- We show the qualitative results for Cardiomegaly and Airspace Opacity
- CheXseg gives better visualizations as compared to the best weakly-supervised method

Cardiomegaly



Airspace Opacity