

### Introduction

The Multispecies Ovary Tissue Histology Electronic Repository (MOTHER) is a web-searchable repository of digital images. Its goal is to maximize the use of ovary histology slides by openly sharing digital images and metadata about each slide. Applications include:

- Comparative studies of female reproductive biology and toxicology
- Development of cell-based computational models
- Use as training data for machine learning algorithms
- Educational resource

Figure 1 is a data flow diagram for the project. This presentation focuses on two components of MOTHER:

1. Development of ezEML+MOTHER, a web-based tool for collecting metadata and sharing histology slide images
2. Preliminary results of a machine learning algorithm to identify ovarian follicles in different stages of development (data segmentation).

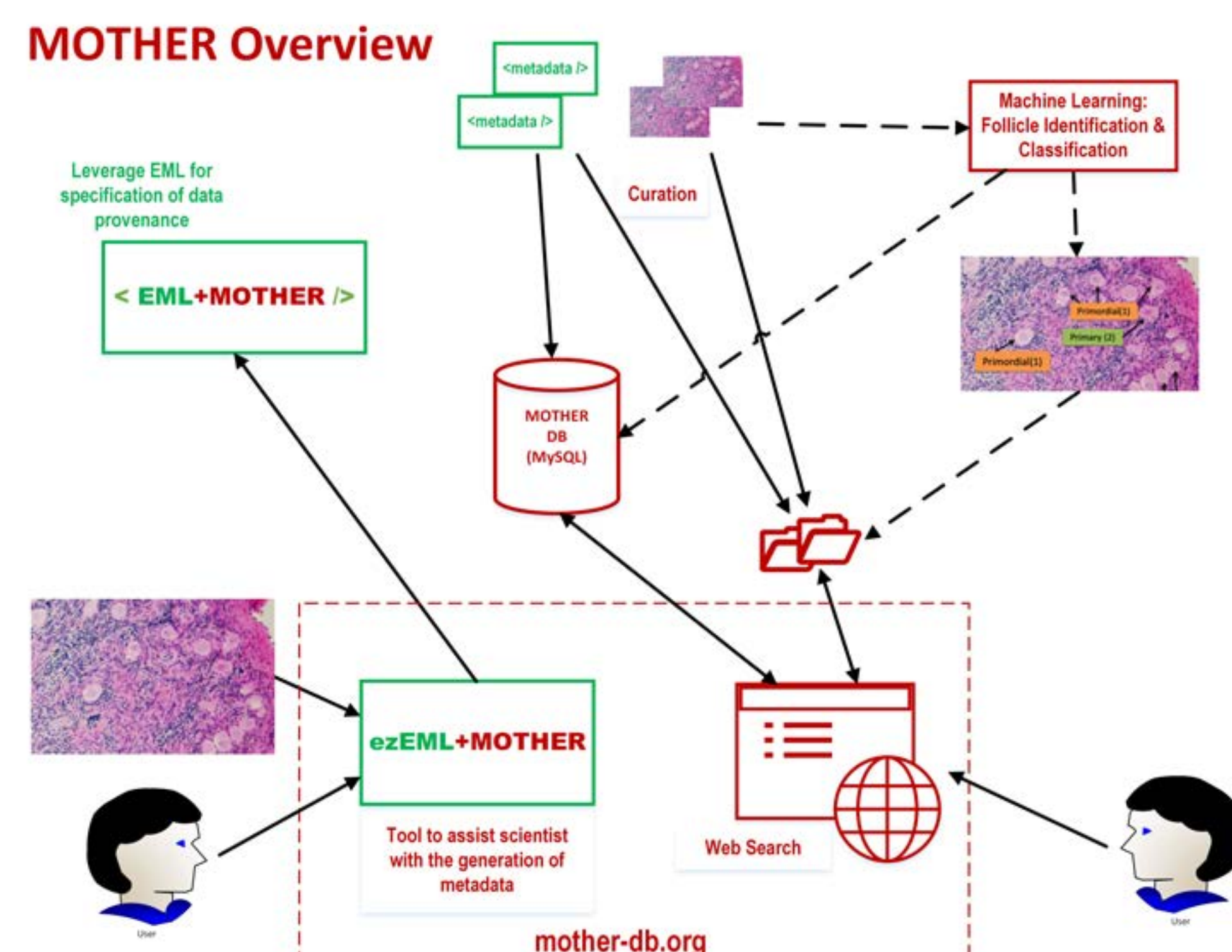


Figure 1: MOTHER Overview. EML = Ecological Metadata Language (Jones et al., 2019); DB = database. Green indicates reuse of existing tools. Red highlights components developed by this project.

### Methods

MOTHER combines a database, semi-automated transfer pipelines for images and metadata, and ezEML+MOTHER for contributing slide images and metadata.

#### ezEML+MOTHER Overview

We extended the ezEML tool for the Ecological Metadata Language (EML) that captures data provenance to include additional metadata needed for MOTHER's ovary histology images:

- Donor Animal and slide metadata
  - Species
  - Animal age
  - Reproductive state
  - Xenobiotic exposure
  - Staining technique
  - Immunohistochemistry

### Methods (Continued)

#### ezEML+MOTHER Sample Web Forms

The screenshot shows the 'Image' and 'Donor' sections of the ezEML+MOTHER web form. The 'Image' section includes fields for Title, Name (32685\_S5L-SC\_omen\_71x10), Image Type (histology), and Data Format (tif). The 'Donor' section includes fields for Donor ID, Sex (female), Years, Days, Life Stage, Specimen Sequence Number, Specimen Tissue (ovary), Ovary Position, Specimen Location, Corpus Luteum Type, Day of Cycle, Cycle Type, Stage Of Cycle, Follucular Values, and Luteal Values. There are also sections for Fixation, Stain, and Magnification.

Figure 2: ezEML+MOTHER Sample Forms - Image and Donor

#### Data Segmentation

- QuPath (Bankhead et al., 2017) used to annotate ovarian follicles
- Annotated histology images (Sluka et al., 2023a,b) serve as training data for a machine learning algorithm to identify follicles.

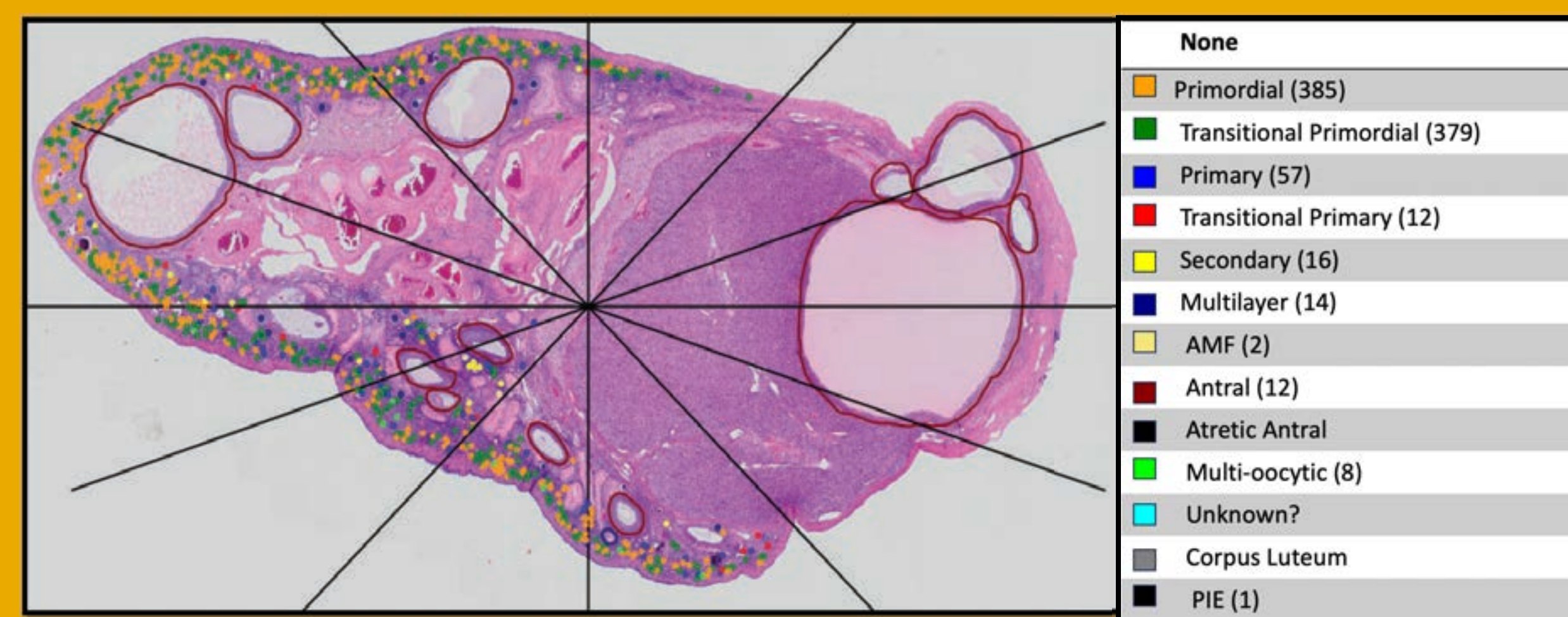


Figure 3: QuPath annotated ovary histology image. Different colors represent follicles in different stages of development

#### Machine Learning Algorithm Development

- Machine Learning Algorithm: ResNet34 (He et al., 2015) is a pre-trained, image classifying, convolutional neural network. We use a *transfer learning mode* and inherit ResNet34's basic feature detectors.
- Default Output Layer: Seven classes - Nothing, Primordial, Transitional Primordial, Primary, Transitional Primary, Secondary and Multilayer.
- Training data consisted of 18 manually annotated (Sluka et al., 2023a,b) images from 11 unique monkeys (seven Rhesus, three Cynomolgus, and one Japanese) and contained 8340 annotations.
- We augment the manual annotations by generating rotations, small offsets and mirror images of the original annotation creating a set of 12 images for each annotation.
- The transfer learning training was done using 75% of the data, with 20% used for testing and 5% used for validation. An image and its augmentations were always in the same data partition.

### Results/Conclusions

Currently, MOTHER contains histology slide images and metadata for macaques and mice available at <https://mother-db.org/>.

#### Machine Learning Preliminary Results

Figure 4 shows machine learning algorithm output for a macaque ovary histology image. The algorithm will enable development of an auto-segmentation tool to classify follicles for histology images in MOTHER.

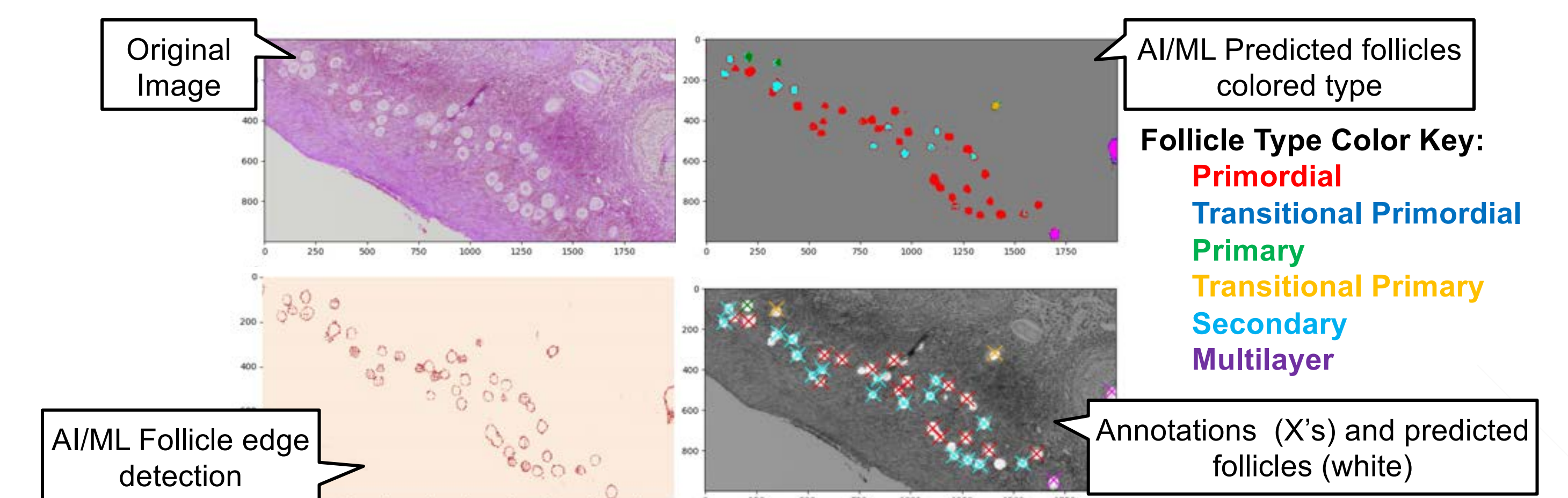


Figure 4: Typical segmentation results from the AI/ML processing pipeline. A portion of a typical ovary H&E slide is shown (upper left), predicted edge-detected features (lower left), predicted follicle types color coded by type (upper right) and an overlay of the feature predictions (white areas) and human annotations (X's)(lower right).

### How to Contribute Slides to MOTHER

1. Use the web-based ezEML+MOTHER tool to enter the slide metadata and upload your slide image(s).
2. Alternatively, create a project at osf.io in the Open Science Framework (Center for Open Science, 2020) and upload histology images with associated metadata saved from the ezEML+MOTHER tool to your OSF project.
3. Contact the MOTHER team and we will help you figure out the best way to contribute your slides. Email: [motherdbproject@gmail.com](mailto:motherdbproject@gmail.com)

### References

- Bankhead, P., et al. (2017). QuPath: Open source software for digital pathology image analysis. *Scientific Reports* 7, 16878, doi:10.1038/s41598-017-17204-5.
- Center for Open Science (2020). *Open Science Framework*, <https://osf.io>.
- He, K. et al., (2015) Deep Residual Learning for Image Recognition, arXiv:1512.03385
- Jones, M. B., et al. (2019). *Ecological Metadata Language version 2.2.0*, <https://eml.ecoinformatics.org>.
- Sluka, J., Watanabe, K. & Ding, Y. MOTHER Ovarian Follicle Annotation using QuPath (2023a). <https://hdl.handle.net/2022/29016>.
- Sluka, J., Watanabe, K., Ding, Y., Zelinski, M. & Dietrich, S. MOTHER Step-by-step supplementary files (2023b). <https://hdl.handle.net/2022/29015>.

### Acknowledgments

The authors thank all the students who worked on the project. This research was supported by: NSF DBI-2054061, GCRLE-0120, and P51 OD 011092 (DPCPSI,ORIP, NIH). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.