

Bern, 15/04/2021

[Bern Data Science Day 2021 - April 23](#)

Contribution - Machine Learning and Medical Images

Semi-supervised Learning, Active Learning.

Fei WU

University of Bern, ARTORG, PhD Student, fei.wu@artorg.unibe.ch

Raphael Sznitman

University of Bern, ARTORG, Prof. Dr., raphael.sznitman@artorg.unibe.ch

Pablo Marquez Neila

University of Bern, ARTORG, Dr., pablo.marquez@artorg.unibe.ch

Abstract

In recent years, the use of machine learning has greatly improved several computer imaging tasks in the area of medical image analysis. This area, normally reserved for physicians with great expertise, has expanded and now welcome the contribution of a greater number of researchers thanks to computer vision. The limiting factor however to this computer approach is the number of labeled images. Labelling an image is the kind of task that artificial intelligence can not solve and still requires the attention of domain experts. We can however build models that are efficient with a few training samples or build methods that can identify the most important unlabeled data in order to assist experts in the annotation process. The latter is the area of active learning and it points out the single images that are most representative of the overall data distribution. We then proceed to label a small pool of images with the highest quality in order to mitigate the lack of quantity. There exist several methods to assess the quality of an image with the entropy being the most well known among them. In this work, we investigate the use of intrinsic elements of neural networks in order to understand which samples are most useful in training models with a limited number of samples. With applications in healthcare a focus as annotations are costly, we show results on different approaches related to the Neural Tangent Kernel on image classification tasks.

