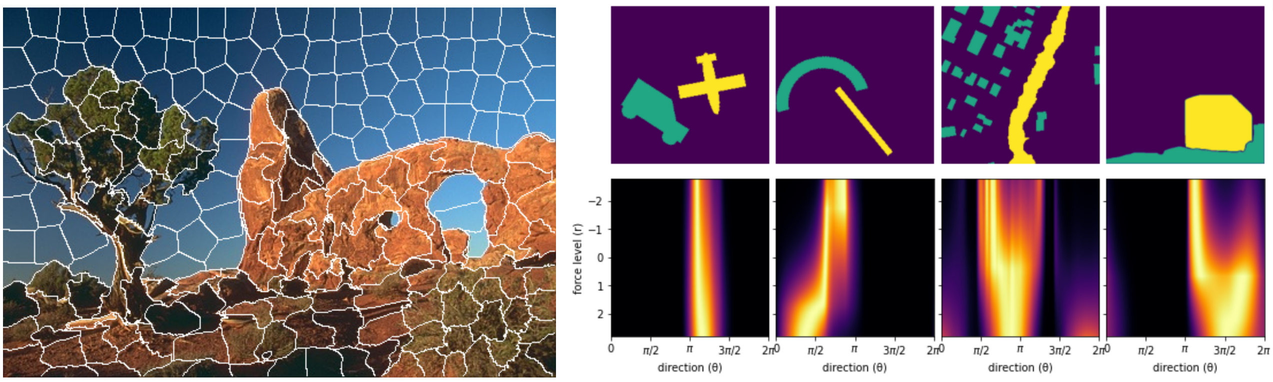
Deep learning models from structural image representations

Environment

Supervisors:	Michaël Clément (Bordeaux INP - LaBRI) michael.clement@labri.fr
	Rémi Giraud (Bordeaux INP - IMS) remi.giraud@ims-bordeaux.fr
Funding:	Ministerial scholarship (€2300 gross per month)
Laboratory:	Laboratoire IMS, campus of Talence (close to Bordeaux)
	Close to other research labs (LaBRI, IMB)
Team:	Signal and Image. 20 permanents, 20 PhD students/interns
Teaching:	Teaching opportunities during the thesis
Keywords:	Deep learning; Computer vision; Image analysis; Segmentation; Spatial relations.



(a) Superpixels (b) Spatial relations

Figure 1: Tools considered during this thesis. (a) Example of superpixel segmentation [1]. (b) 2D representations of spatial relationships between objects in images [2].

Scientific context

State-of-the-art machine learning methods are generally based on training procedures for large neural networks and annotated databases. The resulting models are often large and complex, and can be difficult to evaluate or interpret. In particular, popular computer vision networks make extensive use of convolution blocks, which are pixel-scale operations applied to regular grids. However, these architectures still have certain limitations, particularly when it comes to image size, which must often be limited during the training phase, for reasons of computation time and memory representation. In this context, there is a real challenge to integrate structural or symbolic representations of data in order to efficiently synthesize information and guide the learning process.

A first line of research consists in using a sub-representation of the image, such as segmentation into superpixels [1] (see Figure 1(a)), reducing the number of elements to be processed while respecting the structure and contours of the image objects. However, the absence of a regular structure (shape/size/neighborhood) poses a problem for their use in a convolutional neural network. Initial work has been initiated in [3] to consider hierarchical segmentation into superpixels and in [4] on the definition of convolution operators and attention blocks adapted to superpixels. In this thesis, we propose to deepen and consolidate this work for different image analysis and processing applications.

A second line of research aims to integrate semantic knowledge about the spatial relationships between objects into neural networks. Indeed, the spatial structure of objects in an image constitutes relevant information that can be exploited to interpret their content [2] (see Figure 1(b)). To achieve this, the thesis will draw on our recent work using structural representations in deep learning, for example [5] (graph-based representations of brain structures), as well as on our long-term expertise in learning the characteristics of spatial relationships [6].

Objectives of the thesis

The primary objective of this thesis is thus to successfully overcome the methodological barriers that currently prevent the explicit and natural use of structural representations of images within deep learning models. The research work will be structured around the two methodological axes mentioned above:

1. Design of new neural network architectures that can integrate structural representations (graphs, region hierarchies);
2. Integration into the models of spatial relationship information between structures to semantically guide the learning.

This thesis may proceed through several stages and intermediate objectives:

- Conducting a comprehensive state of the art on structural representations in deep learning: hierarchical segmentation methods, spatial relationship description, graph neural networks, generative models with latent spaces, etc.;
- Design and prototyping of new methods directly integrating structural representations into neural network architectures;
- Experimental validations of these methods in various application contexts: image classification and segmentation, image colorization with user interaction, super-resolution, etc.;
- Dissemination of the work through the publication of scientific articles in leading conferences and journals.

Depending on the results obtained, and the motivations of the doctoral candidate, other applications based on the theoretical contributions proposed could be considered (image synthesis, segmentation of image sequences, etc.).

Collaborations with academic partners (Bordeaux Sciences Agro, Université Paris Cité, University of the Basque Country, University of Valencia) and industrial partners (Thales Mérignac) working on the same themes could also be envisaged.

Required profile

Graduated with a Master's degree or from an engineering school, with a specialization in computer science. Solid technical foundations in programming are required (Python, C), as well as knowledge in image processing and deep learning (PyTorch, TensorFlow). Proficiency in scientific English and writing skills are also very important.

To apply: send a resume, cover letter, and transcripts to the directors.

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. "SLIC superpixels compared to state-of-the-art superpixel methods". *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)* (2012).

[2] R. Deléarde, C. Kurtz, and L. Wendling. "Description and recognition of complex spatial configurations of object pairs with Force Banner 2D features". *Pattern Recognition* (2022).

[3] R. Giraud, M. Boyer, and M. Clément. "Multi-scale superpatch matching using dual superpixel descriptors". *Pattern Recognition Letters, Special Issue on Hierarchical Representations* (2019).

[4] H. Carrillo, M. Clément, and A. Bugeau. "Super-attention for exemplar-based image colorization". *Asian Conference on Computer Vision (ACCV)*. 2022.

[5] H.-D. Nguyen, M. Clément, B. Mansencal, and P. Coupé. "Towards better interpretable and generalizable AD detection using collective artificial intelligence". *Computerized Medical Imaging and Graphics* (2023).

[6] M. Clément, C. Kurtz, and L. Wendling. "Learning spatial relations and shapes for structural object description and scene recognition". *Pattern Recognition* (2018).