Monitoring Eelgrass Wasting Disease Using Deep Learning & Crowd Wisdom

Students: Phoebe Dawkins, Olivia Graham, Brendan Rappazzo, Runzhe Yang

PI's: Carla Gomes, Drew Harvell

THE HARVELL LAB

1. Why Monitor Eelgrass Wasting Disease?

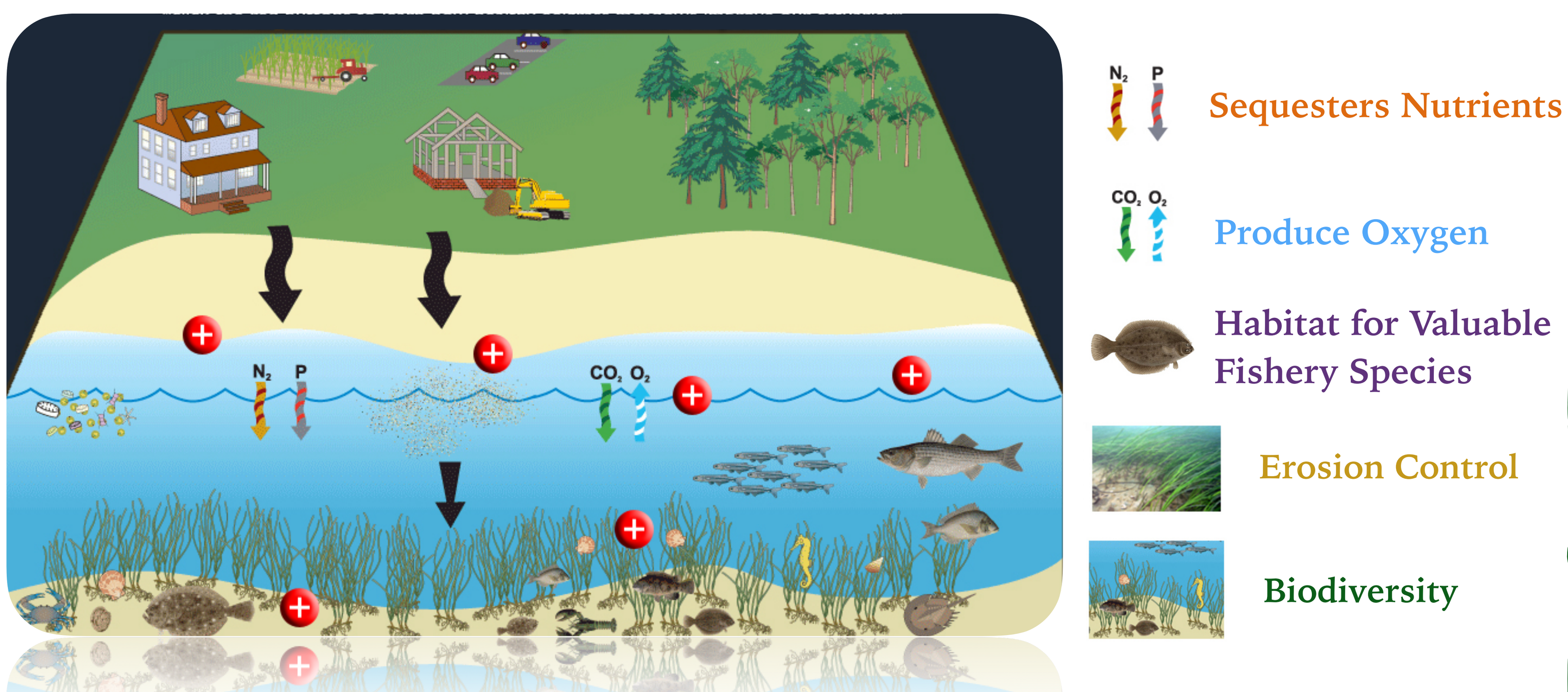
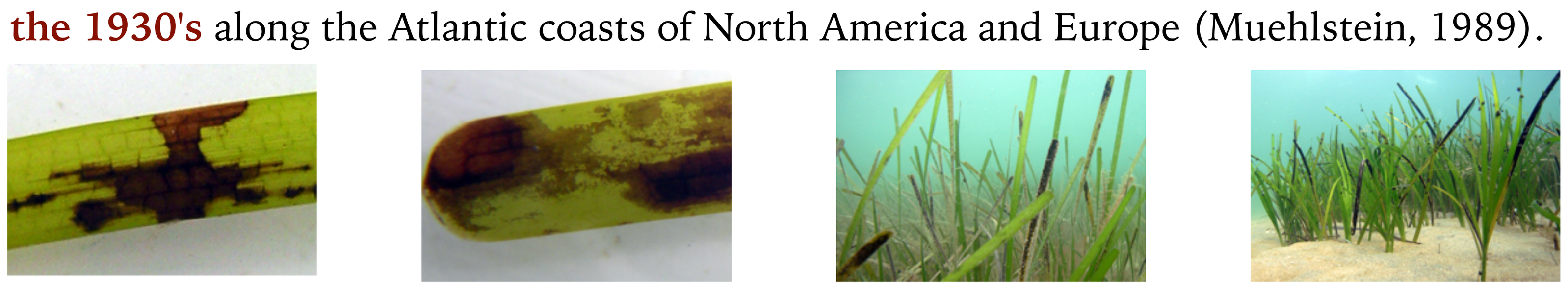


Fig.1 Eelgrass services the ecosystem.

Eelgrass Wasting Disease (EWD) can compromise the ecological services of Eelgrass, and was responsible for the disappearance of 90% of Eelgrass in the 1930's along the Atlantic coasts of North America and Europe (Muehlstein, 1989).



EWD spreads through direct blade to blade contact and can be identified by black-brown lesions, dots or streaks on the blades leading to patches, larger blackened spots, and longer streaks (Short et al., 1988; Muehlstein, 1989). Given the significant value of Eelgrass it is crucial to understand widespread patterns of disease and environmental drivers.

2. Challenges in Eelgrass Monitoring

- HUMAN ERROR:
- Disease prevalence is calculated by manually measuring lesions of collected samples. This process is a bottleneck and is prone to human error/variance.
- Artificial Intelligence (esp. Deep Learning)

- INSUFFICIENT DATA:
- Several agencies and research institutions already collect regular global surveys of Eelgrass health, but do not collect disease data.
- Crowdsourcing with Teaching

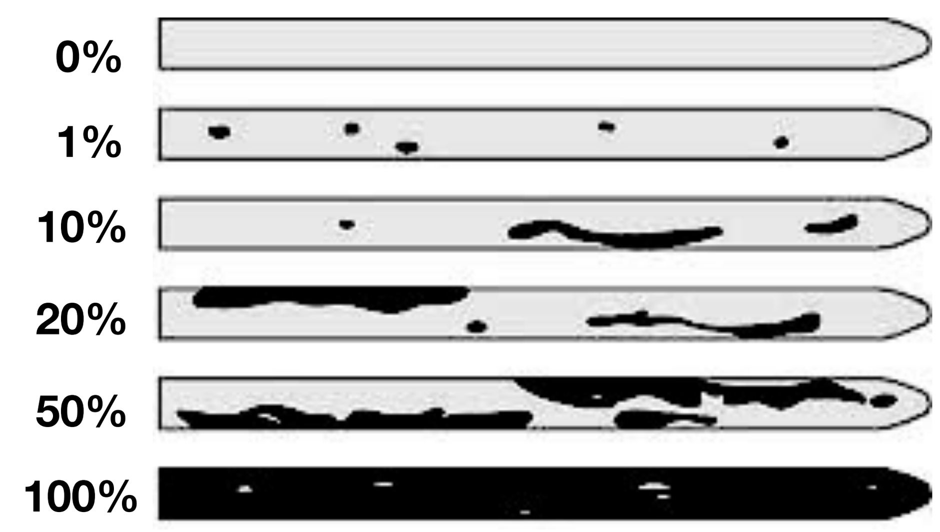


Fig.2 Eelgrass wasting disease index key.
Volunteer monitors can use this key to estimate the disease's presence on the leaves (Burdick et al. 1993)

Fig.3 APP Labeel It™
We designed an app to collect labeled data from human volunteers.

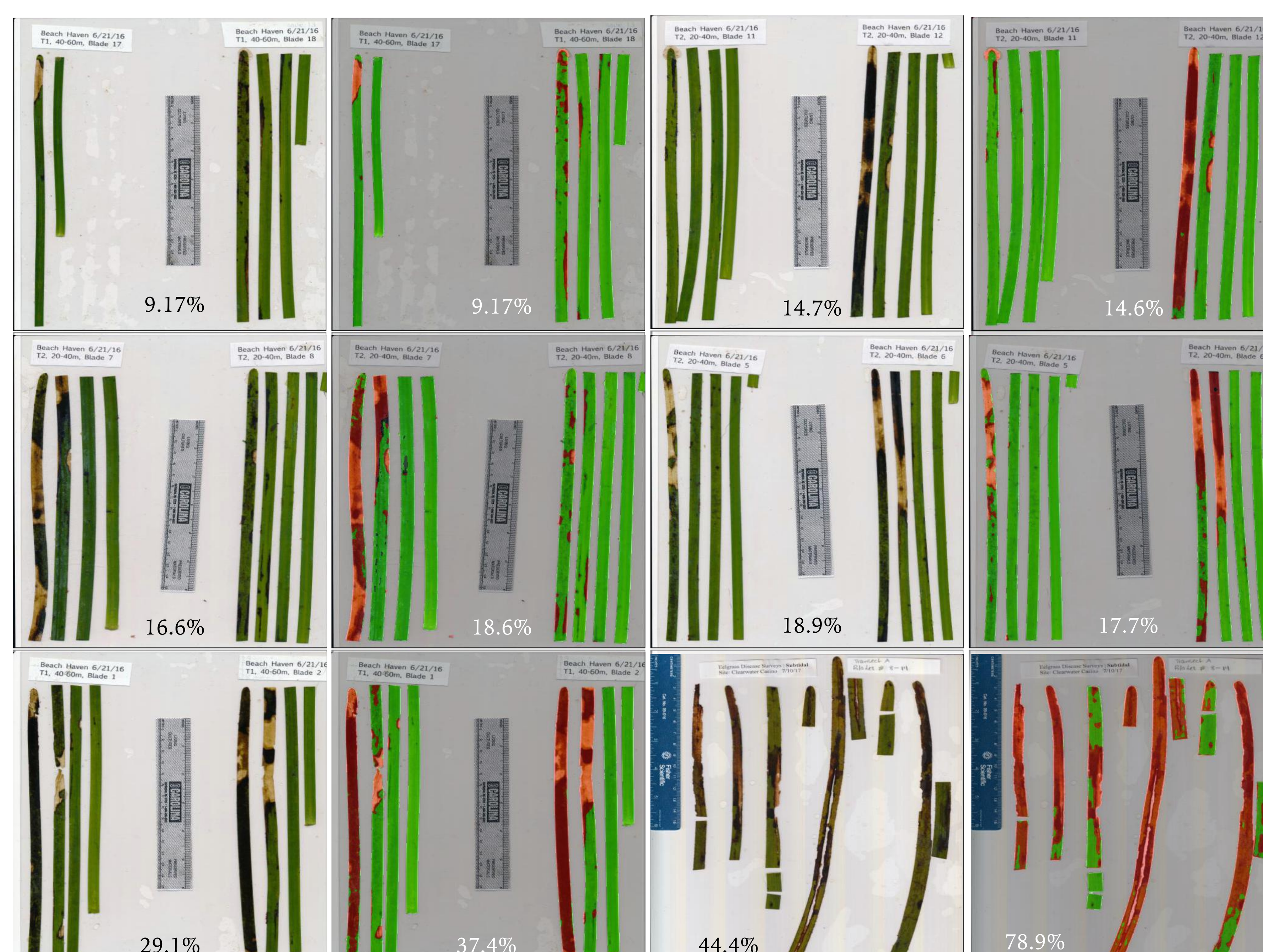


Fig 5. Results of Image Segmentation for Disease Index Calculation

3. Semantic Segmentation for Disease Index Calculation (Fully Convolutional Networks)

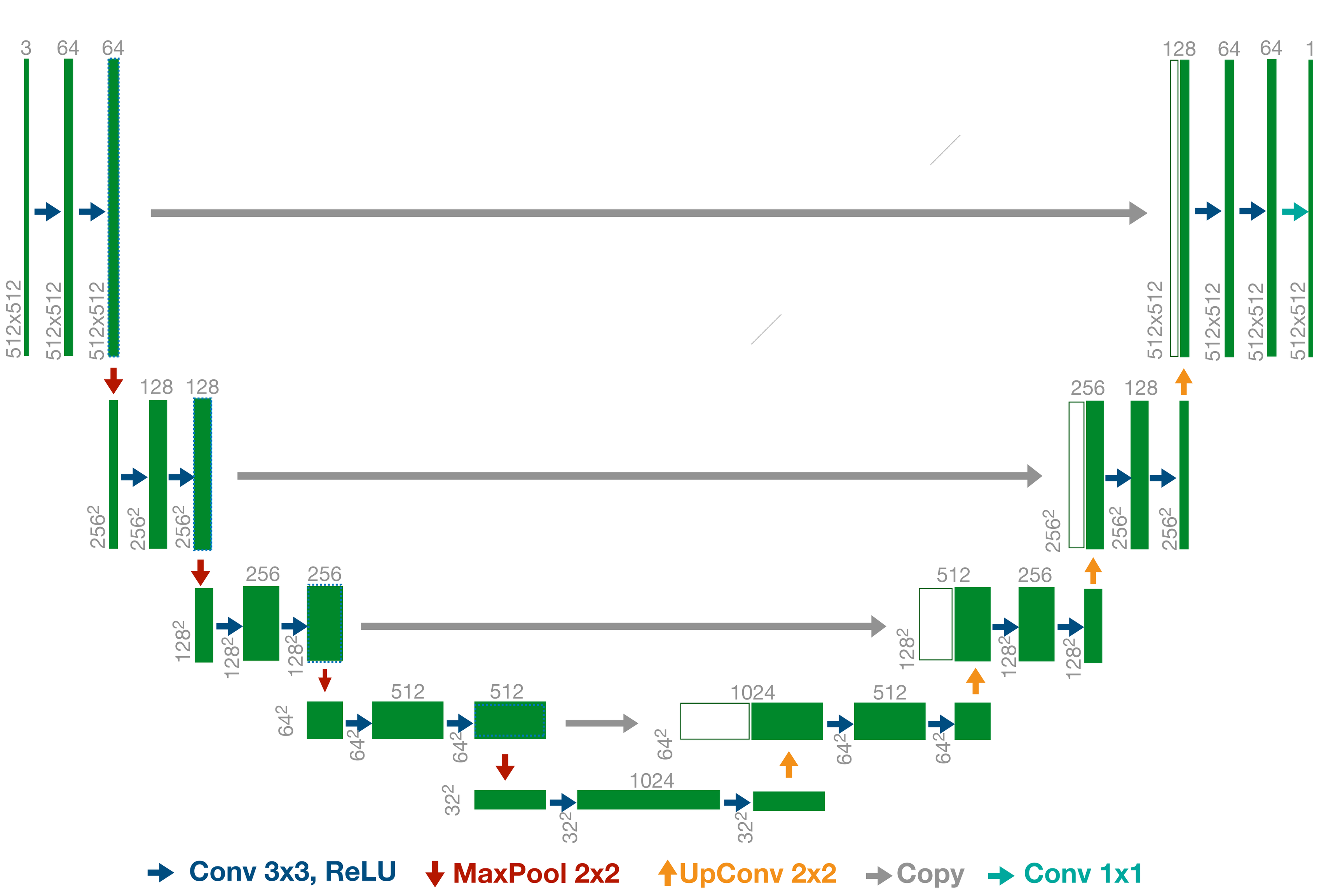


Fig.4 U-Net Architecture for Image Semantic Segmentation

We implemented the U-Net architecture (Ronneberger 2015), a kind of fully convolutional network to achieve per pixel segmentation of the background, healthy plant material and lesioned plant material. The inputs of this model are scans taken by marine biology researchers and the outputs are predicted label masks. We used a modified cross entropy loss to emphasize importance of lesioned areas, despite being represented less than the background or healthy material. We trained the model on a very small dataset (~100 labels).

4. Future Work

[Improving Model Accuracy]
- Continue tuning hyper parameters, and training data manipulation for improved accuracy.
- Try different networks, and combine traditional segmentation methods with deep learning methods.

[Promoting Application]
- Develop application that uses trained model to analyze 1000's of scans, and export relevant information.
- Develop temporal-spatial disease model to predict and prevent spread of disease.

[Improving Data Collection Process]
- Overcome data scarcity via reliable crowdsourcing, such as pedagogical value-aligned crowdsourcing.

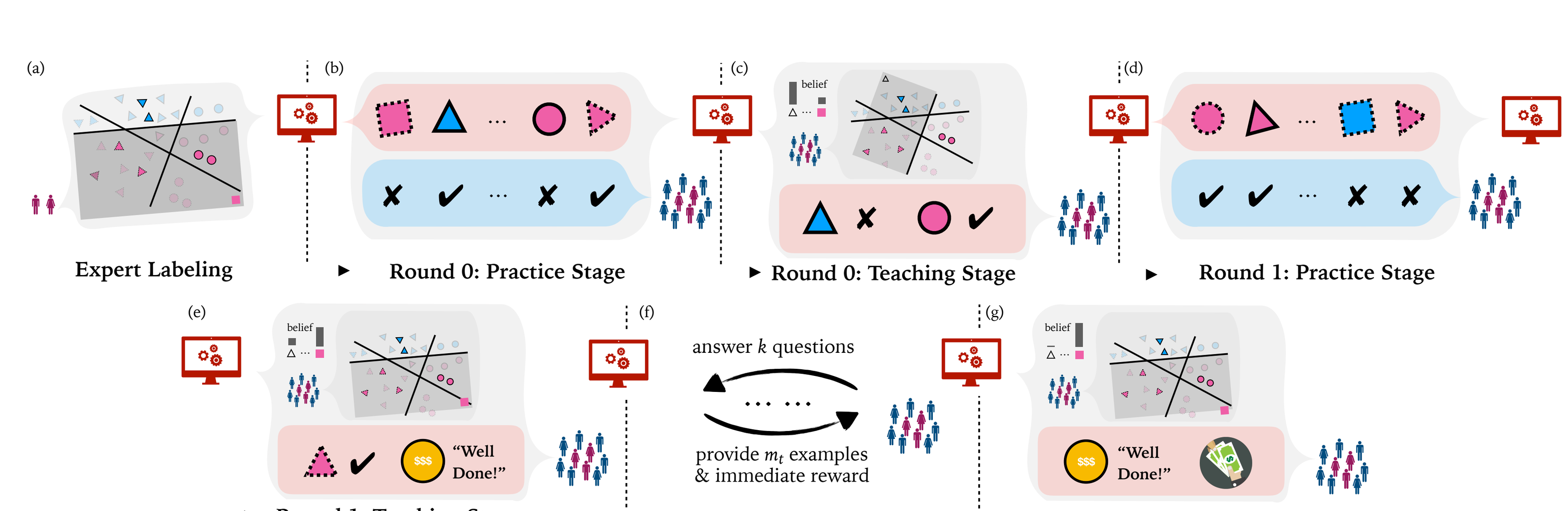


Fig 6. The basic process of pedagogical value-aligned crowdsourcing.

5. Acknowledgments

We are thankful to the National ScienceFoundation, CompSustNet and Cornell University,