

Specimen-GT tool : Ground Truth Annotation tool for herbarium Specimen images

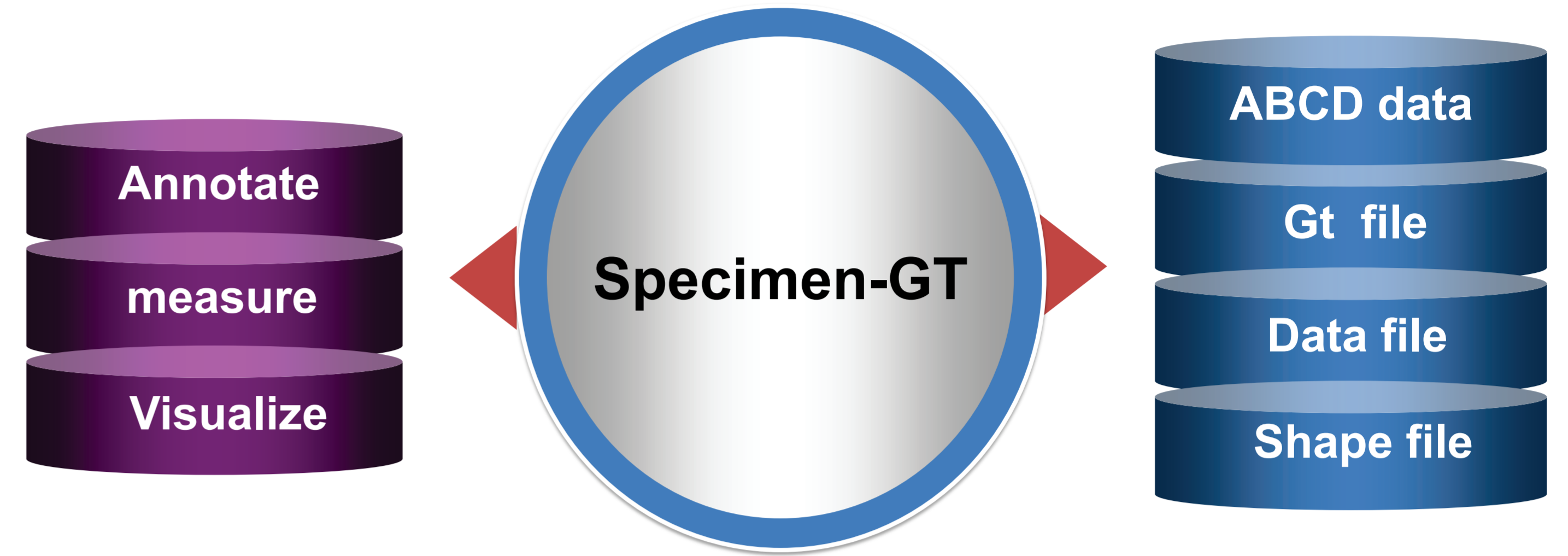
Bassem Bouaziz ¹, Rochdi Ben Ali ¹, Abdelaziz Trikki ¹, Jitendra Gaikwad ^{2,3}



¹ MIRACL/CRNS-University of Sfax, Tunisia, ² Friedrich Schiller University Jena, Germany; ³ German Centre for Integrative Biodiversity Research (iDiv) Halle-Jena-Leipzig

#1 BACKGROUND

- Datasets images annotation is an important process for supervised machine learning algorithms.
- Ground truth regions annotation are the base of the qualitative evaluation of image recognition task.
- Lack of tools dedicated for annotation of digitized herbarium specimen.
- Specimen-GT tool is a JAVA based desktop software developed for annotation of digitized herbarium specimens.
- Specimen-GT tool is one of the outcomes of the Managing Multimedia Data for Science (MAMUDS) project.



Specimen-GT Tool : Global overview

#2 OBJECTIVES

- Develop an open source tool for annotating ground truth objects contained in digitized specimens images stored in Herbarium.
- Integrate trait extraction.
- Interact with Virtual herbarium hosted in University of Vienna to add and retrieve a specimen and its associated primary data.

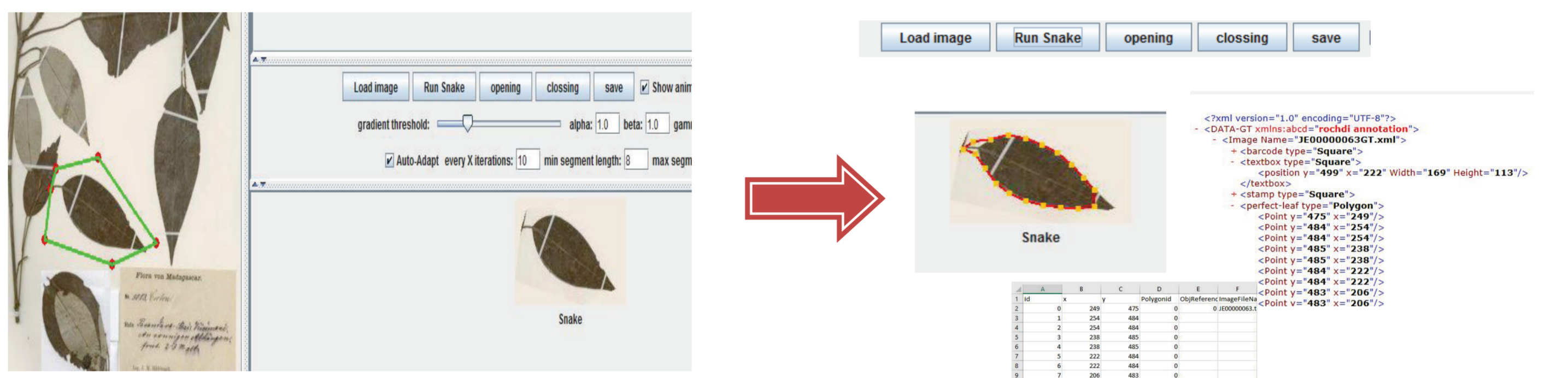


A. Graphical user interface

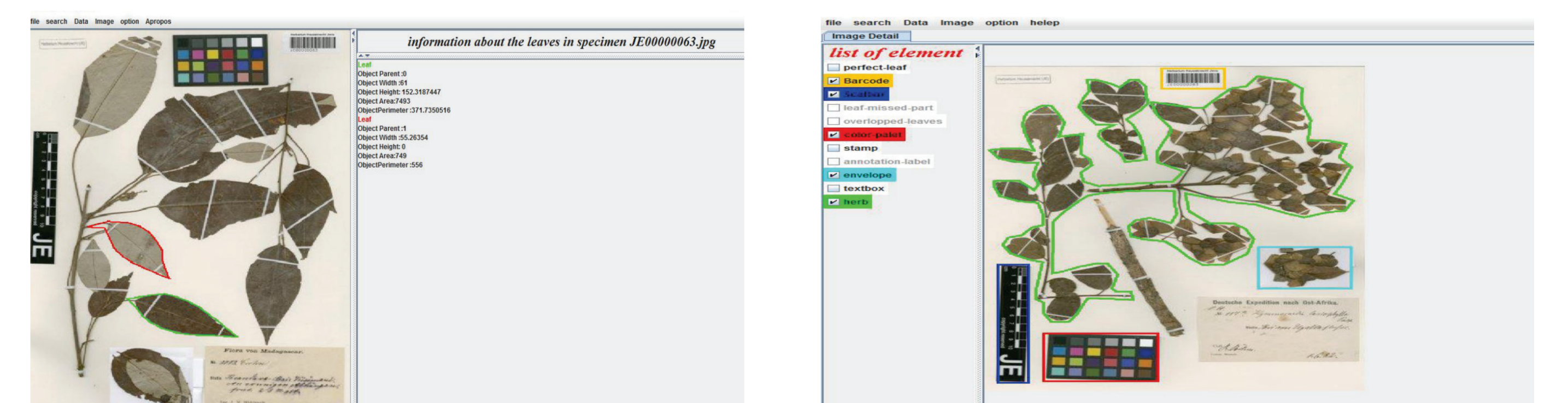
B. Extraction of GT herb Region

#3 FEATURES

- Annotation of objects in specimen images by generating customized multiple class bounding box.
- Produce multiple objects annotation in multiple format.
- Define a customized bounding element to each objects in a given specimen to be annotated (box, polygone, circle, rectangle, square, etc.) for 11 classes of objects.
- Store vector shape for evaluation of automatic object recognition in XML and CSV.
- Extract trait from specimen plant leaves.
- Provide assistance for automatic leaf shape detection based on Active snake.
- Generate Ground truth data in XML GT file, Shape File and CSV.
- Visualize batch annotated specimen images.
- Provide search function of a given specimen in offline from local repository or from Virtual Herbaria JACQ.



C. Active snake to extract bounding polygone leaf GT-data

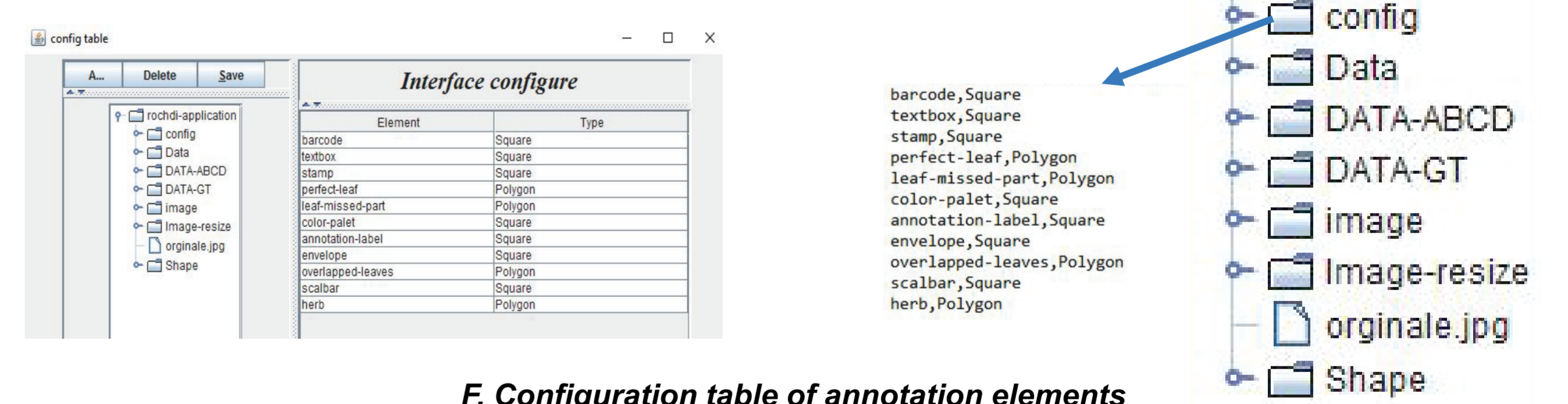


D. Visualize fine annotated leaves

E. Visualize Batch annotated specimen to check quality of GT-annotation

#4 ANNOTATED DATASET

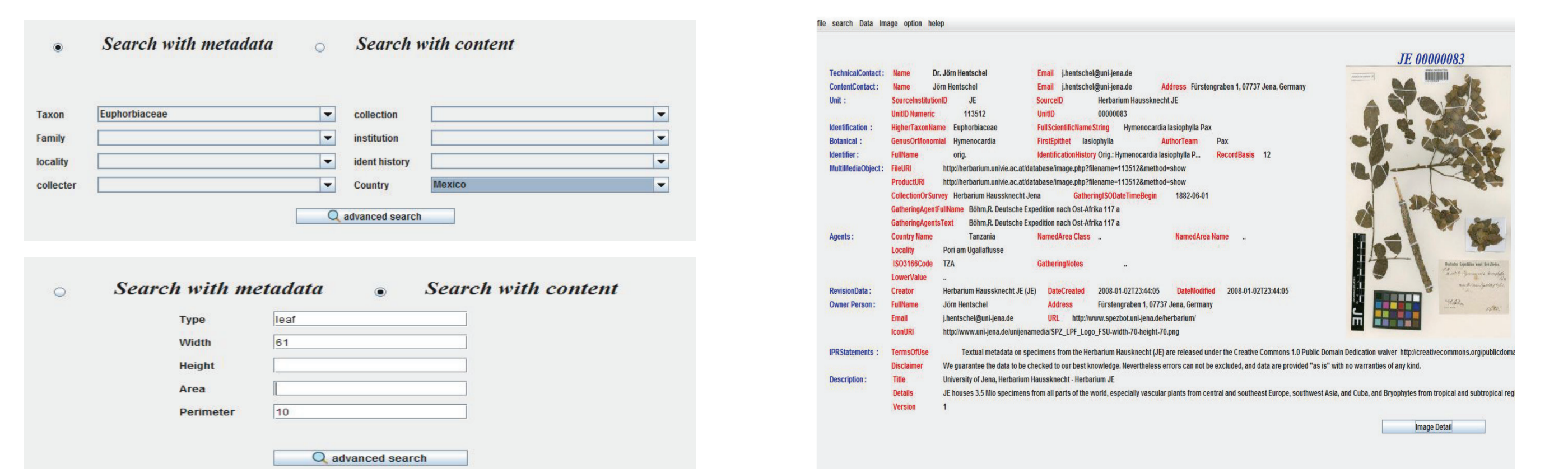
- Annotation of 7 classes of objects within digitized herbarium specimen.
- Created 4500 digitized herbarium specimen with annotations.
- Preparation of the dataset for publication.



F. Configuration table of annotation elements

#5 SIGNIFICANCE

- A research resource for annotating ground truth digitized herbarium specimens in a context of object recognition.
- Provides annotation for supervised Deep learning algorithm in the training stage.
- Enable quantitative and qualitative object recognition evaluation.
- Enrich the indexing and retrieval processes by adding extra attributes.



G. Retrieve online primary data by parsing ABCD descriptor and additional traits

Features and functionalities covered by Specimen GT-tool

Contact

Bassem.bouaziz@isms.usf.tn
jitendra.gaikwad@uni-jena.de

