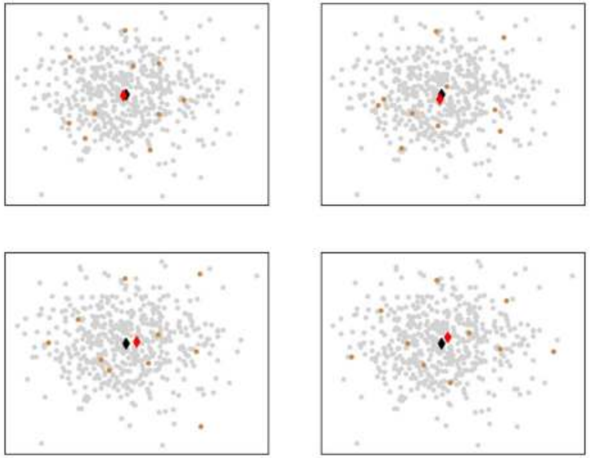
JCJC ENDIVE

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

ENDIVE
Bastien Pasdeloup
14/11/2024

ENcouraging DIVErsity in few-shot learning



(b) Examples of a DPP sampling realization of 10 points (in orange) among 500 (in gray). The mean of the sampled points (resp. of the entire point cloud) is depicted with a red (resp. black) diamond.

(a) Labeled image (class A)

(b) Labeled image (class B)

(c) Unlabeled image to classify

Figure 1: Example FSL image classification task from the miniImageNet dataset [11]. For each possible class, only one labeled example is provided. Here we can observe the ambiguity of the task, as expected class would be A, although the human, the room, or even records in the background can act as confusing elements.

OUTLINE

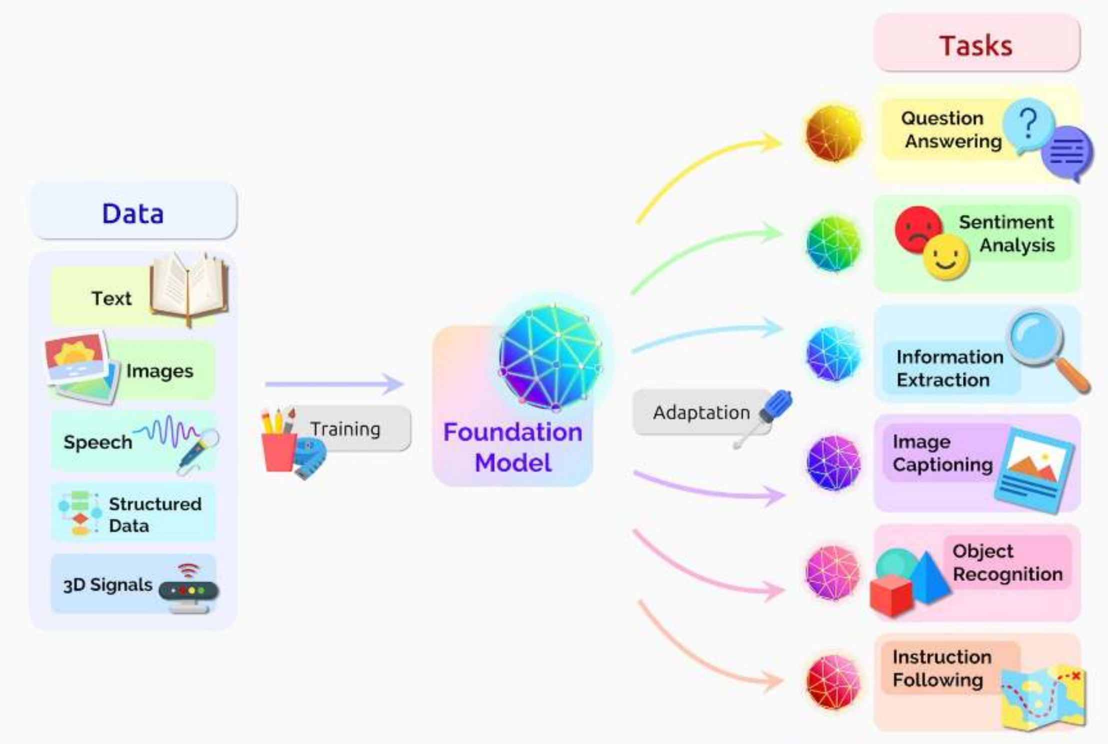
- Context, positioning & objectives
  - General context
  - Objectives
  - State of the art
  - Methodology

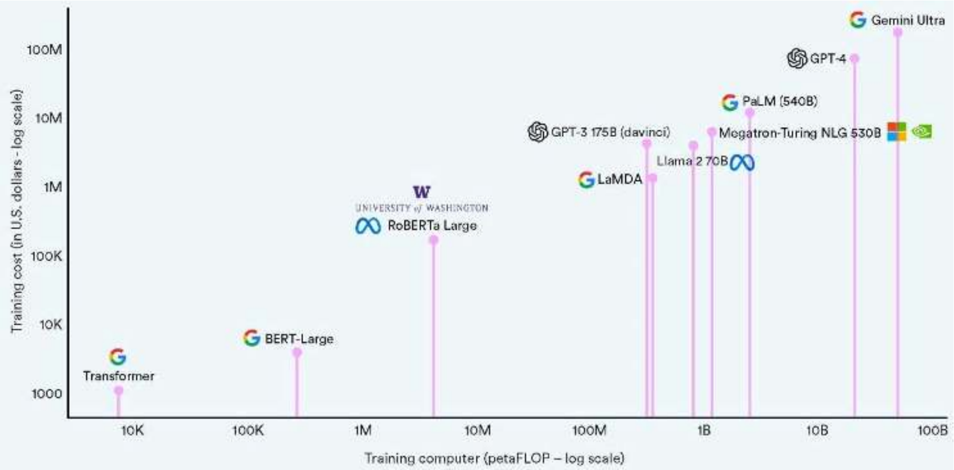
- Organization & implementation
  - Project team
  - Timeline
  - Resources

Context, positioning & objectives

└ Objectives & research hypothesis - General context - Foundation models

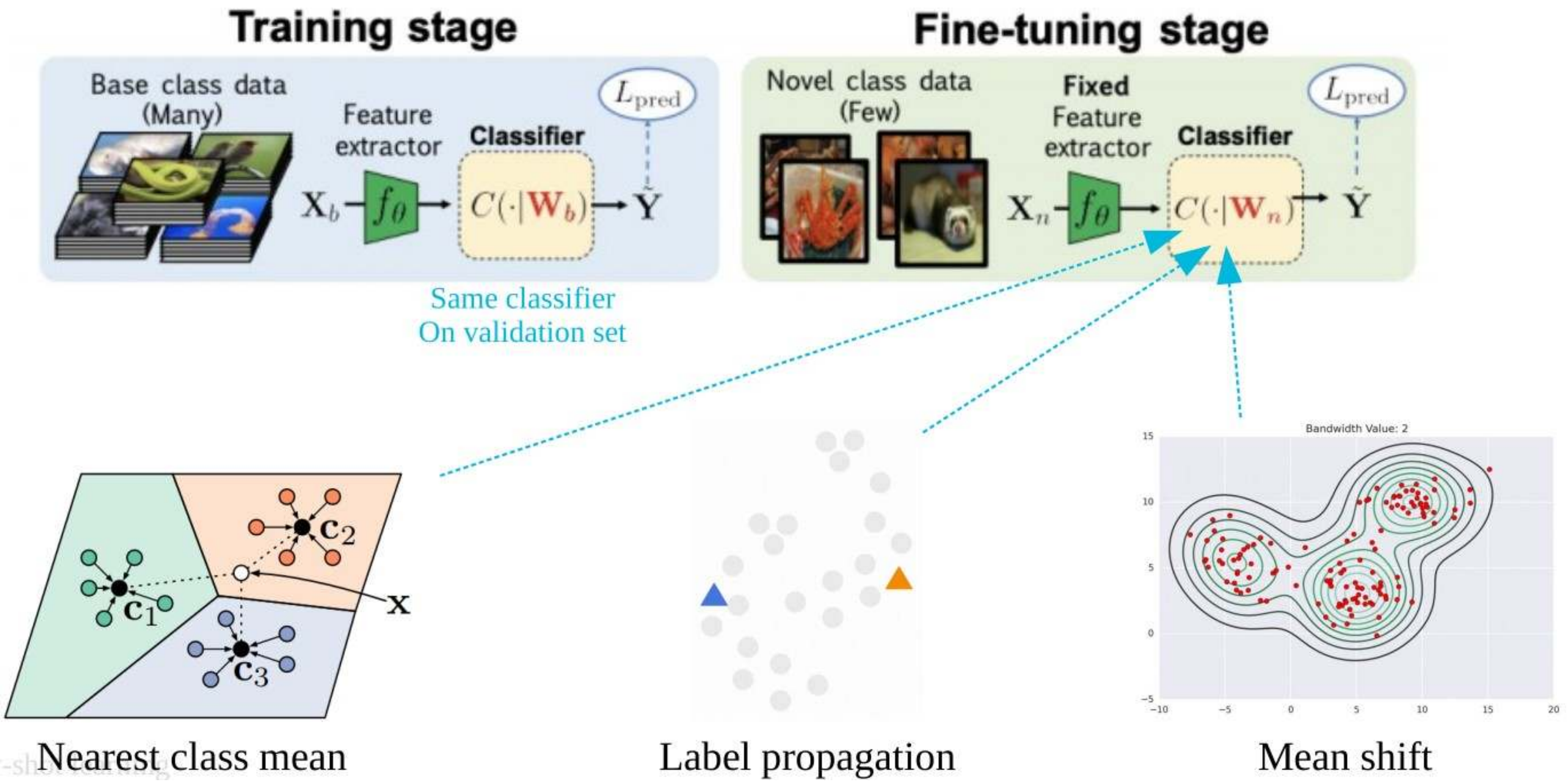
https://blogs.nvidia.com/blog/what-are-foundation-models/
https://pureai.com/Articles/2024/04/23/Open-Source-Models-Cost.aspx
https://www.wired.com/story/meta-llama-ai-gpu-training





« In March, Meta and Nvidia shared details about clusters of about 25,000 H100s that were used to develop Llama 3. In July, Elon Musk touted his xAI venture having worked with X and Nvidia to set up 100,000 H100s. »

Chen *et al.*, A closer look at few-shot classification



Bendou *et al.*, Disambiguation of one-shot visual classification tasks: a simplex-based approach

Fig. 1. Image extracted from the MiniImageNet benchmark for visual few-shot classification. This image was automatically found to be the one in the benchmark that has the most negative impact to performance when it appears in randomly sampled few-shot tasks. Indeed, even though it contains multiple objects in the foreground, the class it is associated with is "library", creating dramatic ambiguities when it is selected as the only example from that class.

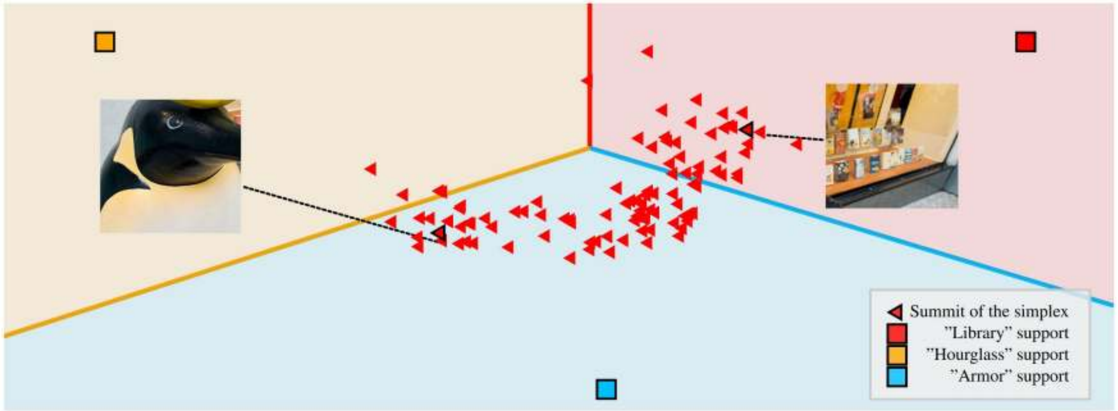


Fig. 2. Representation of the feature vectors associated with 200 random crops from the image depicted in Figure 1. The feature extractor used here, a ResNet-12 from [2], outputs vectors with 640 dimensions. For visualization purposes, we chose to project these high-dimensional vectors onto a 2d-plane containing the centroids for the classes "armor", "hourglass" and "library", the three classes in the MiniImageNet dataset that are the closest from these feature vectors. We also depicted in color the Voronoi cells corresponding to these three classes. We see that the feature vectors display an interesting distribution following what resembles an arc, starting from crops close to the "library" class on the top right to crops close to the "armor" class on the bottom left. The vertices circled in black correspond to those automatically found using the methodology described in 3.1. The images corresponding to the closest crops to these vertices are displayed. We observe that both the object "penguin" and the object "library" have been correctly retrieved from this example.

Barthelmé *et al.*, Tutorial: Determinantal Point Processes and their Application to Signal Processing and Machine Learning

Figure: Example of DPP sampling

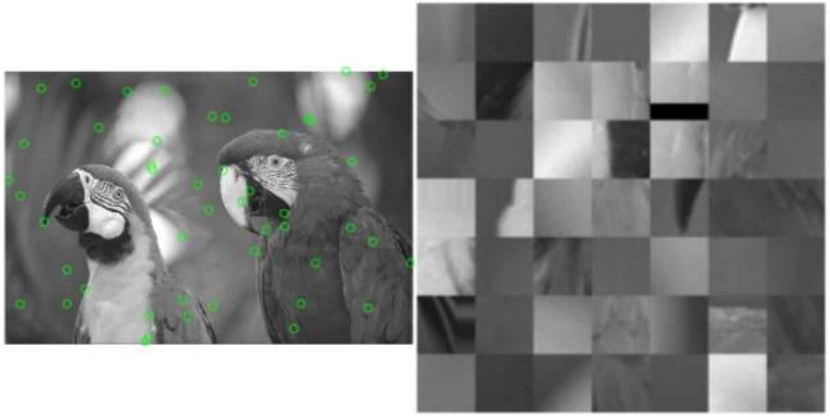


Figure: Example of iid uniform sampling

« Can diversity help improving performance of FSL,
while leveraging powerful state-of-the-art models,
even when it is impossible to train them due
to the requirements in computation power or data availability? »

Data

D1 – Can we predict the difficulty of a FSL task?

D2 – Given an unlabeled dataset, which few points should be labeled to maximize FS classification?

D3 – Are all parts of data relevant?

Features

F1 – Are some features more relevant to FSL than others?

F2 – Can we exploit positional embedding for better FSL classification?

F3 – Can we select good models for a FSL task?

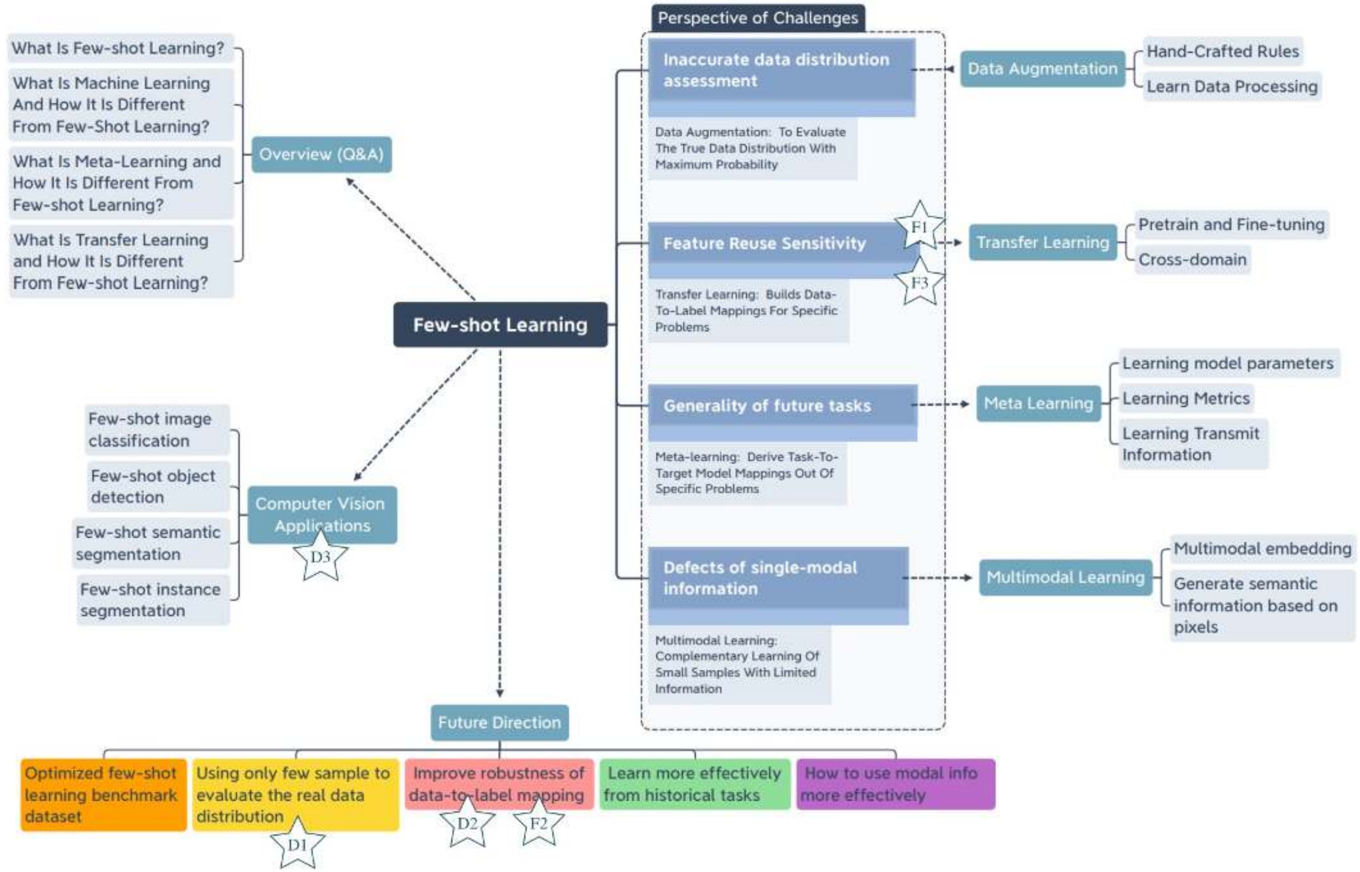
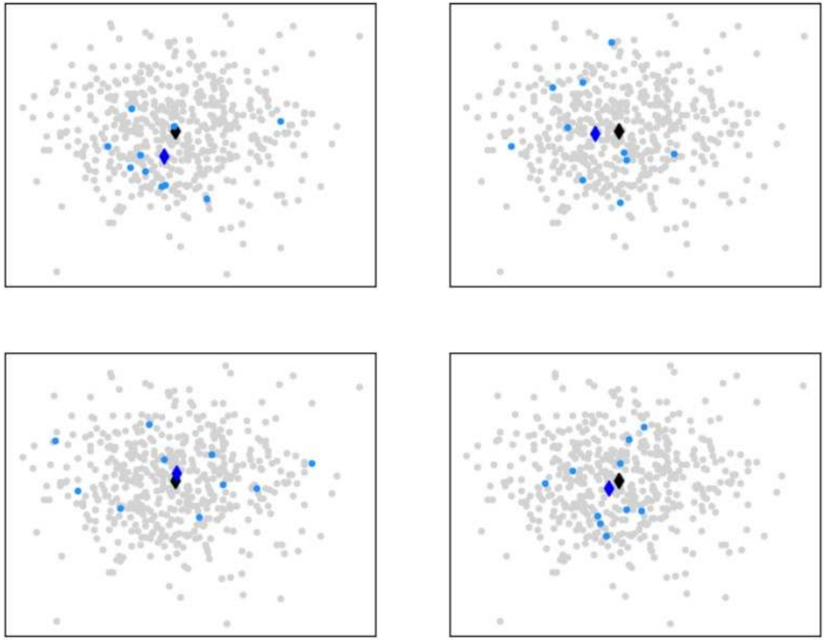


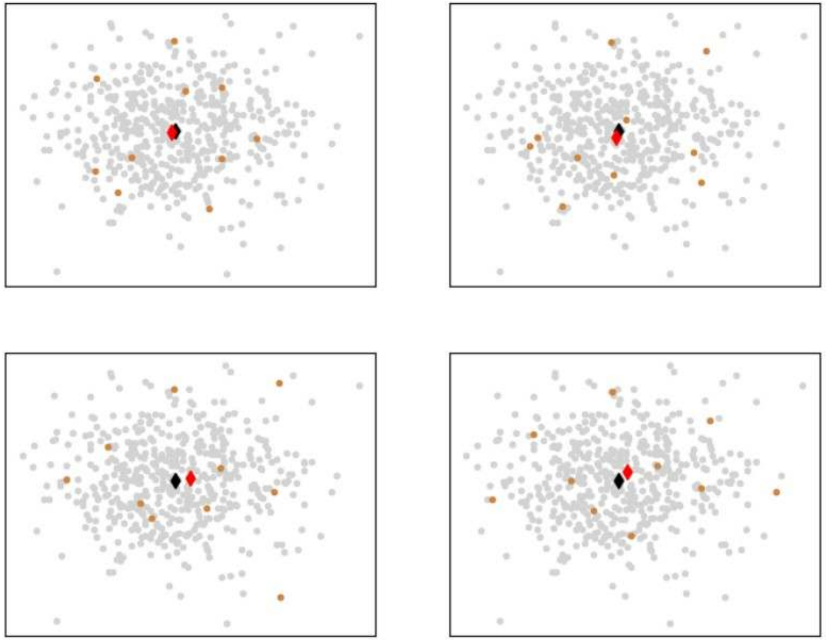
Figure 2: Illustration of the challenges in FSL. Stars indicate elements to which this project will contribute, with associated text corresponding to research questions in Section 1.1.2. Figure adapted from Song *et al.* [2]

Context, positioning & objectives

Objectives & research hypothesis - State of the art - DPP (1)



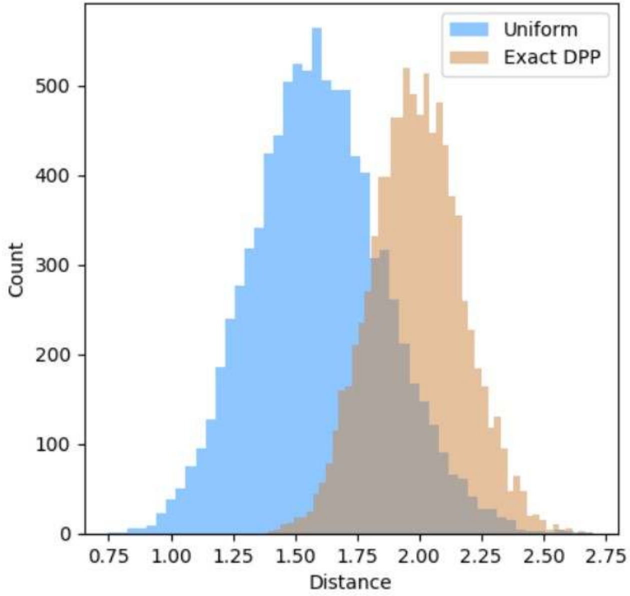
(a) Examples of a uniform sampling realization of 10 points (in blue) among 500 (in gray). The mean of the sampled points (resp. of the entire point cloud) is depicted with a blue (resp. black) diamond.



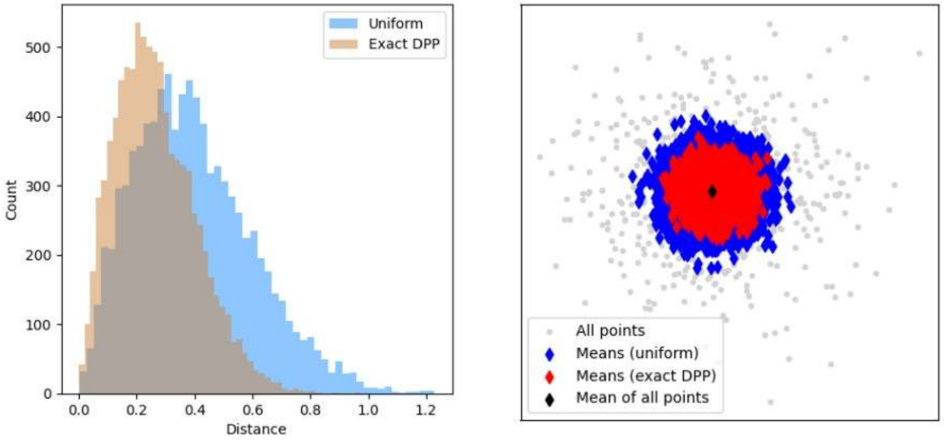
(b) Examples of a DPP sampling realization of 10 points (in orange) among 500 (in gray). The mean of the sampled points (resp. of the entire point cloud) is depicted with a red (resp. black) diamond.

Context, positioning & objectives

└ Objectives & research hypothesis - State of the art - DPP (2)



(c) Distribution of the average inter-point distance, for $9,000$ realizations of each sampling procedure.



(d) Distribution of the distances between the mean of sampled points – *i.e.*, diamonds in (a) and (b) – and the mean of the point cloud, for $9,000$ realizations of each sampling procedure (left); Visual representation of the repartition of these means (right).

Metrics

- Accuracy (standard comparison metric)
- Model size (ease to transfer to the general public)

Datasets

- Vision (miniImagenet, CoOp)
- Brain signals (EEG, fMRI)
- Other medical-related datasets

Open science

- GitHub repository
- Target main open DL conferences
- Contribution to DPPy

Ecological considerations

- No training (as much as possible)
- Leverage pre-trained models

Context, positioning & objectives

└ Objectives & research hypothesis - Methodology - Work packages

WP1 – Evaluating a FSL problem

- ***WP1.1*** – Identify DPP-based metrics that correlate with difficulty
- ***WP1.2*** – Predict performance of a FSL pipeline and model on a new dataset
- ***WP1.3*** – Use previously identified metrics to select the best model for a FSL task

WP2 – Transformers & diversity

- ***WP2.1*** – Exploiting the diversity of instances
- ***WP2.2*** – Exploration of positional embedding to contextualize FSL

WP3 – Selection of relevant parts in data and features

- ***WP3.1*** – Identify data to label from a large unlabeled dataset
- ***WP3.2*** – Determining features that are important for FSL

WP4 – Dissemination

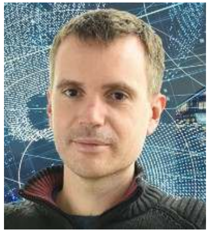
- ***WP4.1*** – Continuous integration
- ***WP4.2*** – Transition toward medical applications

	D1	D2	D3	F1	F2	F3
WP1	×					×
WP2			×		×	
WP3		×	×	×		
WP4	*Dissemination*					

Organization & implementation
Project team

Vincent Gripon

IMT Atlantique, HDR
Deep learning, FSL



Nicolas Tremblay

CNRS (Gipsa-lab)
Machine learning, GSP, DPP

Nicolas Farrugia

IMT Atlantique, HDR
Neuroscience, deep learning

Rémi Bardenet

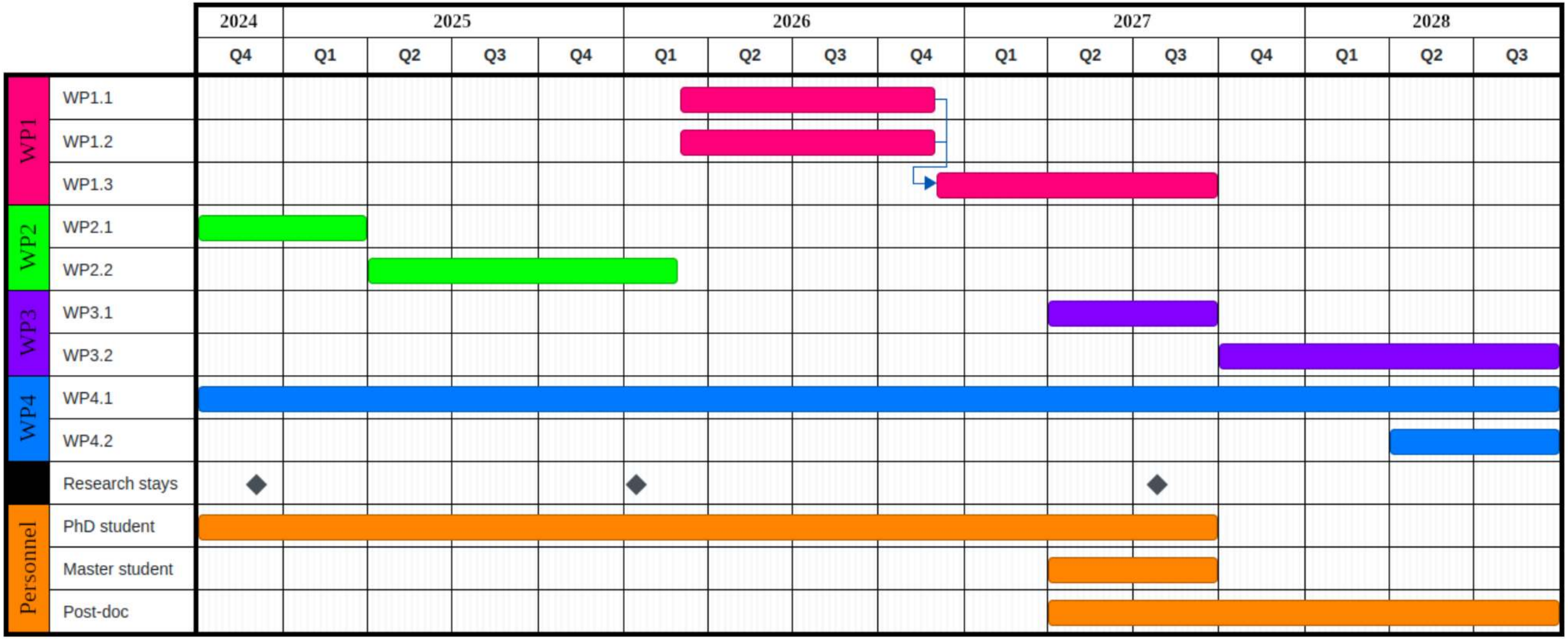
CNRS (CRIStAL), ERC, CNRS bronze medal
DPP, Monte-Carlo methods, AI

Alexandre Reiffers-Masson

IMT Atlantique
Optimization, machine learning

Organization & implementation

└ Timeline



Organization & implementation

Resources

		Partner 1 – IMT Atlantique
Staff expenses, including costs of a partial release from teaching obligations	Permanent (66 p.m.)	278,917.97€
	Non-permanent (60 p.m.)	226,500€
Instruments and material costs		16,000€
Building and ground costs		0€
Outsourcing / subcontracting		5,000€
Overhead costs	Travel costs	11,500€
	Administrative management & & structure costs	37,555€
Sub-total		**575,472.97€**
Requested funding		**296,555€**

JCJC ENDIVE