Breast cancer risk assessment based on anistropic fractal features using deep learning techniques

L3/M1 internship in machine learning applied to medical image processing

Supervision and contact: Patrice Abry, patrice.abry@ens-lyon.fr, CNRS, Research director; Barbara Pascal, barbara.pascal@cnrs.fr, CNRS, Researcher (LS2N)

Application: Send a CV, a grade report, references and motivations to Patrice Abry.

Location: Laboratoire de Physique, École Normale Supérieure de Lyon.

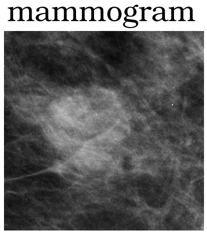
Duration and dates: 3 to 4 months in 2025.

Gratification: Standard net 659.76 euros per month.

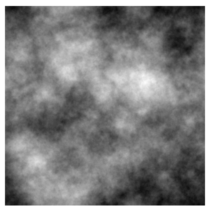
Context: Breast cancer is the most encountered in women, it has been estimated that approximately one woman over eight in North America will receive a breast cancer diagnosis during her lifetime, and one over thirty-four will die from it. Numerous studies have demonstrated that early detection is key in the patient's survival, triggering intense research to improve breast cancer screening techniques. In parallel to the improvement of imaging equipment, advanced computer-aided diagnosis tools has been developed to help radiologists to make a decision.

The main imaging technique used for breast cancer screening consists in X-ray scans of the breast, providing so-called *mammograms*, an example of which being provided on the right (top image). It has soon be noted that the textures appearing in mammograms exhibit some self-similarity, i.e., *fractal*, properties, see on the right the comparison between a zoom on the mammogram (bottom left) and a synthetic fractal texture (bottom right). Further analysis showed that the local *self-similarity exponent* $H \in (0,1)$ of a mammogram reflects the local properties of the breast tissues [1]. Notably it is possible to discriminate between fatty (translucent to X-rays) and dense (absorbing X-rays) tissues, the former corresponding to anti-correlated fractal textures, characterized by a fractal index $H < 1/2$, while the later to correlated fractal textures, characterized by $H > 1/2$.

mammogram



fractal texture



Recent works postulated that the ability of the overall mammographic environment to impedes cancerous cell growth relies on a complex tissue organization, whose destruction could be quantified by the amount of *disrupted* tissues, corresponding to *uncorrelated* fractal textures characterized by $H = 1/2$. They demonstrated that the percentage of the mammogram corresponding to uncorrelated fractal textures differs significantly between tumorous and healthy breasts. Even without seeing a tumor, a high quantity of disrupted tissue indicates high risk of cancer and advocates for more frequent screenings.

Challenge: Previous works leveraging fractal analysis tools to asses breast cancer risk [2, 3] have two main limitations. First, they resort to patch-wise estimation of the fractal features, leading to a significant loss in resolution potentially impacting strongly risk assessment accuracy. Second, only isotropic self-similarity is considered, through a unique fractal attribute H while visual inspection of mammograms raise suspicion of anistropic fractal properties.

Objectives: The main goal of this internship is to design deep learning algorithms for mammogram classification into normal, benign or cancer based on the anisotropic fractal attributes extracted at the pixel level via a recently proposed regularized estimation approach [4].

The first task will be to design a multi-channel image classification neural network for mammogram classification taking as input pixel-wise 12-dimensional fractal features provided by the procedure designed in [4]. Then, a training and a testing datasets will be extracted from a mammogram database shared by the CompUMAINE research group[1] and possibly several training strategies will be devised. Finally, the trained neural networks will be compared on the test set to determine the best deep-learning algorithm for anisotropic fractal attributes-based tumorous vs. normal breast classification.

Prerequisite: The recruited intern is expected to be at ease with the basic concepts of statistics and machine learning. Knowledge in image processing, with a special focus on medical imaging, would be a plus. Good programming skills in Python are required, in particular in the use of PyTorch environments.

References

[1] P. Kestener, J.-M. Lina, P. Saint-Jean, and A. Arneodo. Wavelet-based multifractal formalism to assist in diagnosis in digitized mammograms. *Image Anal. Stereol.*, 20(3):169–174, 2001.

[2] Z. Marin, K. A. Batchelder, B. C. Toner, L. Guimond, E. Gerasimova-Chechkina, A. R. Harrow, A. Arneodo, and A. Khalil. Mammographic evidence of microenvironment changes in tumorous breasts. *Medical physics*, 44(4):1324–1336, 2017.

[3] E. Gerasimova-Chechkina, B. C. Toner, K. A. Batchelder, B. White, G. Freynd, I. Antipev, A. Arneodo, and A. Khalil. Loss of Mammographic Tissue Homeostasis in Invasive Lobular and Ductal Breast Carcinomas vs. Benign Lesions. *Front. Physiol.*, 12:660883, 2021.

[4] L. Davy, N. Pustelnik, and P. Abry. Combining dual-tree wavelet analysis and proximal optimization for anisotropic scale-free texture segmentation. In *IEEE Acoust. Speech Signal Process.*, pages 1–5, Rhodes, Greece, 2023. IEEE.

[1] `https://umaine.edu/compumaine/`