Foss-AI, an exploratory project using deep learning to extract paleobiological and paleoenvironmental data from fossil woods

Anne-Laure Decombeix, Verlingue Killian

IBC2024 XX International Botanical Congress Madrid Spain
July, 21st - 27th, 2024

UNIVERSITÉ DE MONTPELLIER

cnrs

AMAPlab

Foss-AI, an exploratory project using deep learning to extract paleobiological & paleoenvironmental data from fossil woods

Anne-Laure DECOMBEIX & Killian VERLINGUE

CNRS-UMR AMAP, Montpellier, France.

1. CONTEXT & OBJECTIVE OF THE PROJECT

Fossil wood is one of the most common types of plant fossil. It often **preserves fine features of the tissue** (e.g., growth-ring boudaries) **and of the individual wood cells themselves** (e.g. , wall ornamentation) [Fig.1]. This is made possible by a process called permineralization, during which water rich in minerals infiltrates and preserves in 3D the cellular structure of the wood. This process has allowed the preservation of woods from **all geological periods since the origin of wood, 400 million years ago** [1].

Fossil woods yields detailed information

- on **extinct plant physiology**, including for example defense mechanisms [2] or biomechanical and hydraulic properties of fossil stems [3, 4].
- on **past environments** since the production of wood cells by the cambium is influenced in a large part by external factors [e.g. 5].

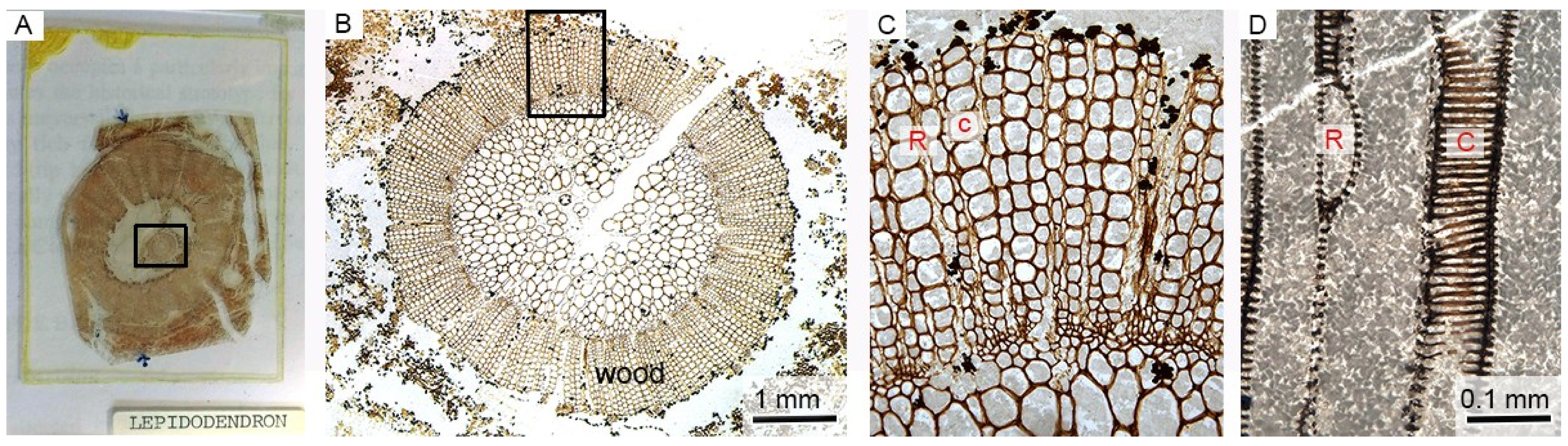

FIG. 1. Illustration showing the amount of details preserved in permineralized fossil plants, here a woody branch of a Lepidodendron tree from the Late Carboniferous, about 300 million years ago

Problem: Identifying correctly cells and structures in fossil wood is a time consuming task that requires the experience of a researcher trained in fossil plant anatomy. The small number of experts means that **fossil wood data is currently largely underexploited**

Deep learning is the part of artificial intelligence (AI) that replicates the way humans gain certain types of knowledge. It allows models to recognize patterns in data and to automate tasks that would normally require human expertise. In botany, **several recent studies have also demonstrated the possibility to automatically identify plant tissues and cells from histological thin-sections** [e.g., 6].

This project aims to use AI assisted image analysis to automate the extraction of paleobiological and paleo-environmental data from thin-sections of fossil wood [Fig.2]

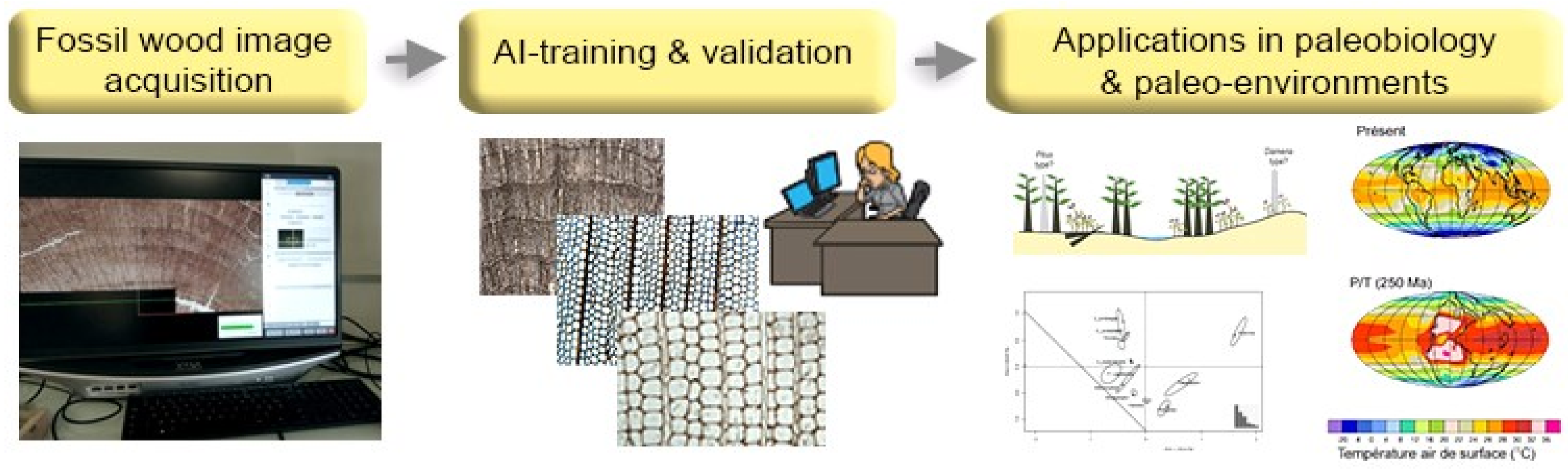

FIG. 2. General workflow of the project

2. MATERIAL & METHODS

The project focuses on Paleozoic woods because of their simple anatomy with only 2 main types of cells: tracheids and parenchyma cells Transverse thin-sections of woods from the **Devonian, Carboniferous, and Permian of several regions of the world** were selected to represent a **diversity of anatomies and types of preservation** [Fig 3].

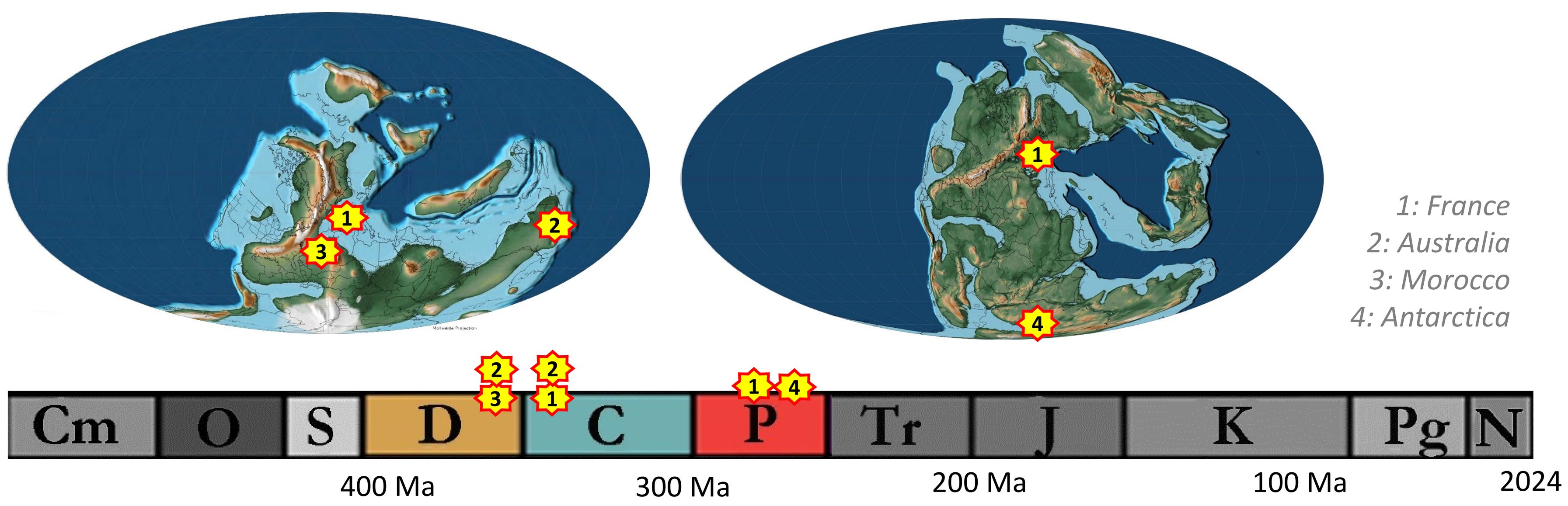

FIG. 3. Paleogeographical origin and age of the fossil woods used in the study. Maps from CR Scotese Paleomap Project.

Very high resolution images of each thin-section are produced by automatically stiching x200 photos of the whole section with a Keyence VHX-7000 digital microscope. These files can then be used as a source of images of various resolutions that can be either **(1) manually annotated with LabelMe** [7] **to train the AI** or **(2) analyzed by the AI** [Fig 4].

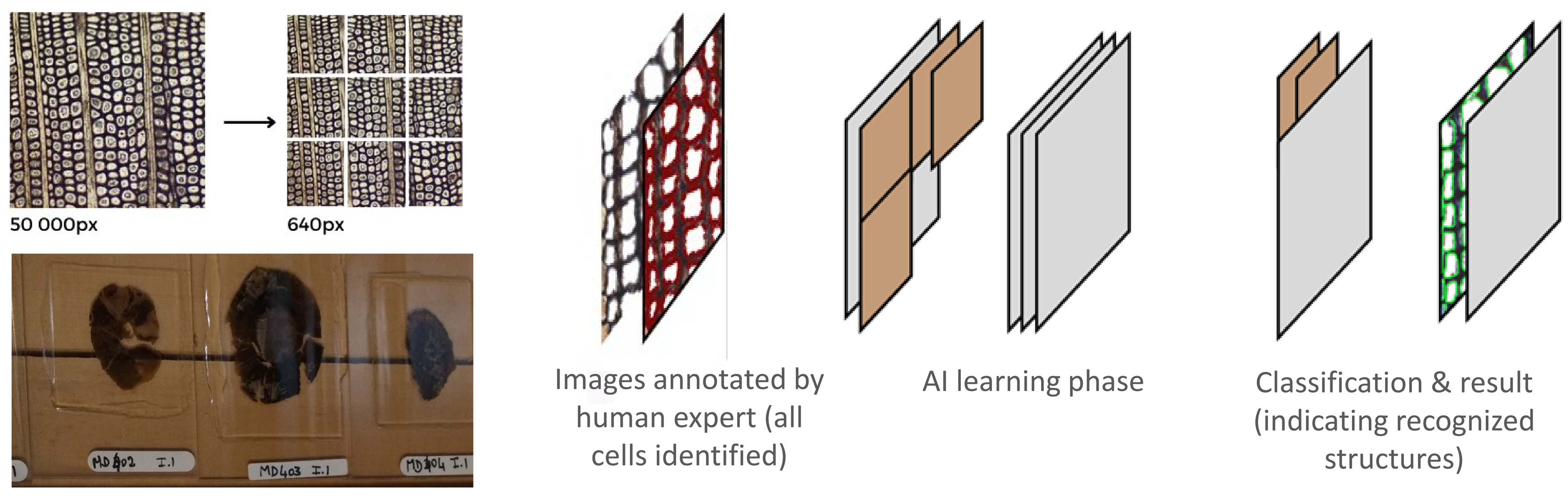

FIG. 4. AI training (illustrations by Killian Verlingue).

3. FIRST RESULTS (Spring 2024)

A total of **34 thin-sections** were imaged and annotated and some portions were analyzed with **YOLO v8** [8] , a single-shot detector that uses a convolutional neural network (CNN) to process an image. YOLO can discern multiple objects in an image and assign each a bounding box, indicating its position within the input image

Parenchyma cells of rays were rare on training images and as a result not well recognized. On the other hand, **tracheids were well recognized**, altough with some errors (e.g., double detections). This allowed extracting data on **the density of tracheids on a section** and on the **surface of individual tracheids** [Fig. 5]. This is, to our knowledge, **the first AI-assisted extraction of cell data from thin-sections of fossil wood.**

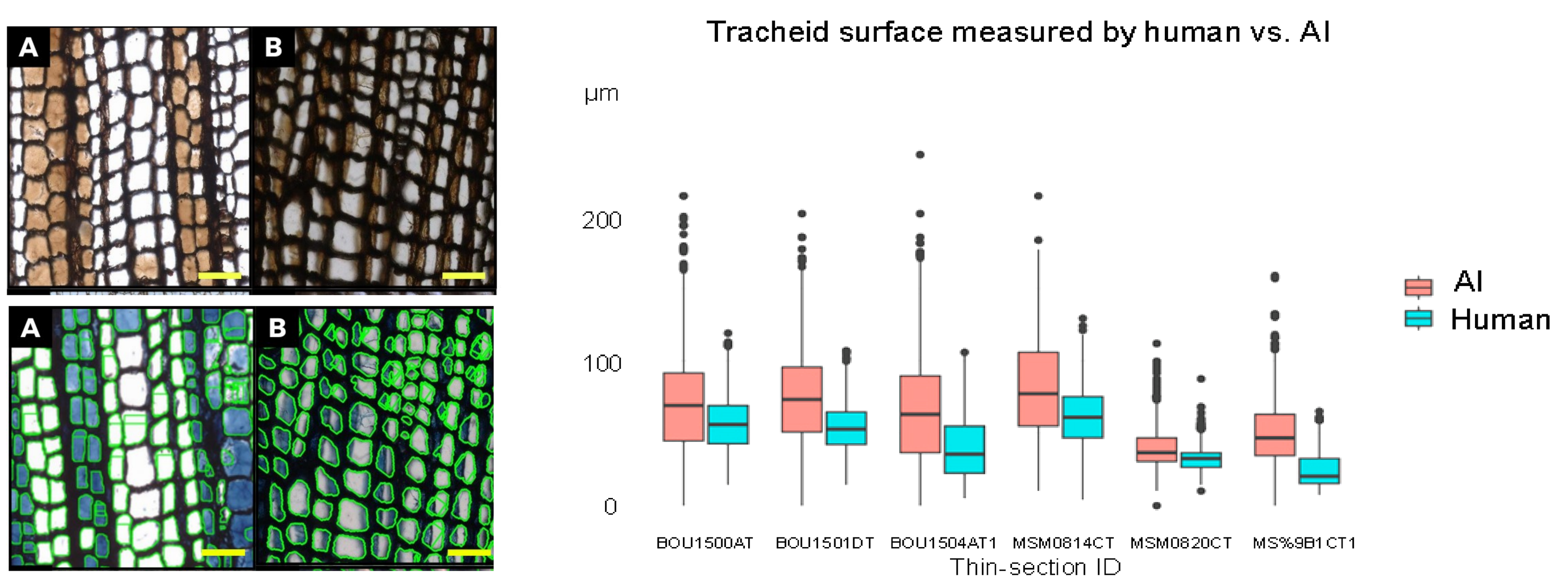

FIG. 5. Results of automatic tracheid detection and measurements on several early Carboniferous (ca. 350 Ma) woods

4. WHAT NEXT? (2024-2025)

Based on promising initial results the project will now focus on **testing new Artificial Intelligence tools**, including some already trained on other types of histological sections, in collaboration with P. Tresson (AMAP)

We will use both human and AI outputs to compare thin-sections of woods from
(1) different parts of a same fossil tree,
(2) different trees in a same assemblage (same and different taxa), and
(3) different localities of a same age

This will provide

- A large database of annotated images of Paleozoic fossil woods
- New data on wood anatomical variations within a same fossil tree & between coexisting trees
- New data on differences in wood growth and structure between paleoclimatic zones
- A framework for future large scale studies of fossil woods using AI tools

❖ REFERENCES

[1] Decombeix A.-L., Boura A. & Tomescu A.M.F. 2019. Plant hydraulic architecture through time: lessons and questions on the evolution of vascular systems. *IAWA J* 40: 387-420 [2] Decombeix A.-L., Harper C.J., Prestianni C., Durieux T., Ramel M. & Krings M. 2023. Fossil evidence of tylosis formation in Late Devonian plants. *Nature Plants* 9: 695-698 [3] Decombeix A.-L. & Rowe N.P. 2018. Functional significance of cambial development in *Vertebraria* roots: How do unusual xylem traits serve life at a high latitude? In: *Transformative Paleobotany*: M. Krings, C. Harper, N. Cuneo & G. Rothwell (eds), pp. 629-643. [4] Tanrattana M., Barczi J.-F., Decombeix A.-L., Meyer-Berthaud B. & Wilson J.P. 2019. A new approach for modelling water transport in fossil plants. *IAWA J* 40: 466-487 [5] Schweingruber F. 2007. Wood structure and environment. Berlin: Springer [6] Wu S, Al Dabagh R, Jacobsen AL, Holmlund HI & Scalzo F. 2022. Deep learning-based classification of plant xylem tissue from light micrographs. In *Advances in Visual Computing: 17th International Symposium*, San Diego, USA, pp 237–248 [7] http://labelme.csail.mit.edu/Release3.0 [8] https://docs.ultralytics.com

Acknowledgements: Stéphane FOURTIER (AMAP). This project is funded for 2024 & 2025 by the Programme d’Excellence I-SITE MUSE - Soutien à la Recherche – Projets Exploratoires