Integration of an AI engine for content-based image retrieval in a radiological image database

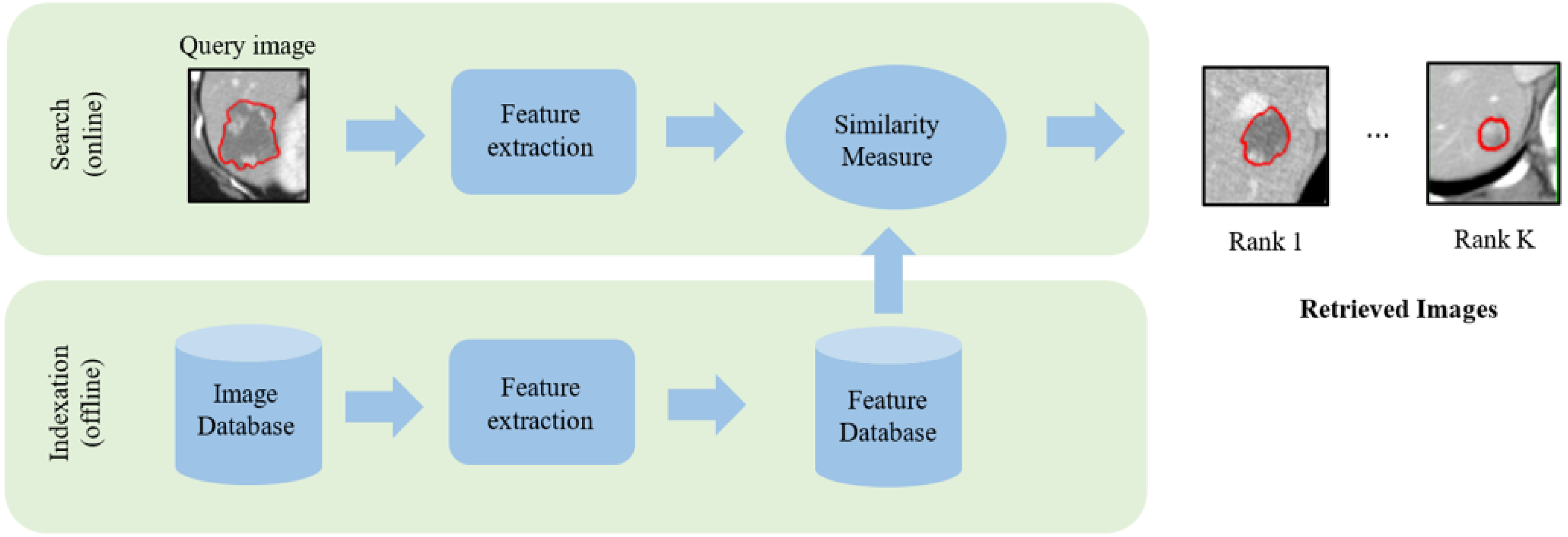
Thibault SAURON[1,3], Florence CLOPPET[1], Camille KURTZ[1], Antoine OLIVIER[2], Laure FOURNIER[3]
1: LIPADE, Université Paris Cité, Paris, France
2 : PHILIPS Research France
3 : Hôpital Européen Georges Pompidou

Context

- **PACS** (Picture Archiving and Communication Systems) are the database systems used in hospitals to store, edit and retrieve patient information
- **The goal is to perform CBIR** (Content Based Image Retrieval) on those PACS to allow query by example to be performed in clinical routines

Constraints :
- Multi sequence images
- Metric is visual similarity
- Lack of annotated data

Neural Networks for CBIR : overview

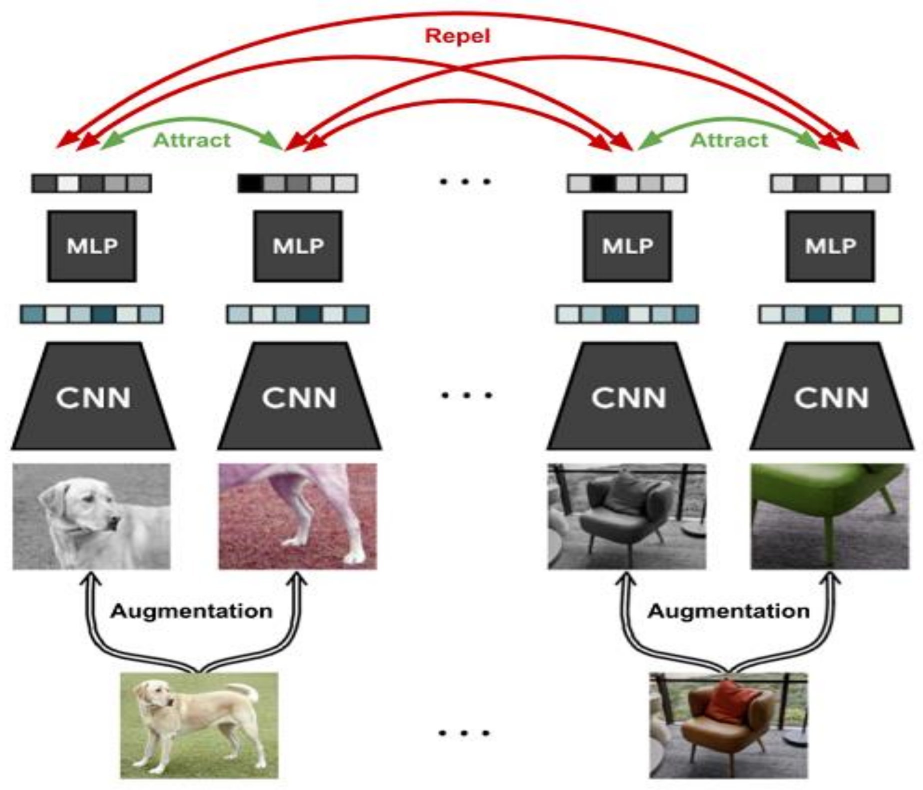


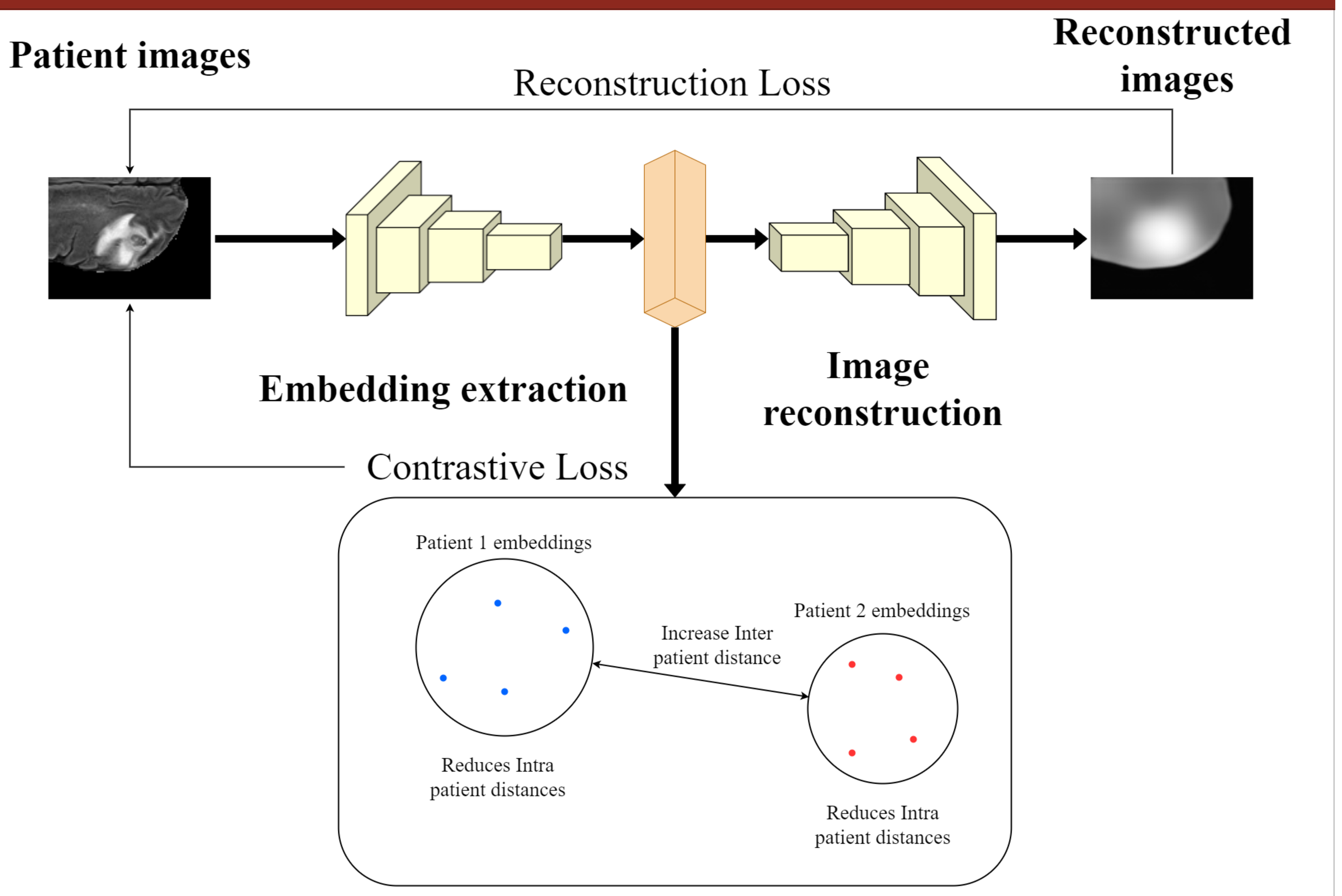
Leveraging patient information for unsupervised multi-sequence Contrastive training

Contrastive learning principles :
- Pulling closer together in the latent space the positives examples
- Pushing apart the negatives examples
- Positives can be hard annotation or different view of the same image (SIMCLR [1])

Contrastive learning advantages :
- Well suited for **multi sequence** tasks, we can consider every MR image to be a different sequence
- Adapts well to **unsupervised** tasks

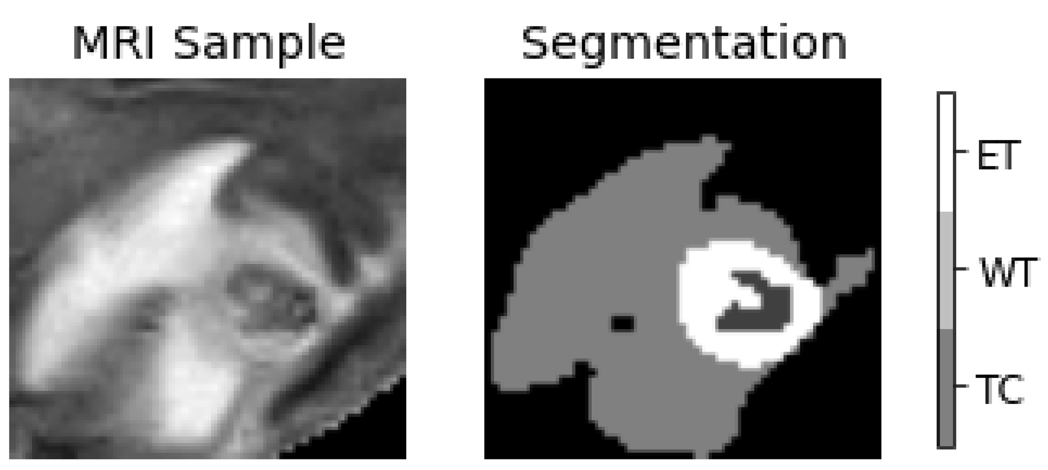
Training strategy :
- Considering every parametric MR image as a positive sample from a patient in the supervised contrastive setting [2]
- Adding a **visual reconstruction** path from the contrastive embeddings [3]



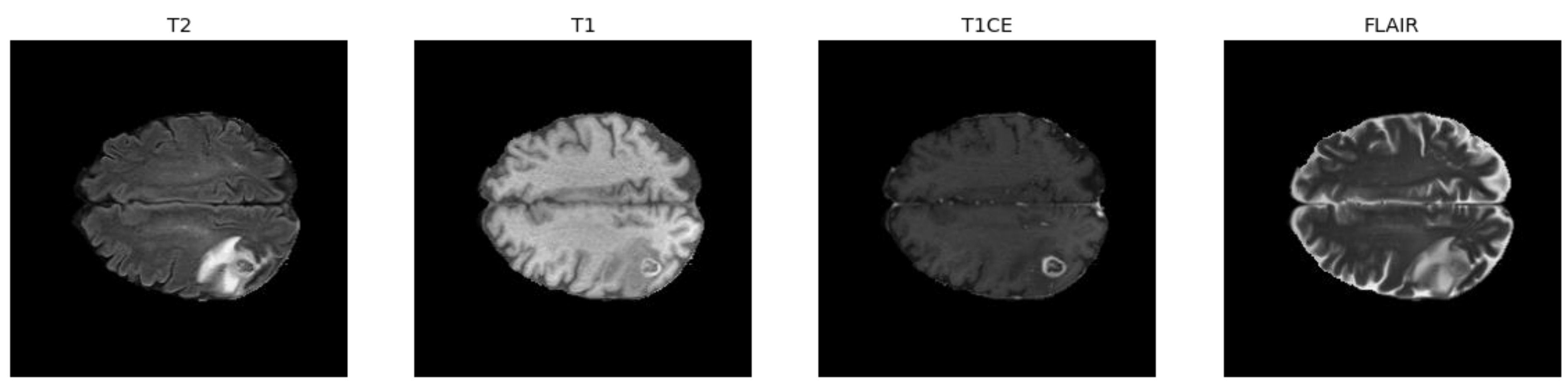


Dataset : RSNA-ASNR-MICCAI Brain Tumor Segmentation (BraTS) Challenge 2021 [4]

- Dataset is a mpMRI segmentation dataset of brain glioblastoma split in 3 sub-regions : the "enhancing tumor" (ET), the "tumor core" (TC), and the "whole tumor" (WT)



- 4 MRI parameters : T2, T1, T1CE, FLAIR



Evaluation strategy and results

Assessing visual similarity
- Using combination of standard image metrics to compute a visual distance between two images
- Using **nDCG** as the metric for retrieval evaluation to take the rank into account :
- $nDCG@n = \frac{DCG@n}{iDCG@n}$ $DCG@n = \sum_{i=1}^{n} \frac{rel(i)}{\log(1+i)}$ iDCG is best score possible,
- We used IOU, gray histogram distances and GFD distances as relevance function

Training
- Using **Resnet18** [5] as images are not to complex
- Testing different training variants to assess the impact of each isolated change
- Networks are trained on the best slice of the volume

Network	Score (nDCG@ 5,10,20)
Resnet18 trained from scratch	0.7587, 0.7287, 0.718
Resnet18 pretrained, small batch size	0.7704, 0.7351, 0.7215
Resnet18 pretrained, big batch size	0.7804, 0.7462*, 0.7336*
Resnet18 pretrained with image augmentation	0.7902, 0.7412, 0.7389
Resnet18 pretrained with reconstruction loss	0.7964*, 0.7451, 0.7198
Resnet 18 pretrained on imagenet, no retraining	0.7882, 0.7406, 0.7254
Random	0.41444, 0.4392, 0.47198

Perspectives

- Testing different images augmentation and their impact, using the whole brain volume instead of only the best slice
- Weighting the contrastive and the reconstruction loss, improving the reconstruction loss
- Training separate neural networks for every MR sequence
- Using another database [6] with annotations more suited to the evaluation task and closer to the real-world scenario. However this database is not yet **anonymized**, sufficiently **annotated** and **curated**.

References

[1] T. Chen, et al., “A Simple Framework for Contrastive Learning of Visual Representations”, arXiv:2002.05709, 2020
[2] A. Maschinot, et al., “Supervised Contrastive Learning”, arXiv:2004.11362, 2021
[3] K. Kobayashi, et al., “Decomposing normal and abnormal features of medical images for content-based image retrieval of glioma imaging”, MIA, 2021
[4] U.Baid, et al., "The RSNA-ASNR-MICCAI BraTS 2021 Benchmark on Brain Tumor Segmentation and Radiogenomic Classification", arXiv:2107.02314, 2021
[5] K. He, et al., “Deep Residual Learning for Image Recognition”, CVPR, 2015
[6] I. Thomassin-Naggara et al. “Ovarian-Adnexal Reporting Data System Magnetic Resonance Imaging (O-RADS MRI) score for risk stratification of sonographically indeterminate adnexal masses”, JAMA network open 1, 2020

PHILIPS
LIPADE Laboratoire d'Informatique PAris DEscartes
ASSISTANCE PUBLIQUE HÔPITAUX DE PARIS