Project RAF-MedIA
2-year Post-Doc position

Final application: 5Th April 2024
Audition: 25th April 2024
Start date: Autumn 2024

Title:
Designing a framework for benchmarking self-supervised deep-learning representations: how robust and fair are you?

PI: Elsa Angelini, Department Image-Data-Signal, Telecom Paris.

Co-I: Loic Le Folgoc, Department Image-Data-Signal, Telecom Paris.

Abstract: Multiple challenges are faced when developing new deep-learning methods on medical images, including: (1) between-sample variability of image quality, (2) within-sample variability of visual features contrast and orientation, (3) need for patch sampling to handle very large image size, (4) high frequency of presence of image artifacts, (5) high-class imbalance within and between samples.
These challenges are poorly documented in publications while they drastically affect the level of "robustness and fairness" (RAF) of the underlying AI encoding tool, and bear great risk of poor performance on new image cohorts, if not handled carefully.
Regarding deep-learning tools, encoding of image information via **self-supervised learning** of visual features on proxy image transformation tasks is driving major breakthroughs in computational photography (e.g. for multi-class classification tasks). Translation of this novel AI encoding to medical images is a natural avenue pursued by few groups. Open **challenges** with self-supervised learning include added learning computational cost and lack of explainability for a downstream classification or labeling task.
This **project** will **aim** to propose image transforms and dedicated metrics documenting the robustness and fairness (RAF) level of self-supervised learned visual features in the context of specific medical imaging analysis tasks (e.g. organ segmentation, lesion detection, scan-level classification). This project will also prepare reference train and test data sets, based on existing open cohorts of images, to be used to enforce and demonstrate the RAF level of a given algorithm. These reference datasets will be carefully documented regarding the presence of rare visual characteristics, rare scanner types and rare disease categories.

Keywords: deep learning, medical imaging, self-supervised learning, domain transfer. Fair AI, explainability

Scientific Scope:

The **overall scope** of this project is to come up with a Robustness and Fairness (RAF) score measured on reference imaging datasets with various levels of complexifying context. Our use case will focus on the use of recent **self-supervised learning (SSL) or fondational models** architectures to encode visual features from medical images for various downstream analysis tasks, including pixel-level segmentation and image-level scoring of pathological signs (eg. presence of tumors, fibrosis).

We are interested in both **theoretical modeling** on the *a-priori* risks of lack of robustness and emergence of biases for different **learning design setups** (e.g encoding architecture, transfer learning domains, quality of ground-truth, choice of proxy task in SSL, choice of foundational model) as well as **experimental quantification of observed lack in RAF performance** from such **learning design setups** on benchmark clinical image datasets, to the benefit of the community of researchers.

The **theoretical modeling** part can take two parallel paths: (1) derive mathematical models of image variability effects on the learning design setups to predict RAF metrics ; (2) Design new test metrics in latent space (eg. analogous to the inception metric used on GANs) to early detect biases and patterns of classification errors.

The project also comprises the **preparation of benchmark datasets** for different tasks (image segmentation, image classification). We will focus on diversifying our datasets in terms of: (1) controlling the degree of out of distribution cases (i.e. rare or absent patterns from typical training data in some statistical sense to define), (2) proportion of challenging cases (e.g. with artefacts, poor image quality, rare disease,...). Preparing these datasets will involve **designing new paradigms** for sub-cohorts **aggregation** and image-level **alterations**. These benchmark datasets will be accessible via a **dedicated web portal**: one part is downloadable (for users to use for training) and one part remains hidden and is accessed for testing uploaded user codes.

Some **examples of self-supervised learning strategies** of interest include: Contrastive Predictive Coding, Rotation prediction, Jigsaw puzzles, Relative patch location, and Exemplar networks.

The choice of the use case will be decided based on the post-doctoral candidate's expertise and interest. Examples of use cases include: lung CT scans for airway segmentation (eg. with and without underlying disease), cardiac ultrasound image segmentation (image quality from poor to high).

References:

- Finnveden, Jansson and Lindeberg (2021) "Understanding when spatial transformer networks do not support invariance, and what to do about it", *Proc. International Conference on Pattern Recognition (ICPR 2020)*, pages 3427-3434, extended version in arXiv:2004.11678.
- Taleb A, Loetzsch W, Danz N, Severin J, Gaertner T, Bergner B, Lippert C. "3D self-supervised methods for medical imaging", Advances in neural information processing systems. 2020;33:18158-72. With code: https://github.com/HealthML/self-supervised-3d-tasks

Project Timeline:

T6M: First benchmark cohort with documented statistical properties.
T12M: 1st RAF index on two self-learning approaches for two tasks: pixel-level segmentation and image-level classification.
T18M: 1st version of web-portal submission on shared and hidden datasets.
T24M: User-community paper in high-impact journal (eg. Nature Methods)

Future Perspectives:

This project aims to create new tools, shared datasets and overall build a community of users toward new validation standards of deep-learning tools in medical imaging.
One model of a community-driven proposition for standardized evaluation approaches is inspired by the consortium on "Metrics Reloaded" (https://arxiv.org/abs/2206.01653)

Future perspectives will focus on several expansion directions:

(1) Extend the RAF index to several medical image domains (eg. MRIs, CT scans, ultrasound) for various applications (brain, cardiac, cancer) and aggregating various sources of shared cohorts (open challenges, UKBiobank, TCIA, ,..).
(2) Extend the RAF index to multi-modal deep-learning tasks (eg. combining omic and image data).
(3) Extend the RAF index to biological images (via ongoing collaborations with Ecole Polytechnique and Institut Pasteur)

Applications procedure:
Send an email with title "Application to RAF-Media postdoc position" to elsa.angelini@telecom-paris.fr and loic.lefolgoc@telecom-paris.fr including you CV+1-page motivation letter in single pdf. Also check this website for context on the funding and selection process: https://www.ip-paris.fr/en/research/research-departments-laboratories-centers-and-projects/disciplinary-and-interdisciplinary-centers/e4h-interdisciplinary-center-engineering-health/postdoctoral-fellowship-program