Towards Robust and Adaptable Diagnosis of Pneumonia from Chest X-ray Data

Raphaël Attias (287613), Riccardo Cadei (321957)
CS-503 Progress Report II

I. Milestone Progress

A main assumption in Statistical Learning is that the data are all independent and identically distributed (*i.i.d.*) and then, through the Law of Big Number we can approximate the Expected Risk:

$$\mathcal{R}(f) := \mathop{\mathbb{E}}_{(\boldsymbol{x},y)}\left[\ell(f(\boldsymbol{x}), y)\right] \tag{1}$$

with the Empirical Risk:

$$\mathcal{R}(f) := \frac{1}{|\mathcal{D}|}\sum_{(\boldsymbol{x},y)\in\mathcal{D}} \ell(f(\boldsymbol{x}), y) \tag{2}$$

and minimize it (ERM). Unfortunately in real world applications this assumption doesn't hold and the data is generally split in different environments following different distributions (i.e. ray-x images produced by different machines). First we reproduced the Experiment 1 proposed by Bellot et al. [1] using both a simple CNN and CheXNet [2], a SOTA Deep Convolutional Neural Network based on DenseNet for Pneumonia detection. In Table I we summarize the percentage of accuracy of both the models using or not a robust training routine (IRM [3]).

Model	Method	Train (Env 1)	Train (Env 2)	Test
Baseline	ERM	87.564	88.940	61.154
	IRM ($\lambda = 0.1$)	63.512	64.249	61.346
	IRM ($\lambda = 3$)	49.794	49.022	43.307
CheXNet	ERM	77.492	78.909	71.346
	IRM ($\lambda = 0.1$)	72.914	74.845	83.269
	IRM ($\lambda = 3$)	59.9691	62.191	74.807

Table I

Our results are consistent with the results showed in the original study: ERM fails in presence of spurious features and overfit on shortcuts in the training set. Using a more robust training routine (i.e. IRM) the gap among training and OoD Test is reduced but our hypothesis is that a robust training routine suppresses not only the spurious features, but also the confounders (i.e. style) which are also useful for prediction.

Then, we have designed a whole pipeline (architecture, training method, training routine, transfer protocol) to train a robust and (also) adaptable model, we called RoAdaNetNet, for diagnosis of Pneumonia.

First, an auto-encoder (i.e. U-Net [4]) should be trained in a robust (i.e. IRM) and self-supervised way, getting an encoder Φ able to produce a robust latent representation of the input. No labels are needed for this part and it could be run a priori with a lot of images even from different environments. A second encoder Φ_{style} should be designed to encode the style of the environments and it could be also trained in a unsupervised way using a contrastive loss separating (in the latent space) the images

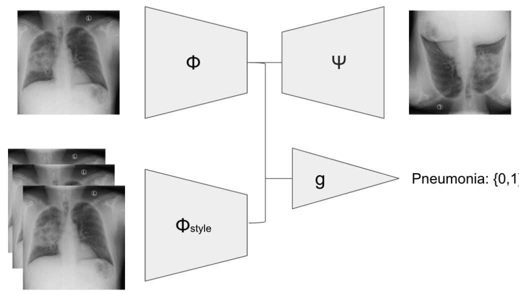

Figure 1.

from different environments. The output of both the encoder should be concatenated and given in input to a simple final classifier g trained a posteriori in a standard (ERM) supervised way. In Figure 1 a simple diagram representation of the whole network. The main advantage of such a modular architecture is that its (robust) back-bone can be trained a priori with heterogenous data not labelled.

II. Discussion & Next Steps

The main challenge in this project are the limited resources. Even if the dataset is medium-size (thousands of medium quality images), the information extraction is not as simple (even for binary classification), the models are very big (6'954'881 parameters ...) and the training last from 6h up to a day even using modern GPUs (we tried both Azure, Google Cloud and our personal GPUs). This constraint avoid us to properly finetune all the hyper-parameters and we are afraid we will not be able to properly pretrain the auto-encoder in RoAdaNet with enough data. The main task for the final deadline is to implement and train RoAdaNet.

III. Author Contribution Statement

Raphaël took care of setup different clusters (Azure, Google Cloud, Google Colab) to get modern GPUs in add to our personal machines. Both Raphaël and Riccardo run the experiments to produce Table 1. Riccardo studied the literature and proposed the design of RoAdaNet. Both Raphaël and Riccardo started the implementation of RoAdaNet.

References

[1] A. Bellot and M. van der Schaar, "Accounting for unobserved confounding in domain generalization," *arXiv preprint arXiv:2007.10653*, 2020.

[2] P. Rajpurkar and J. Irvin, "Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning," *CoRR*, vol. abs/1711.05225, 2017. [Online]. Available: http://arxiv.org/abs/1711.05225

[3] M. Arjovsky, L. Bottou, I. Gulrajani, and D. Lopez-Paz, "Invariant risk minimization," *arXiv preprint arXiv:1907.02893*, 2019.

[4] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.