



HAL
open science

Improved glioma delineation by autofluorescence and 5-ALA-induced PpIX fluorescence spectroscopy using machine-learning data normalization and augmentation

Hermine Quardon, Arthur Gautheron, Thiébaud Picart, Cédric Ray

► To cite this version:

Hermine Quardon, Arthur Gautheron, Thiébaud Picart, Cédric Ray. Improved glioma delineation by autofluorescence and 5-ALA-induced PpIX fluorescence spectroscopy using machine-learning data normalization and augmentation. *Manutech Sleight #13*, Jan 2025, Saint-Etienne, France. 2025. hal-04869575

HAL Id: hal-04869575

<https://hal.science/hal-04869575v1>

Submitted on 7 Jan 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

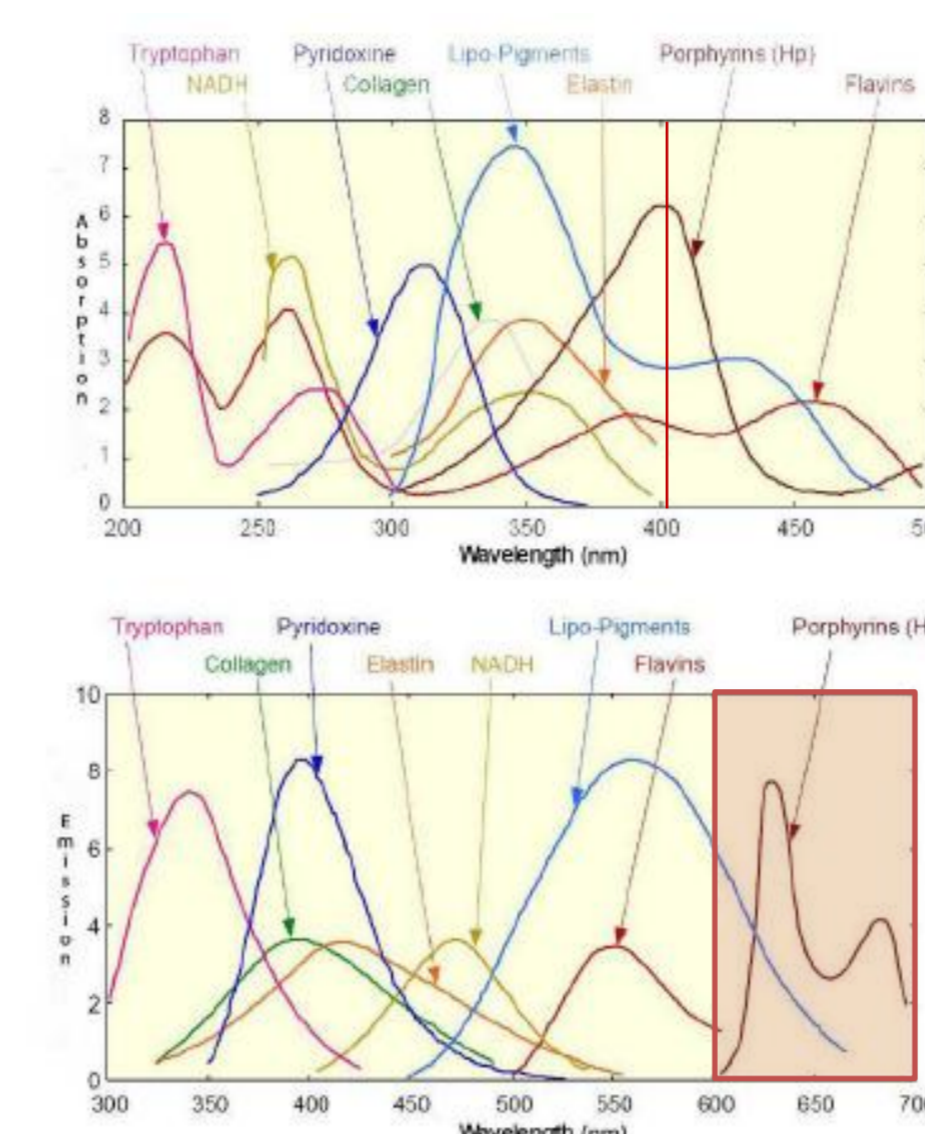
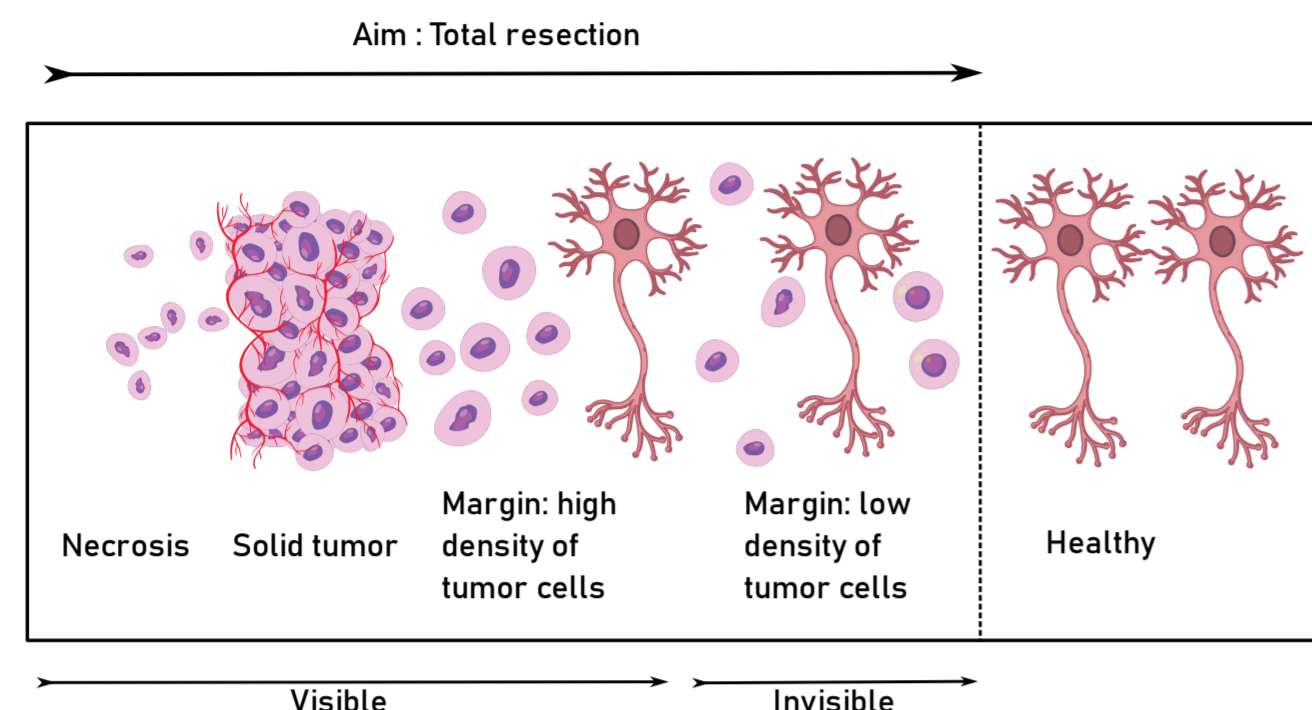
Improved glioma delineation by autofluorescence and 5-ALA-induced PpIX fluorescence spectroscopy using machine-learning data normalization and augmentation

Hermine Qardon¹, Arthur Gautheron^{1,2}, Thiébaud Picart^{1,3} and Cédric Ray¹

¹CREATIS, CNRS (UMR 5220); INSERM (U1294); INSA Lyon, Université de Lyon, France; ²CPE, Lyon, France; ³Department of Neurosurgery, Hôpital Neurologique Pierre Wertheimer, Hospices Civils de Lyon, 69677 Bron, France;

Between healthy cells and gliomas : an invisible margin

Gliomas are infiltrative primary brain tumors, meaning that their margins are **not well-defined** and can be **difficult to distinguish** from healthy tissue. Resecting far from the tumor is not a viable option as the neurosurgeon has to **preserve functional** brain areas. Consequently, the main challenge is to identify the **invisible areas infiltrated** by the tumor, especially in Low Grade Gliomas (LGG) which are more infiltrative than High Grade Gliomas (HGG).



The 5-ALA molecule, orally taken by the patient, is a precursor of **protoporphyrin IX (PpIX)**, which selectively accumulated in tumor cells. PpIX emits **fluorescence** when **excited by a laser (two wavelengths are used – 375 and 405 nm)** and **eases the detection of tumor cells**. Endogenous fluorophores present in tumor cells are also excitable and produce **autofluorescence** in the same spectral bandwidth as the PpIX one. Yet, omitting an endogenous fluorophore or using an incorrect emission spectrum **can lead to an over or under evaluation** of the fluorescence signal of the PpIX. This work aims to develop a biopsy-selection-assisted per-operative tool.

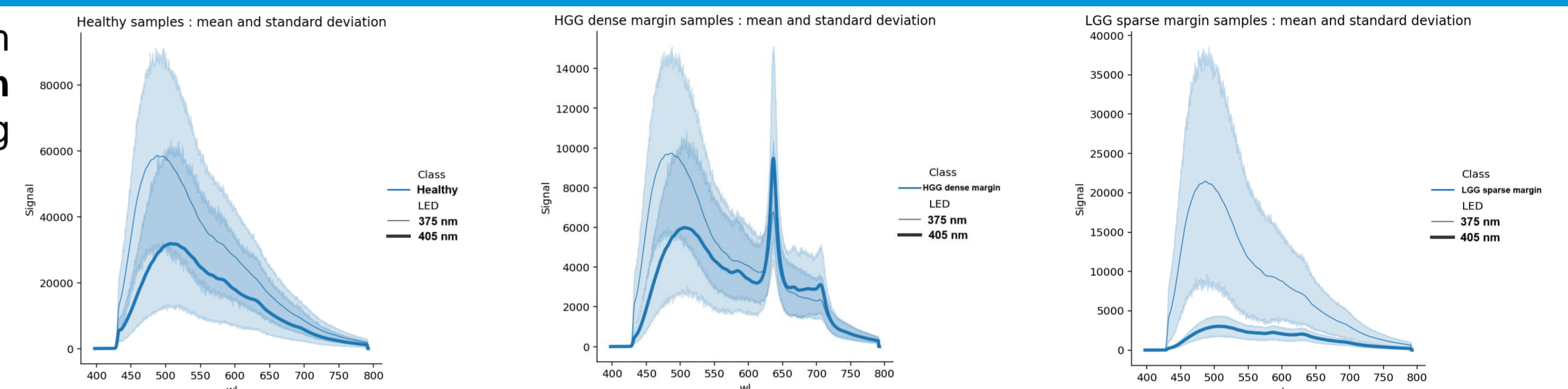
Fluorescence spectra normalization and data augmentation

The interest of normalizing data acquired from different patients

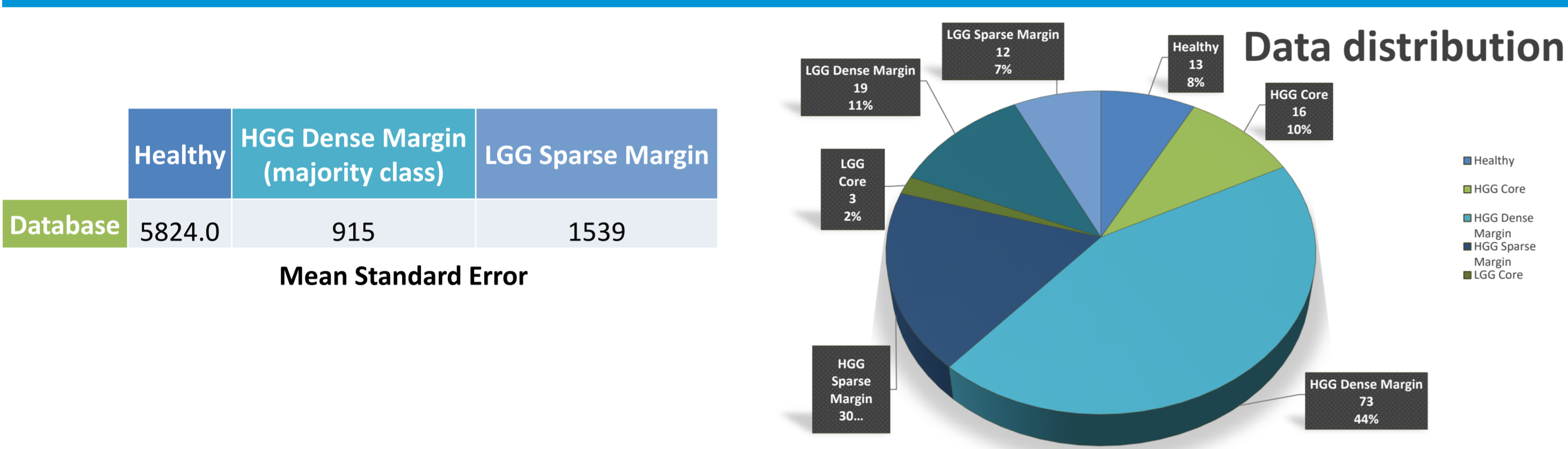
The fluorescence spectra obtained from 35 patients with the experimental system developed by A. Gautheron during his PhD thesis [1] **highly vary in intensity for each patient**. Thus, it is required to **normalize** the spectra before applying machine learning algorithms.

Several ways to normalize the data have been explored :

- Dividing each spectrum by its respective emission laser **power** ;
- Dividing each spectrum by the total **autofluorescence** intensity ;
- Dividing each spectrum by the **healthy samples** mean intensity.



How to manage a small number of unbalanced data ?



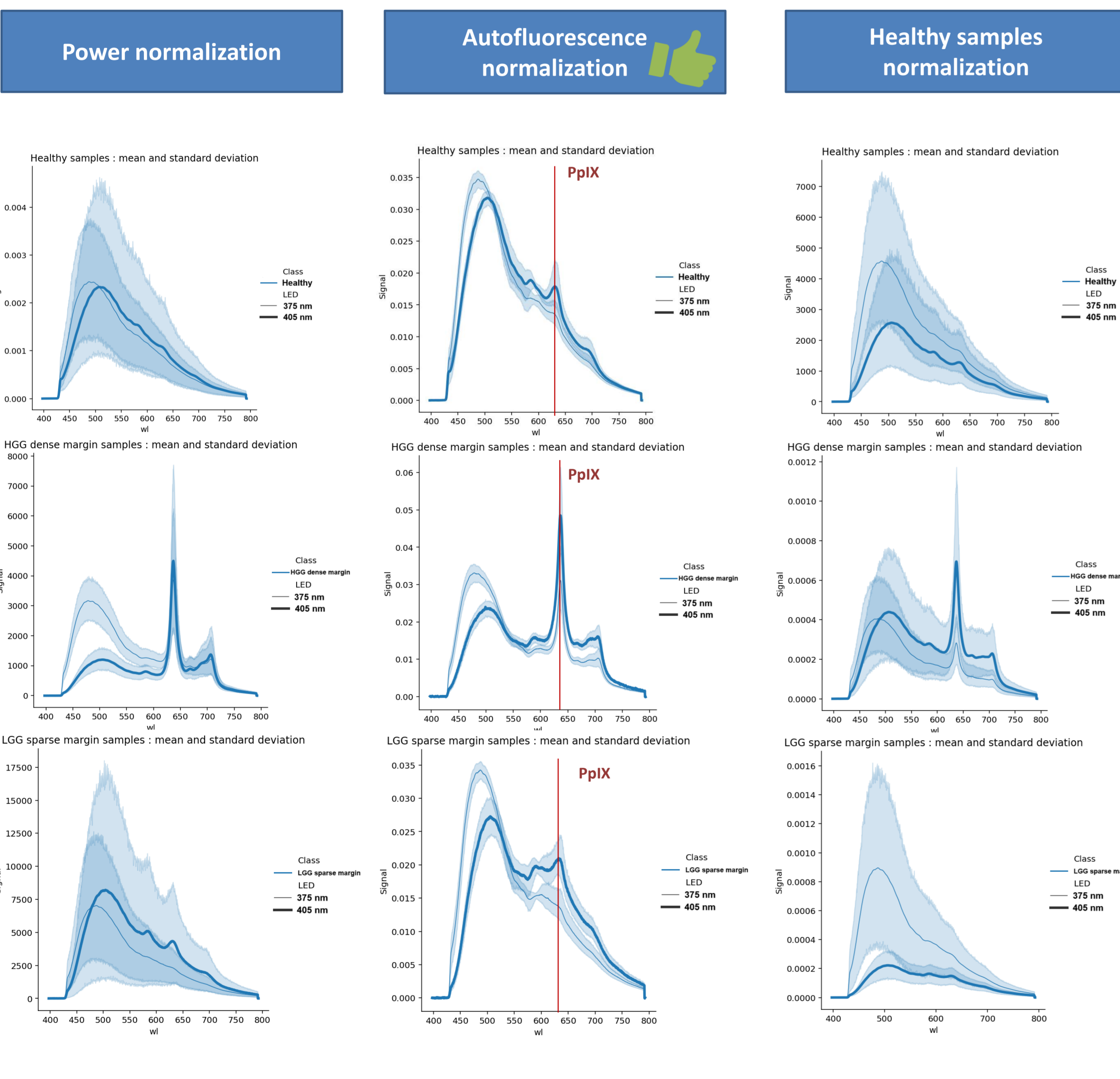
Unfortunately, the size of the database and the balance between the different classes depends on the **number of patients**, the **type** of their tumors and the number and location of the **samples**.

Two data enhancement algorithms have been tested :

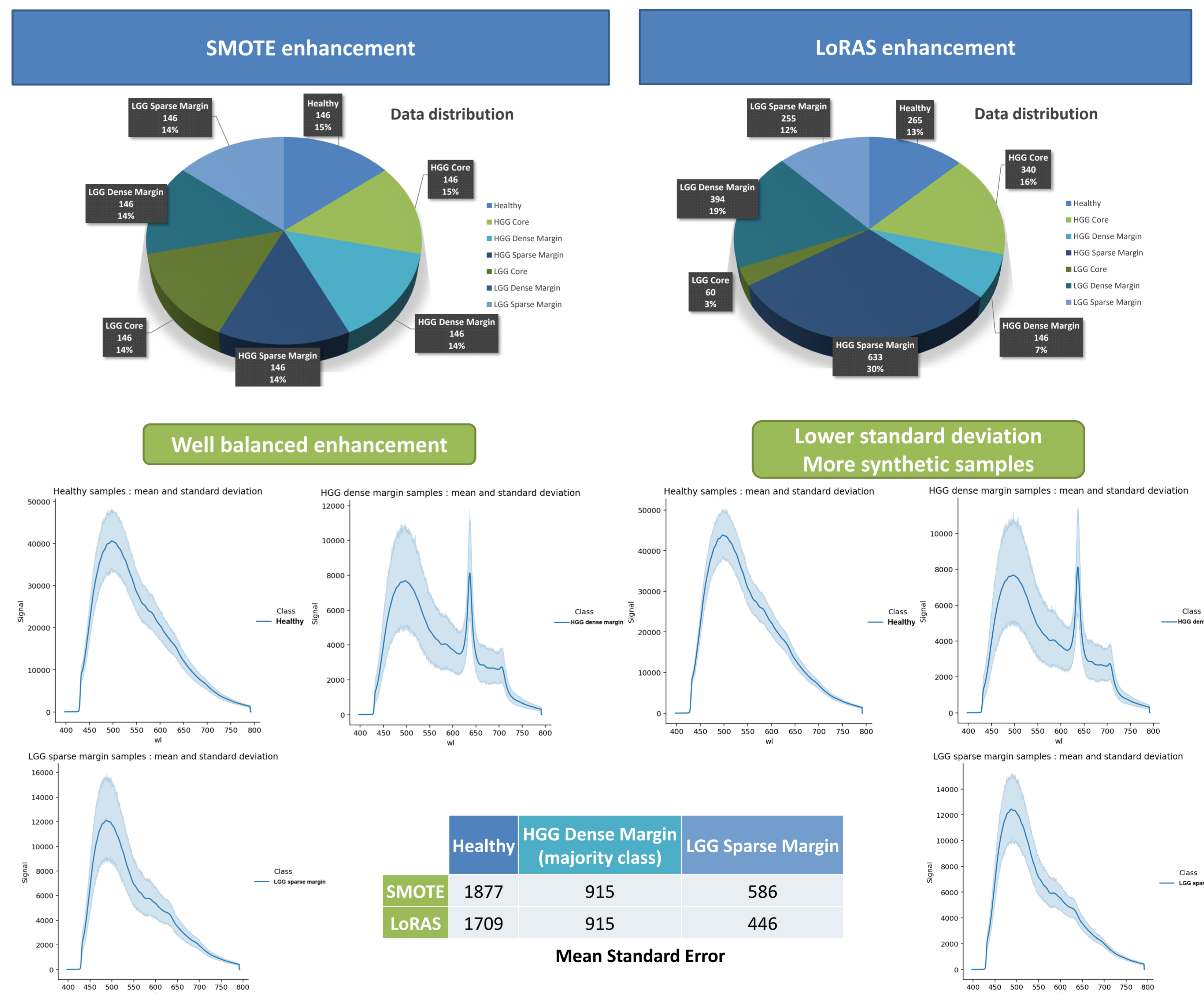
- **Synthetic Minority Over-sampling Technique (SMOTE)** [2] : introduce synthetic examples in the neighborhood of the minority class samples.
- **Localized Random Affine Shadowsampling (LoRAS)** [3] : create random affine combinations of noised samples to keep the synthetic samples in the same manifold as the minority class.

Results and Discussion

Data Normalization



Data Augmentation



Sample dependent

Molecular and genetic differences taken into account

No PpIX enhancement

Data normalization

Unbalanced data machine learning

Acknowledgments

This work was funded by the European Union's Horizon Europe research and innovation program under grant agreement No 101071040. This work was also funded by the Graduate School Medical Device Engineering of the Université Claude Bernard Lyon 1 within the program "Graduate +" (21-SFRI-0001) operated by the French National Research Agency (ANR). This material is based upon work done on the PILOT facility (PILOT, INSA LYON, Bât. L. de Vinci, 7 Av. Jean Capelle 69621 Villeurbanne).

References :

- [1] Arthur Gautheron. Towards quantitative multispectral excitation fluorescence spectroscopy for intraoperative assistance in glioma resection in neurosurgery. Biotechnology. Université Claude Bernard Lyon 1, 2022. English. NNT : 2022LYO10145. tel-04092778
- [2] Leland McInnes, John Healy, James Melville, UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, arXiv:1802.03426 [stat.ML]
- [3] S. Bej, P. Srivastava, M. Wolfien and O. Wolkenhauer, "Combining uniform manifold approximation with localized affine shadowsampling improves classification of imbalanced datasets," 2021 International Joint Conference on Neural Networks (IJCNN), Shenzhen, China, 2021, pp. 1-8, doi: 10.1109/IJCNN52387.2021.9534072

Perspectives

- Enhance the database with new acquisitions during future surgeries.
- Apply several machine learning algorithms and tune the hyperparameters to get the best classification model.

