DeVision Spring 2025

Sydney Lawson
Lauren Bankston
Emma Kate Conner
Dominique Piccolo
Dow Draper
Carmen Neill

MATH 4997:
Vertically Integrated Research

Dr. Nadejda Drenska

April 29th, 2025

Background

The Math Consultation Clinic, coined "MC^2," aims to connect undergraduate and graduate students to real-world problems by partnering with businesses, government agencies, and research institutions. Through a capstone integrated research course, students apply advanced mathematical modeling and computational techniques to address challenges faced by clients, allowing the students to gain hands-on experience and professional skills. MC^2 enhances the educational mission of the Department of Mathematics at LSU, strengthens the ties between LSU and the Louisiana economy, and provides clients with affordable technical consulting.

This semester, the DeVision group, a group specifically catered to fulfilling the needs of *Xenopus laevis* laboratories, partnered with the Aquatic Germplasm and Genetic Resources Center at LSU. The DeVision group's main goal was to create an accurate classification system for frog embryos utilizing deep learning techniques and StarDist.

History and Goals

The DeVision project began in the summer of 2023 with the primary focus of counting frog embryos in petri dishes. The group began with a dataset of 144 images. In the summer of 2024, the project shifted towards classification, specifically labeling each embryo as either "viable" or "nonviable." Following this phase, during the fall 2024 semester, the DeVision team utilized the VGG Image Annotator (VIA). We used this software to annotate frog egg images this past semester, and it allowed us to assign more detailed labels to each embryo with the goal of creating a classification system. A key objective throughout this project has been to train a machine learning model using the StarDist Training Interface with a target accuracy of at least 90%, which was successfully attained in summer of 2024. Reaching this level of accuracy would require consistent annotation, data curation, and refinement of the classification criteria across multiple project phases.

Annotation

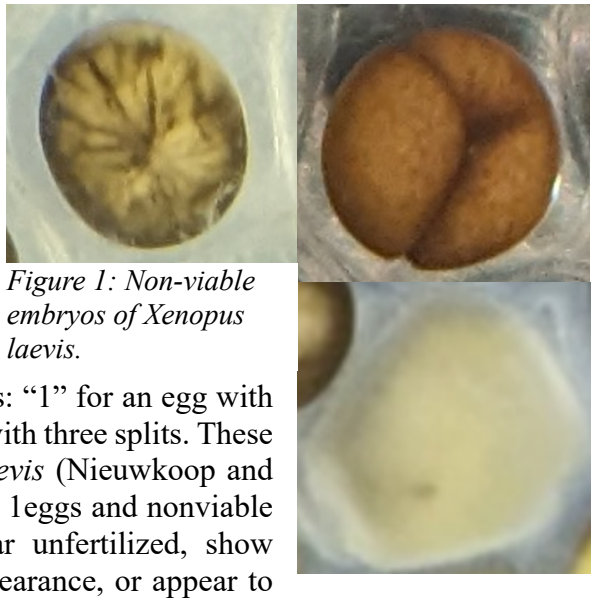

Figure 1: Non-viable embryos of Xenopus laevis.

The process of training the machine learning model begins with manual annotations. First, the frog embryo images in .jpg format are uploaded into the VGG Image Annotator (VIA), an open-source browser-based tool designed for manual image annotation. Each embryo is then manually outlined using circular regions, with partially obscured embryos intentionally excluded from the annotation process to maintain consistency. For each annotated region, the attribute used is labeled as "class_names." Then, based on the observed cell division stage, each embryo is assigned to one of the four categories: "1" for an egg with one split, "2" for an egg with two splits, and "3" for an egg with three splits. These labels reflect the NF stages 2, 3, and 4 of the *Xenopus laevis* (Nieuwkoop and Faber, 1994). The final label, "0," corresponds to NF stage 1eggs and nonviable embryos. This category includes eggs that either appear unfertilized, show abnormal features such as paleness, have an exploded appearance, or appear to

have three splits. Examples of this category can be seen in Figure 1. Consistency in annotation is critical, as poor labeling can lead to misclassification during model training.

The annotation process presented several challenges. First, it was extremely time-consuming, as each embryo had to be carefully outlined. Second, distinguishing between the embryos was sometimes difficult, especially when the images were unclear. Third, maintaining consistent labeling across different annotators was a bit challenging, requiring detailed guidelines and frequent communication between the members of the group to minimize differences in interpretation.

Once all annotations are complete, the data is exported in a .csv file format. This file includes columns such as the embryo count, region ID, shape (circle), and the corresponding category label. The ".csv" file generated for a labeled image (Figure 2A), along with the original ".jpeg" images (Figure 2B) serve as the input for training the StarDist machine learning model.

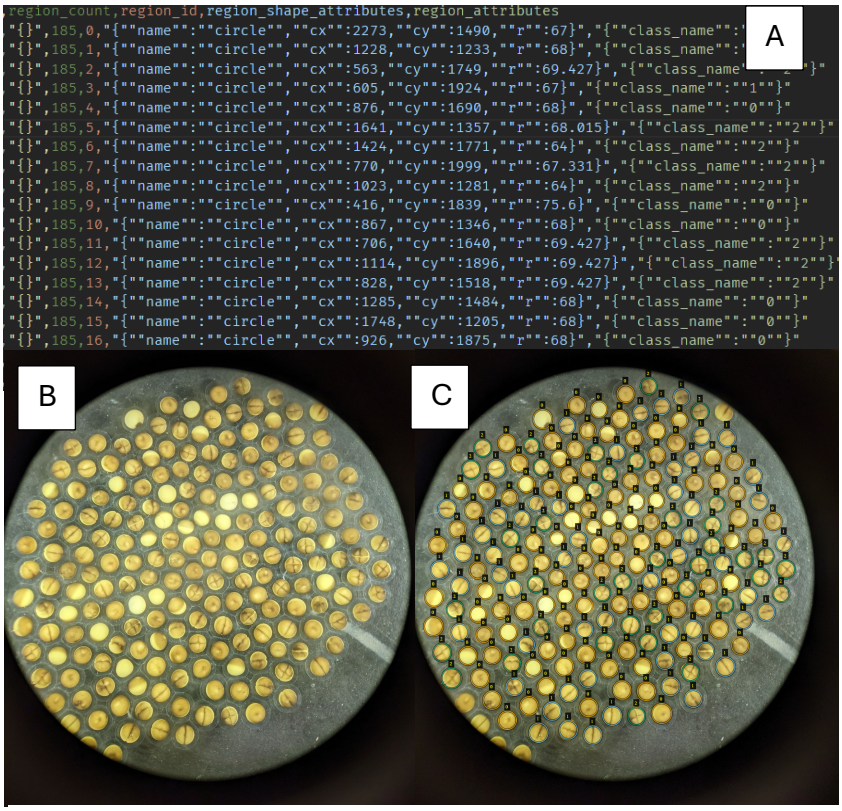

Figure 2: Result of annotation. A. ".csv" file. B. original ".jpeg" image. C. Manually annotated mask image.

The annotations are drawn onto an empty 1 channel image the same size as the color image, with values corresponding to the field "region_id" +1. This value is then mapped to the corresponding "class_name" +1 under the column "region_attributes." This can be seen in the snapshot of a ".csv" file in Figure 2A. A label mask is then generated that displays the output of the entire ".csv" file's data which can be seen in Figure 3.

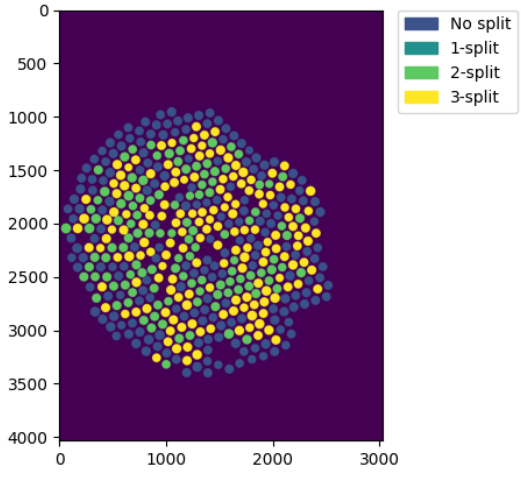

Figure 3: Label mask generated by ".csv" file

Deep Learning

Deep learning is a branch of machine learning that uses neural networks with many layers to model complex patterns in data. A neural network is made up of interconnected nodes organized into the following layers: an input layer, one or more hidden layers, and an output layer, allowing the network to learn representations of the data. During training, the network adjusts the strength of connections based on the error between its predictions and

the true values. StarDist is a deep learning model specifically designed for object detection and segmentation, particularly for images containing star-convex shapes like cells or embryos, and it can be used through a graphical user interface (GUI) that simplifies the training and prediction process. StarDist predicts the distance to an object's boundaries along 32 radial distances that are fixed and evenly spaced. In the DeVision project, we used the StarDist GUI, seen in Figure 4, to train a model on our own manually annotated frog embryo images and classify embryos based on their division stage.

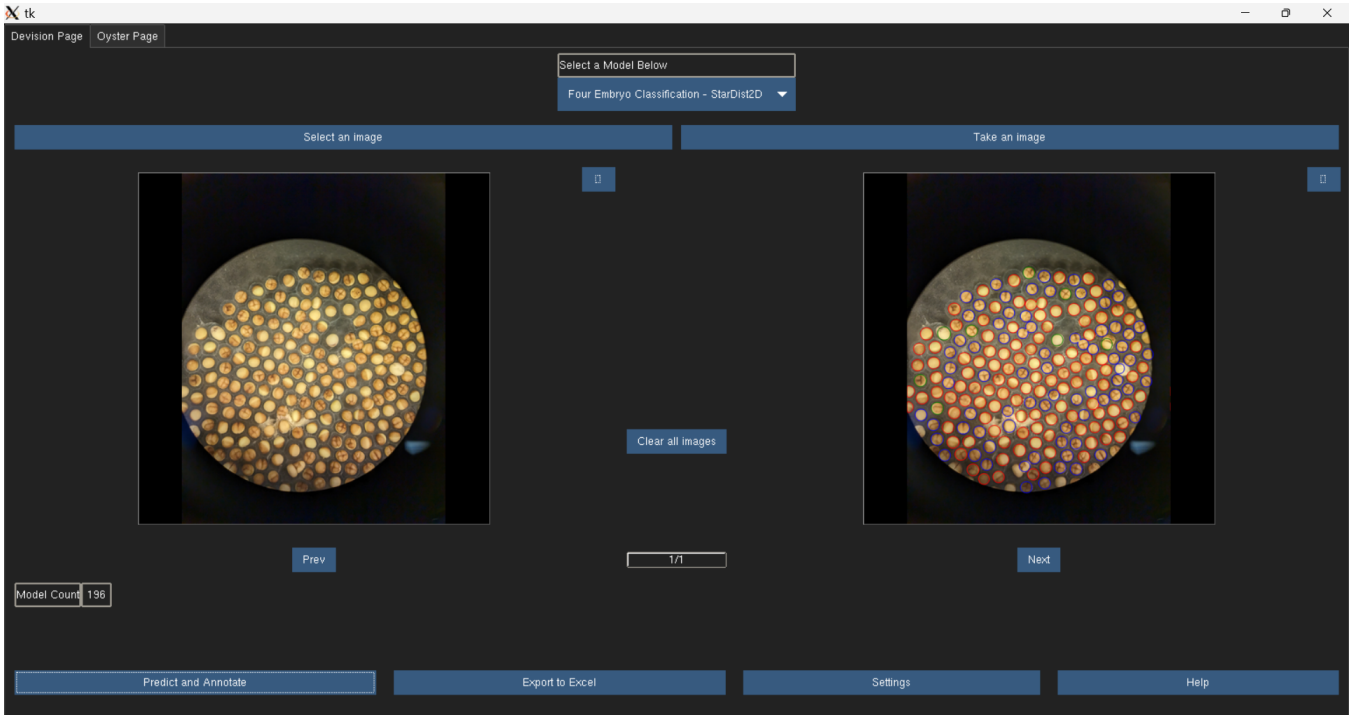

Figure 4: Graphical User Interface (GUI)

As previously mentioned, the goal for this semester's DeVision project is to develop a StarDist model that can classify embryos. To train a multiclass StarDist model with color using the Python bindings, three things are required. First a tensor- which is a higher dimensional matrix representing image data- of the input color image in the shape " $w \times h \times 3$", a tensor of the label image in the shape "$w \times h \times 1$," and a mapping between object instances and the desired predicted classes. The GUI makes it easy for the user to upload the training data, adjust model parameters, and monitor the progress of training. Throughout the process, we aimed to improve the model's accuracy by ensuring that the annotations were consistent. Once trained with several images, the model was evaluated based on its ability to correctly classify images into one of the four defined categories.

Results

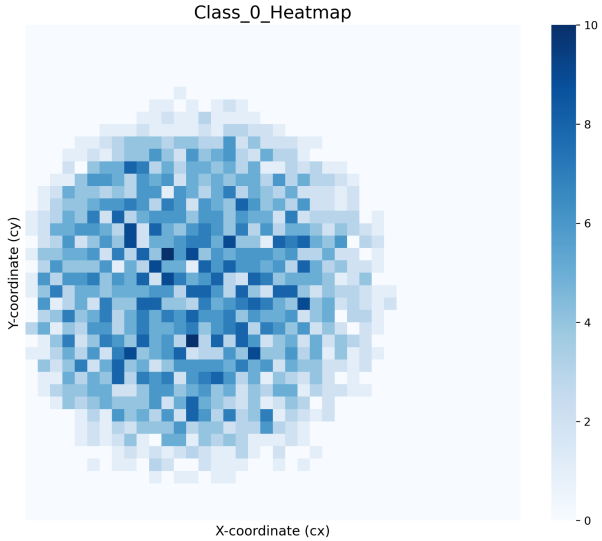

Figure 5: Density of Class 0 Embryos

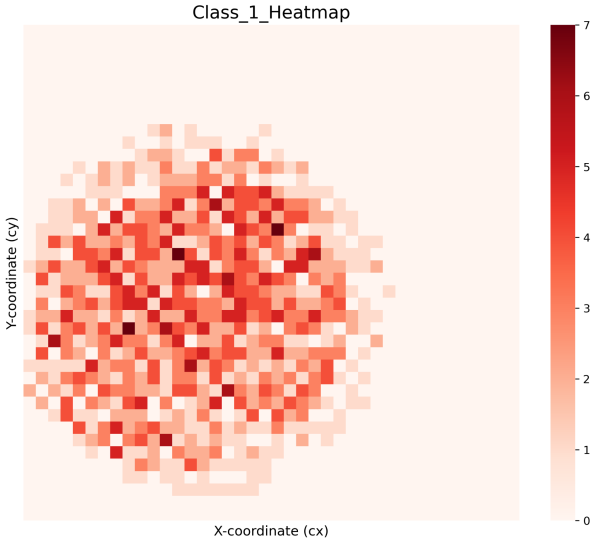

Figure 6: Density of Class 1 Embryos

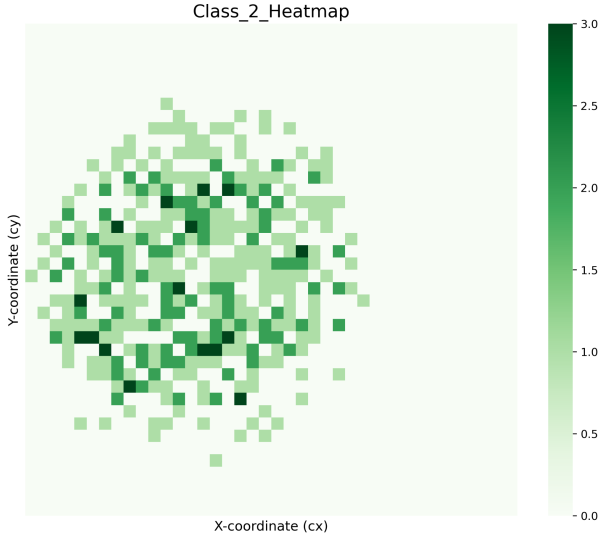

Figure 7: Density of Class 2 Embryos

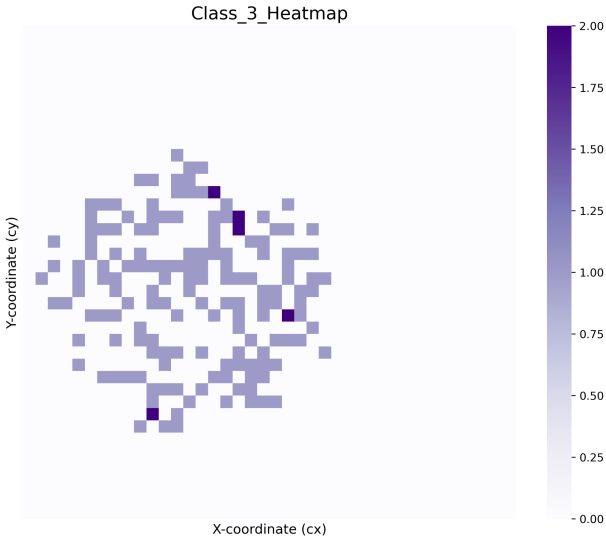

Figure 8: Density of Class 3 Embryos

The loss function, a metric of the model's accuracy, being minimized:

$$L(p, p', r, r') = L_{\text{prob}}(p, p') + p' L_{\text{dist}}(p, p', r, r')$$

where (p, r) are the predictions and (p', r') is the ground truth.

$$L_{\text{prob}}(r, r') = -r' \ln(r) - (1 - r') \log(1 - r)$$

$$L_{\text{dist}}(r, r') = \frac{1}{n} \sum_{k} |r_k - r'_k|$$

Figure 9: Loss function to determine model accuracy

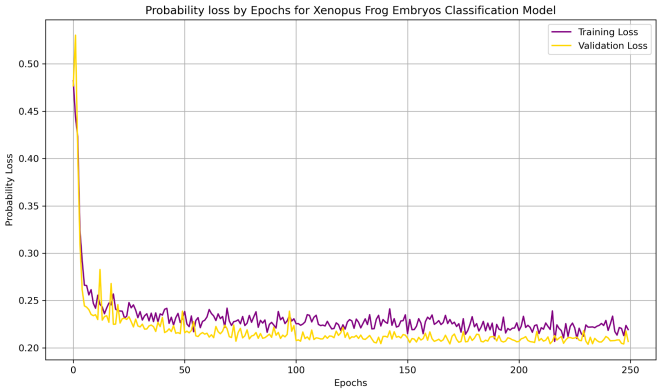

Figure 11: Probability loss graph

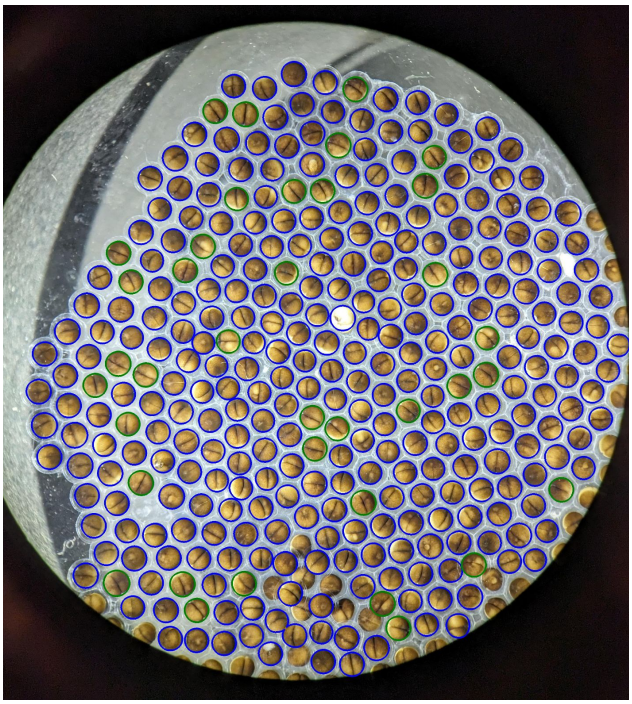

Figure 13: Successful Predicted Label Instances

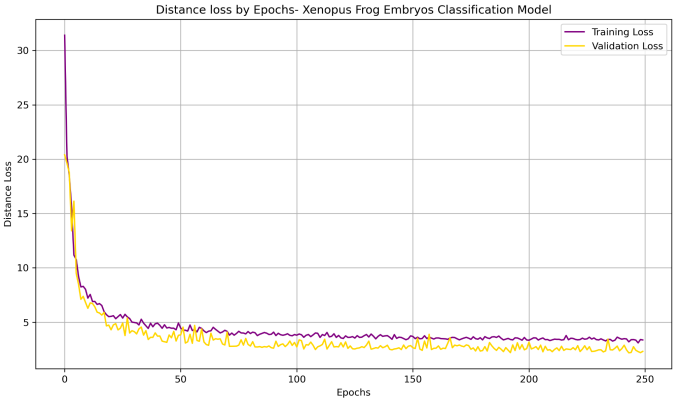

Figure 10: Distance loss graph

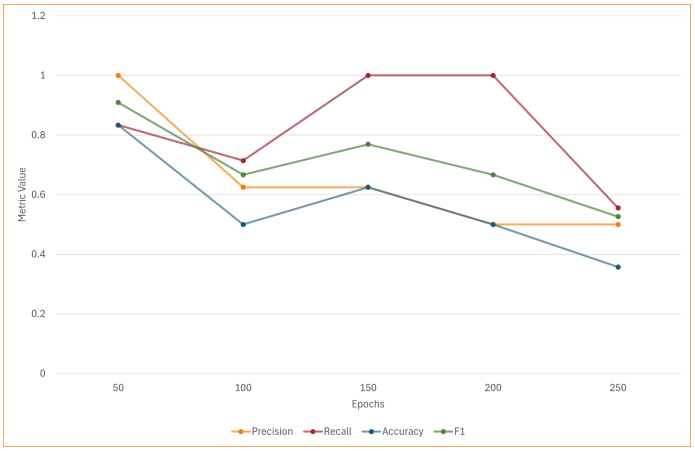

Figure 12: Metrics with IOU at different epochs

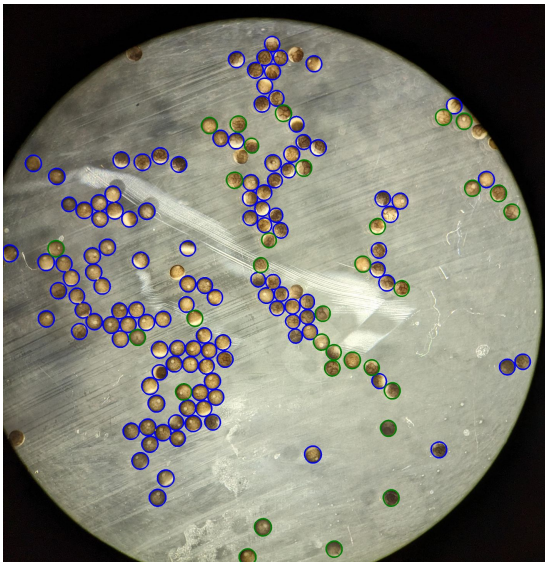

Figure 14: Unsuccessful Predicted Label Instances

Discussion

The heatmaps seen in Figures 5, 6, 7, and 8 illustrate the distribution of embryos across the different classes. Each heat map highlights where embryos of a particular stage are concentrated within the images. This visualization allows us to quickly identify trends, such as whether certain classes are more clustered or evenly spread out.

A loss function, seen in Figure 9, was used to evaluate the model's accuracy. The result of applying this function is seen in Figures 10 and 11. The distance loss for both training and validation losses started around 30 and decreased to 5 by epochs 50, 100, 150, 200, and 250. The probability loss for training and validation started at 0.5 and dropped sharply to around 0.25 by epoch 10. The training and validation curves for both losses progressed closely across all epochs, representing the model's consistent learning without major overfitting. In Figure 12, the model's performance metrics- precision, recall, accuracy, and F1 score- are presented. These values provide a comprehensive evaluation of how well the model can classify the embryos after several epochs.

Predicted label instances from the trained StarDist model are shown in Figures 13 and 14. Figure 13 displays the model's output with a classification accuracy of 59%. This result indicates that while the model was beginning to learn patterns in the data, further training and refinement were necessary to improve performance. Figure 14 shows another result, where the model performed poorly, achieving 0% precision and recall. This suggests that the model was unable to correctly classify any embryos in that instance, likely due to issues such as overfitting, insufficient training data, or inconsistencies in the annotations. These results highlight the challenges involved in training deep learning models and the importance of careful evaluation and adjustment throughout the process.

Future Work

There are several directions for future work on the DeVision project. First, expanding the training dataset by manually annotating more embryo images would likely improve model performance, particularly in underrepresented classes. Another important goal would be to continue increasing the model's accuracy, precision, and recall percentages through additional training and parameter tuning. Improving annotation consistency across all annotators will also be crucial to minimize variability in the dataset. The StarDist GUI that was used only supports equations for circles, but future teams may also explore augmenting the dataset by applying rotations, flips, or other transformations to existing images to artificially expand the training set.

We would love for researchers to one day use this model within their own laboratories. Ultimately, a successful model would significantly reduce the amount of manual labor needed for embryo classification. By automating the classification process, researchers could save valuable time and resources, allowing them to focus more on experimental design and data analysis rather than tedious manual counting and categorization. A reliable model would also help standardize embryo classification across different laboratories, minimizing human error. In the long term, this tool could be adapted for other species or similar biological studies that require early developmental stage assessment. With continued improvement, the DeVision model has the potential to become an essential resource in developmental biology research.

Conclusion

Over the course of the Spring 2025 semester, the DeVision group made significant progress toward developing a deep learning model capable of classifying frog embryos by their division stage. Beginning with manual annotation through the VGG Image Annotator, we worked to build a consistent and detailed training dataset. Despite challenges related to annotation time, image clarity, and classification consistency, we successfully prepared data for use in the StarDist GUI and trained a multiclass model. While early results showed moderate accuracy and some inconsistencies, the work established a strong foundation for future improvements. Through additional training, expansion of the dataset, and further refinement of parameters, the model's performance is expected to improve. Ultimately, this project demonstrates the potential of combining mathematical modeling, deep learning, and biological research. Our goal remains to develop a model that researchers can reliably use in their laboratories. The DeVision project continues to grow, and future work will build on this semester's accomplishments to move closer to that vision.

References

Abhishek Dutta and Andrew Zisserman. 2019. The VIA Annotation Software for Images, Audio and Video. In Proceedings of the 27th ACM International Conference on Multimedia (MM '19), October 21–25, 2019, Nice, France. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3343031.3350535.

Aquatic Germplasm and Genetic Resources Center. (n.d.). Development of a sperm cryopreservation pathway for Xenopus. AGGRC. https://aggrc.com/xenopusResearch.php

Liu, Y., Blackburn, H., Taylor, S. S., & Tiersch, T. R. (2019). Development of germplasm repositories to assist conservation of endangered fishes: Examples from small-bodied livebearing fishes. Theriogenology, 135, 138–151. https://doi.org/10.1016/j.theriogenology.2019.05.020

Schmidt, U., Weigert, M., Broaddus, C., Myers, G. (2018). Cell Detection with Star-Convex Polygons. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018.

Weigert, M., Schmidt, U., Haase, R., Sugawara, k., Myers, G. (2020). Star-convex polyhedra for 3D object detection and segmentation in microscopy- IEEE (WACV) 2020.

Weigert, M., Schmidt, U. (2022). Nuclei Instance Segmentation and Classification in Histopathology Images with Stardist - IEEE (ISBIC) 2022.

Zahn et al. (2022). Normal Table of Xenopus development: a new graphical resource. Development. 149. 10.1242/dev.200356.