Editor's note: Josselin Lefèvre was awarded the Best Student Paper prize at the International Conference on Discrete Geometry and Mathematical Morphology (DGMM 2022) for his paper titled "Join, Select, and Insert: Efficient Out-of-core Algorithms for Hierarchical Segmentation Trees." For more news from DGMM 2022, follow this link to the TC18 report in January 2023 IAPR Newsletter.

~ Heydi Méndez Vázquez, EiC

Josselin Lefèvre

Josselin completed his master's degree in computer science engineering from ESIEE Paris (France) in 2021, after which he joined the Laboratoire d'Informatique Gaspard Monge (LIGM), a joint unit of Université Gustave Eiffel and Centre National de la Recherche Scientifique (CNRS), in collaboration with Thermo Fisher Scientific for his Ph.D. thesis.

His research primarily focuses on developing scalable algorithms for his hierarchical segmentation of images and data.

How did you get involved in pattern recognition research?

My journey towards pattern recognition started during my computer engineering studies at ESIEE Paris, where I got involved in research-oriented projects related to image processing. While working on a project that involved a method combining learning, hierarchical analysis, and image segmentation, I was excited by the potential of these methods in solving real-world problems. This experience inspired me to pursue research in this field and to explore opportunities for further study.

As a result, I landed an internship at Safran Aircraft Engines, where I collaborated with the Center for Mathematical Morphology on non-destructive testing by detecting regions of interest on tomographies.

This experience further solidified my interest in pattern recognition, and I knew that I wanted to pursue a Ph.D. in this field.

After seeking advice from my master's program professors, I had the opportunity to continue working in mathematical morphology in an industrial context. I began my CIFRE thesis, which involves a collaboration between the LIGM laboratory and Thermo Fisher Scientific.
This specific modality of thesis is particularly valuable for me because of its application and concrete dimension. As a scientist at Thermo Fisher Scientific, I am involved in a workflow that goes from acquisition to image analysis. The algorithm developed during my thesis will be part of the software suites provided to scientists. Thus, while pushing the limits of knowledge in my field, I also assist other fields such as life or material sciences.

What technical work have you done and what are your current and future research interests?

Image segmentation is one of the oldest problems in computer vision. Among the many existing algorithms, the watershed algorithm is a fundamental brick for numerous computer vision workflows. The core idea of this algorithm is to see the image as a topological relief. The regions correspond to catchment basins associated with the local

minima of the image. This method is often used in an interactive way by substituting the minima of the image with user-defined markers. This procedure makes it possible on the one hand to avoid over segmentation and, on the other hand, to introduce semantic information into the output segmentation.

Nowadays, advanced watershed segmentation tools are based on hierarchical representations that describe how relatively unimportant regions are progressively merged to the most significant structures of the relief according to some regional measure of importance such as depth, size, or volume of regions.

However, in recent years, electron microscopy has experienced a dramatic increase in resolution (near atomic) resulting from improvements in multiple fields, but especially better detectors and image processing methods. For example, cryogenic electron microscopy (cryo-EM) allowed for the first visualization of coronavirus and made it possible to observe proteins structures. These acquisition modalities are greatly facilitating the development of drugs. At the same time, this increase in resolution has led to the generation of images ranging in size from gigabytes to terabytes, and processing such large 3D volume data requires suitable computation strategies.

Unfortunately, the watershed algorithm and the structure we use for hierarchical analysis relies on strictly non-local properties and therefore cannot be applied to these large images that will not fit in memory in one piece. As a result, some images cannot be processed by these useful algorithms and some workflows simply cannot be used anymore.

Therefore, my research focuses on writing new algorithms for hierarchical analysis that take into account the memory constraint. Prior to my thesis, an algebraic framework was proposed based on three operations, which allow the building of a key data-structure for hierarchical analysis, the binary partition tree, in an out-of-core context. This allows loading into memory only part of the image to be processed. Subsequently, my first publication focused on effective algorithms to implement these three algebraic operations.

My current and future work consist in the extension of this out-of-core framework, allowing, for example, the calculation of attributes for connected filtering and the creation of the first out-of-core (hierarchical) watershed algorithm. It would also be interesting to redefine the algorithms, making it possible to obtain the hierarchy of quasi-flat zones or constraint connectivity in this framework.

Overall, my goal is to offer a complete out-of-core workflow in order to be able to process large volumes of data, regardless of their size.

How can the IAPR help young researchers?

I think that promoting the work of young researchers is a fantastic way to get them involved in their community. I believe this can be done mainly through the organization of conferences, workshops, and other events, where we can share our work, get feedback and advice, and learn about new research areas and techniques. Such gatherings provide great opportunities to network with other professionals in the field, including established researchers, educators, and industry professionals. I also think that it is very motivating for a young researcher to feel part of a group organized around a common field, and I think it makes it easier to project oneself into a future research career.

In my own personal case, attending the DGMM conference was a fantastic opportunity to associate faces with some of my references and to be able to chat live with many researchers. This exchange between scientists allowed me to learn about past work that I had missed and about work in progress that I would otherwise have found difficult to learn about.

By connecting young researchers to members of their community, IAPR can help them navigate the complexities of research and build successful careers.

~ Josselin Lefèvre