Seminar Presentation: Mosquito Vision

EN.601.456 Computer Integrated Surgery II

Project 5

Team: Akash Chaurasia, **Alan Lai**, Parth Vora

Mentors: Balazs Vagvolgyi, Dr. Russell Taylor

Project Overview

Goal: To create a ROS-integrated computer vision system for mosquito detection and keypoint identification to guide an automated mosquito dissection robotic system for live malaria vaccine production.

Team Members:

Akash Chaurasia, Alan Lai, Parth Vora

Mentors:

Balazs Vagvolgyi, Dr. Russell Taylor

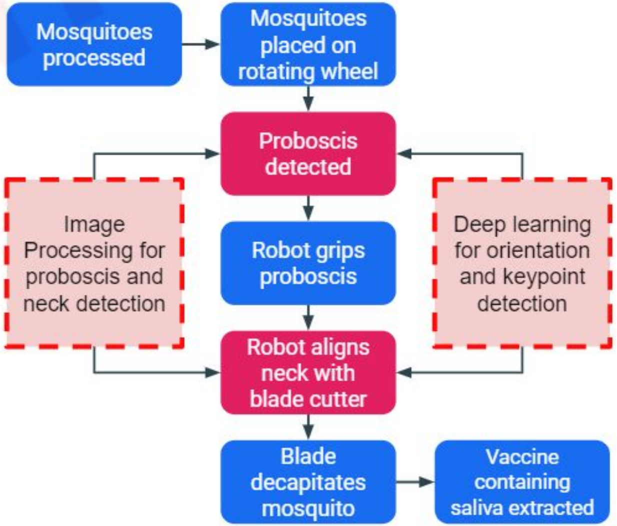


Figure 1. Process flow diagram for mosquito keypoint localization

Paper Selection and Relevance

Multi-Mosquito Object Detection and 2D Pose Estimation for Automation of PfSPZ Malaria Vaccine Production

Hongtao Wu, Jiteng Mu, Ting Da, Mengdi Xu, Russell H. Taylor, *Life Fellow, IEEE*, Iulian Iordachita, *Senior Member, IEEE*, and Gregory S. Chirikjian, *Fellow, IEEE*

- Usage of deep learning to tackle Mosquito keypoint detection
    - Relevant keypoints to workflow
    - Clear description of deep learning architectures
- Comparison of deep learning with image processing techniques

Problem Statement

- Computer vision support for automated Mosquito Micro-dissection System (MMS) to identify location to hold and manipulate mosquito (proboscis) and location to decapitate mosquito (neck)
- Comparison of deep learning and image processing approaches

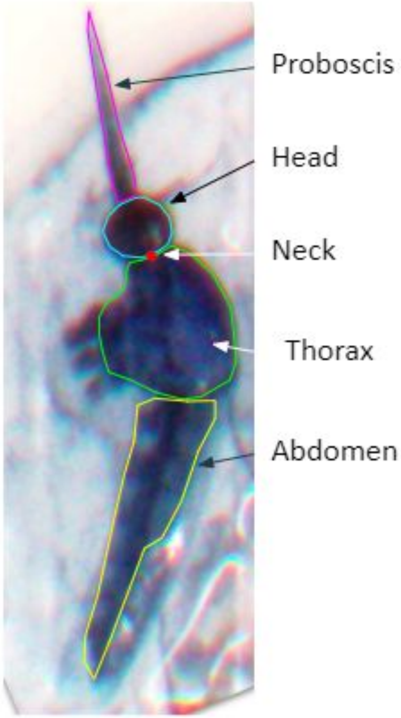


Figure 2. Parts of a Mosquito

Key Results and Significance

The deep learning approach outperforms the image processing approach in both mosquito detection and body part detection

- Deep learning approach is more robust and versatile
- Image processing has a faster runtime
- Mask R-CNN and DeeperCut effective

Approach	Deep Learning	Image Processing
Detection mAP (IoU>0.5)	0.97	0.80
Detection Recall	0.97	0.90
Head Position RMSE	1.61 pixels	2.70 pixels
Proboscis Orientation Error	14.3°	24.7°
Processing Speed	2.5 fps	20 fps

Figure 3. Table from Wu et al. comparing results of Deep Learning and Image Processing approaches (Wu)

Background

- Malaria infected more than 200 million people and caused more than $12B USD loss globally in 2017
- Sanaria has developed a PfSPZ vaccine shown to be very effective in clinical trials
- Previous workflow involved laborious manual processing to extract vaccine

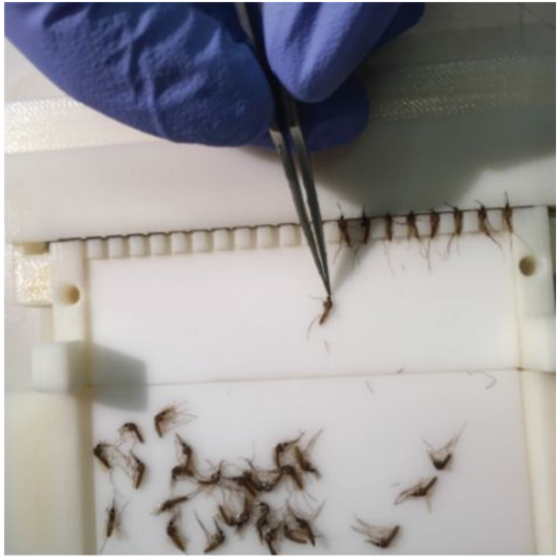


Figure 4. Semi-automatic Mosquito Micro-dissection System (saMMS) (Schrum)

Background

- Proposed automatic MMS

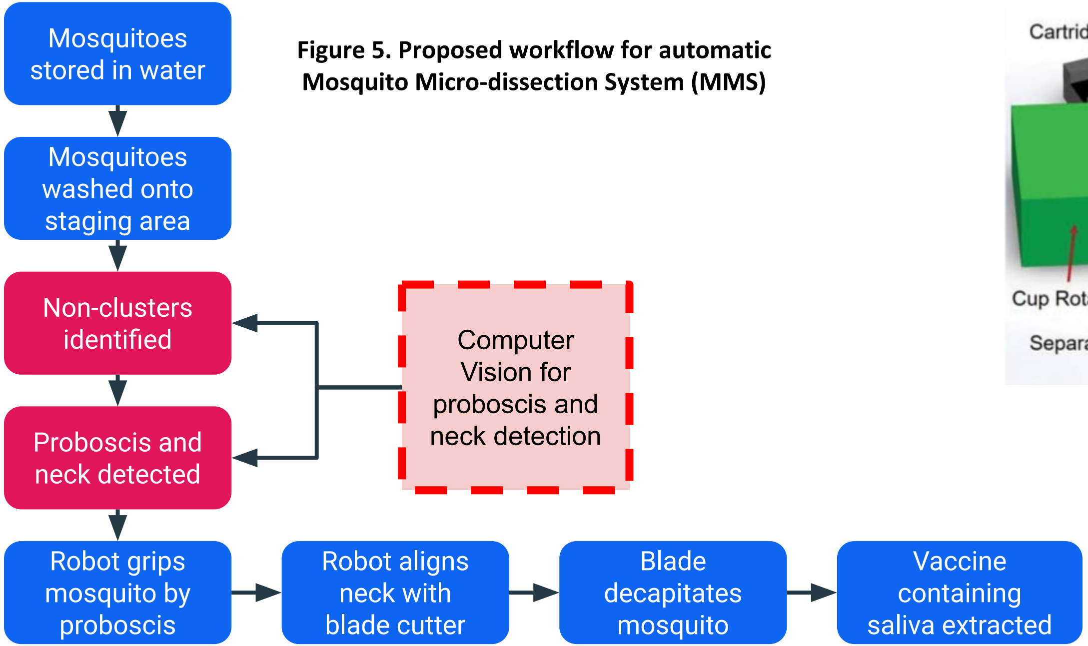


Figure 5. Proposed workflow for automatic Mosquito Micro-dissection System (MMS)

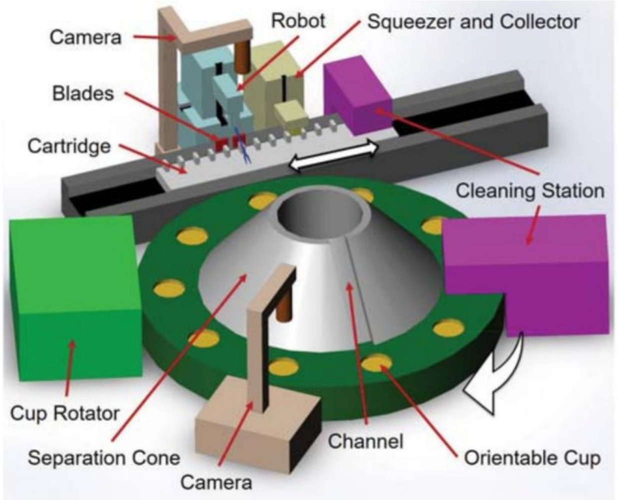


Figure 6. Proposed MMS (Phalen)

Experimental Approach: Dataset

- Manual placement and automatic water transfer
- Varied water flow, scale, and luminance conditions
- 1460 images total, all manually labeled
- 7:2:1 split for training, validation, testing

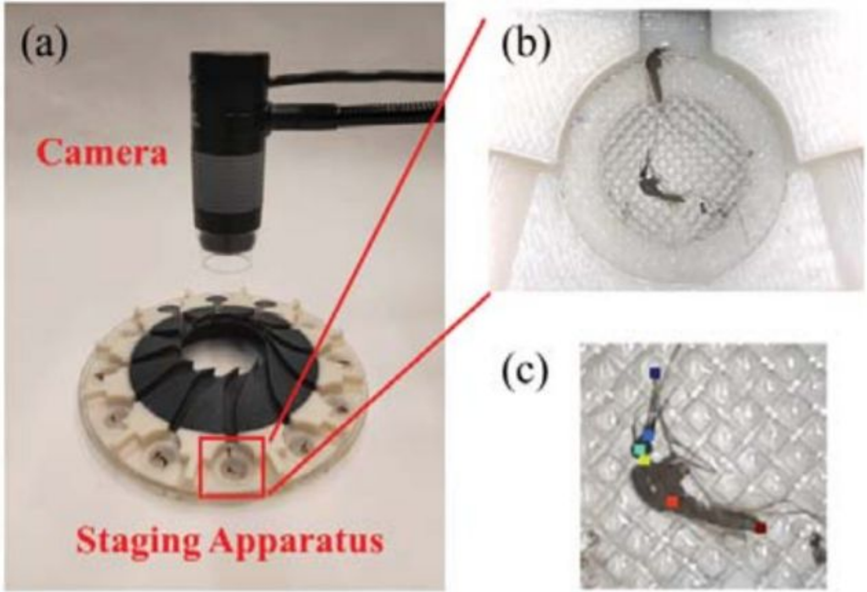


Figure 7. Dataset collection. (a) camera staging apparatus for MMS (b) image taken from camera (c) keypoints labeled on mosquito (Wu)

Experimental Approach: Deep Learning

- Transfer learning (Mask R-CNN on COCO, DeeperCut on ImageNet)

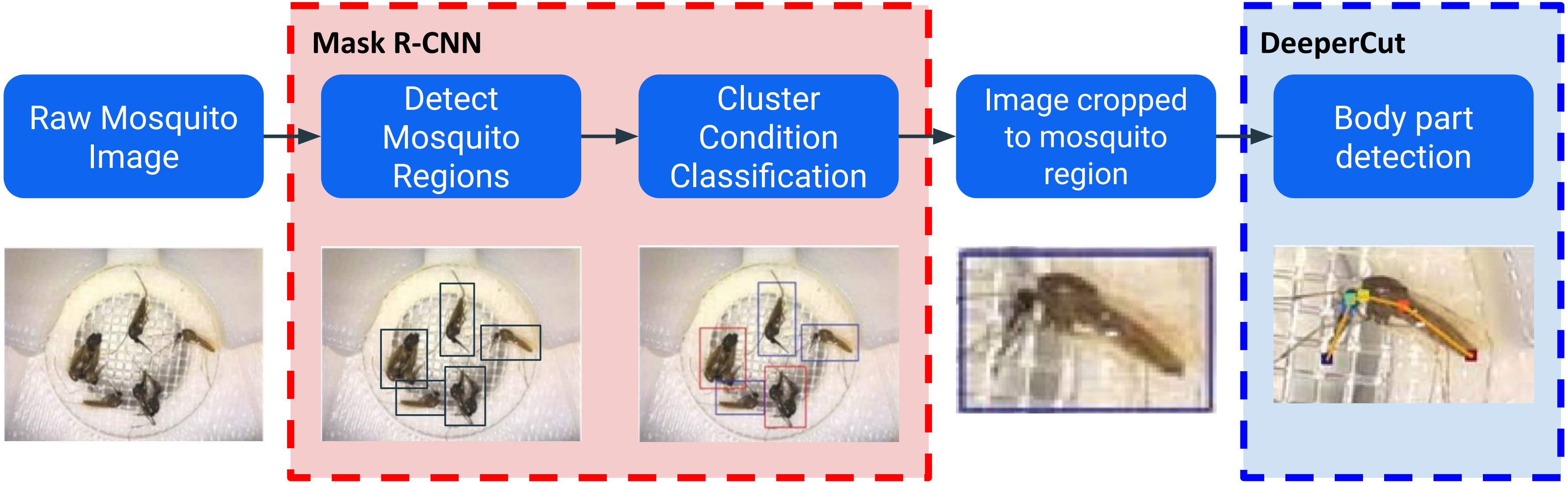


Figure 8. Process workflow for deep learning approach (Images from Wu et al.)

Images from: Wu, H., Mu, J., Da, T., Xu, M., Taylor, R. H., Iordachita, I., & Chirikjian, G. S. (2019). Multi-mosquito object detection and 2d pose estimation for automation of PfSPZ malaria vaccine production. In 2019 IEEE 15th International Conference on Automation Science and Engineering, CASE 2019 (pp. 411-417). [8842953] (IEEE International Conference on Automation Science and Engineering; Vol. 2019-August). IEEE Computer Society. https://doi.org/10.1109/COASE.2019.8842953

Experimental Approach: Image Processing

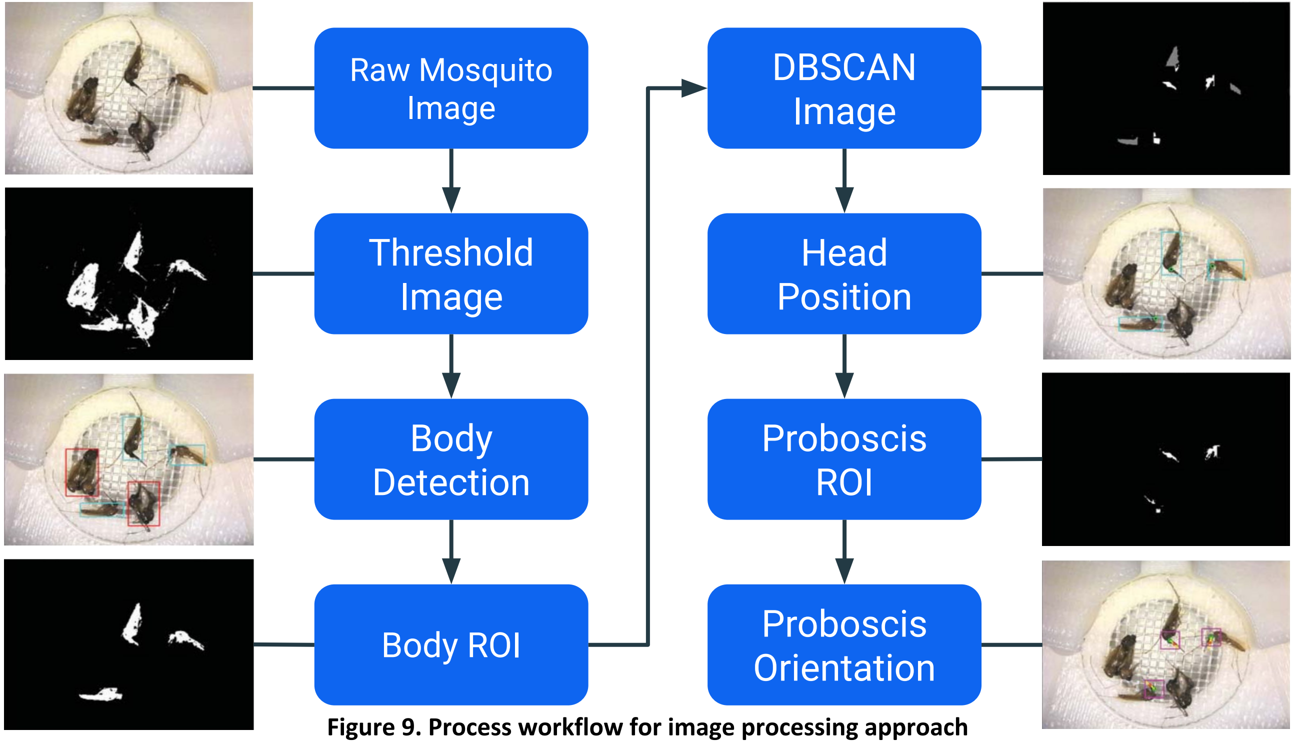


Figure 9. Process workflow for image processing approach (Images from Wu et al.)

Images from: Wu, H., Mu, J., Da, T., Xu, M., Taylor, R. H., Iordachita, I., & Chirikjian, G. S. (2019). Multi-mosquito object detection and 2d pose estimation for automation of PfSPZ malaria vaccine production. In 2019 IEEE 15th International Conference on Automation Science and Engineering, CASE 2019 (pp. 411-417). [8842953] (IEEE International Conference on Automation Science and Engineering; Vol. 2019-August). IEEE Computer Society. https://doi.org/10.1109/COASE.2019.8842953

Experimental Approach: Final Output

Figure 10. Output of Deep Learning Approach (Wu)

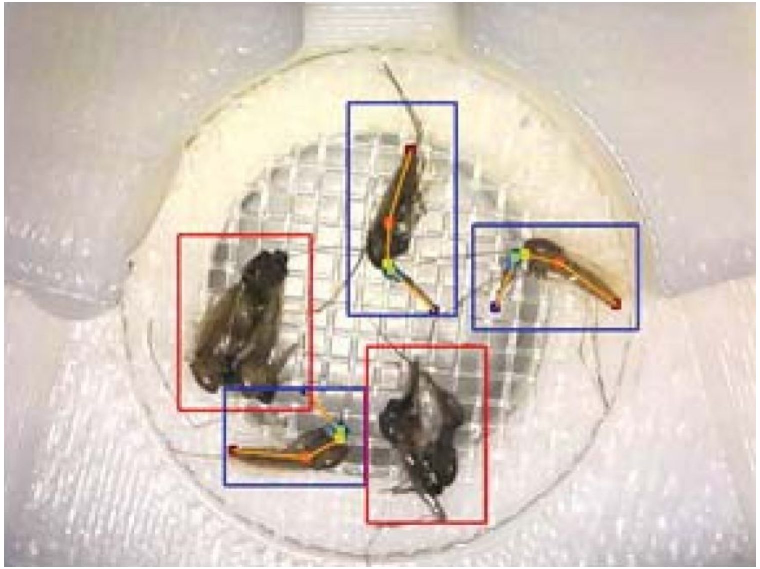
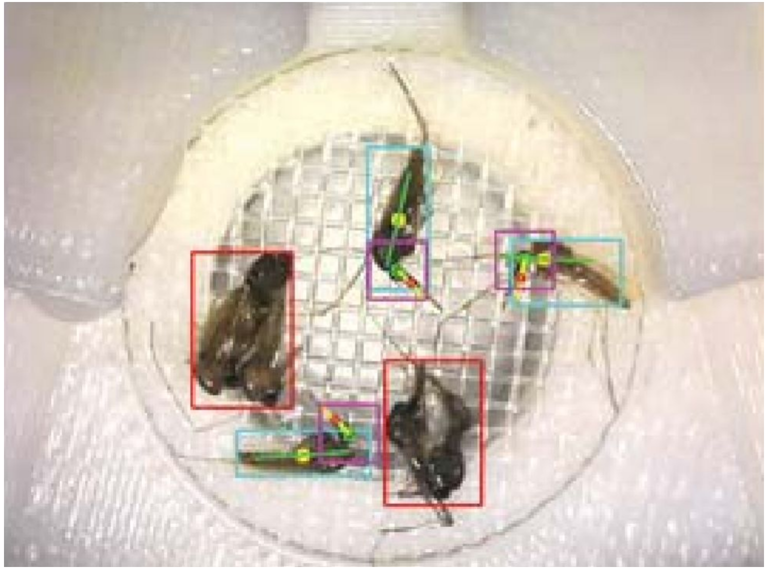


Figure 11. Output of Image Processing Approach (Wu)



Results: Deep Learning

For IoU > 0.75

- Precision - clustered: 0.88, non-clustered: 0.86
- Average precision - clustered: 0.82, non-clustered: 0.85
- Mean average precision - 0.84

For IoU > 0.5, mean average precision - 0.96

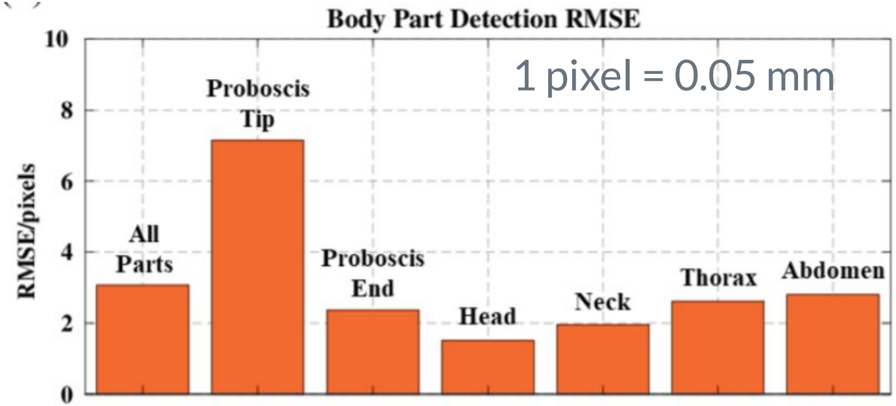


Figure 12. Body Part Detection RMSE for Deep Learning Approach (Wu)

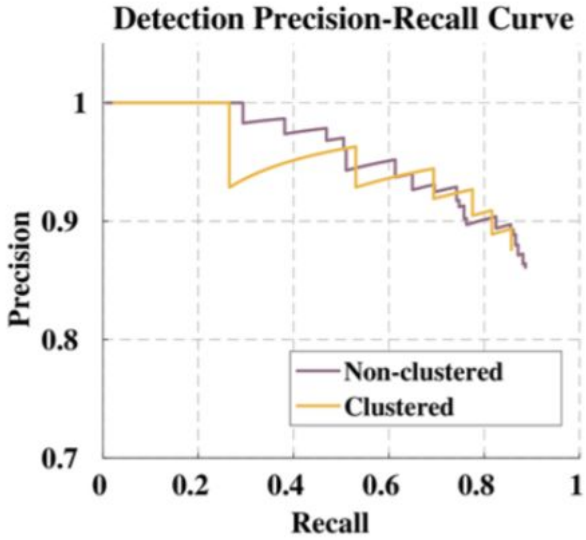


Figure 13. PR Curve for Deep Learning (Wu)

Results: Comparison

- Image processing variant to both luminance and scale

Approach	Deep Learning	Image Processing
Detection mAP (IoU>0.5)	0.97	0.80
Detection Recall	0.97	0.90
Head Position RMSE	1.61 pixels	2.70 pixels
Proboscis Orientation Error	14.3°	24.7°
Processing Speed	2.5 fps	20 fps

Figure 14. Table from Wu et al. comparing results of Deep Learning and Image Processing approaches (Wu)

Paper Assessment

Validated the usage of deep learning as a method for mosquito detection and keypoint identification within the proposed workflow of an automatic mosquito microdissection system

Relevance

- Framework/process flow for deep learning of keypoint detection
- Expected precision, mAP, and RMSE for deep learning approaches
- Validation of deep learning as approach
- Identified pitfalls and limitations of deep learning and image processing approaches

Strengths

- Comprehensive dataset
- Complete system workflow implemented for both deep learning and image processing
- Comparative study between image process and DL

Limitations

- Insufficient result reporting for image processing method
- Comparison of two models does not consider problem context
    - No comparison of proboscis end accuracy

Next Steps and Future Work

Work done after paper publication

- Change of proposed MMS system to current system (left)
  - Manual placement
  - Change in staging area, especially with regards to background
- Addition of the need for orientation detection of mosquito
- New image processing approaches for proboscis detection implemented

Current work related to paper

- Collect more data, especially with new MMS setup
- Reimplementation of more robust deep learning system
- Documentation of deep learning system
- Integration of new DL system with ROS and the MMS.

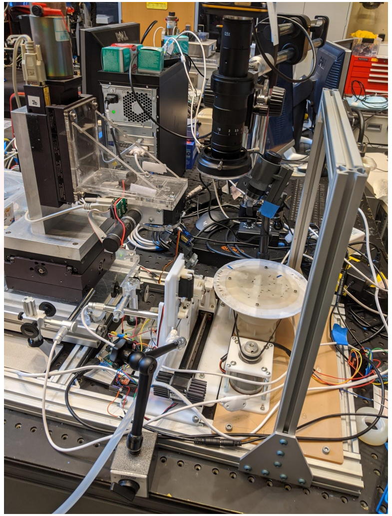


Figure 15. Current MMS system

References

[1] Wu, H., Mu, J., Da, T., Xu, M., Taylor, R. H., Iordachita, I., & Chirikjian, G. S. (2019). Multi-mosquito object detection and 2d pose estimation for automation of PfSPZ malaria vaccine production. In 2019 IEEE 15th International Conference on Automation Science and Engineering, CASE 2019 (pp. 411-417). [8842953] (IEEE International Conference on Automation Science and Engineering; Vol. 2019-August). IEEE Computer Society. https://doi.org/10.1109/COASE.2019.8842953

[2] M. Schrum, A. Canezin, S. Chakravarty, M. Laskowski, S. Comert, Y. Sevimli, G. S. Chirikjian, S. L. Hoffman, R. H. Taylor, "An efficient production process for extracting salivary glands from mosquitoes", arXiv:1903.02532, 2019.

[3] H. Phalen, P. Vagdargi, M. Pozin, G. S. Chirikjian, I. Iordachita, R. H. Taylor, "Mosquito pick-and-place: Automating a key step in pfspz-based malaria vaccine production", Accepted to the 2019 15th IEEE International Conference on Automation Science and Engineering (CASE 2019).

Thank you!

Appendix: Metrics

Precision - % of predictions correct

Recall - how well you find all positives

IoU - intersection over union

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1 = 2 \cdot \frac{precision \cdot recall}{precision + recall}$$

TP = True positive

TN = True negative

FP = False positive

FN = False negative

$$IoU = \frac{\text{area of overlap}}{\text{area of union}}$$

AP (Average Precision)

mAP (Mean Average Precision)

$$\text{AP@k} = \frac{1}{\text{GTP}} \sum_{i=1}^{k} \frac{\text{TP seen}}{i}$$

$$\text{mAP} = \frac{1}{N} \sum_{i=1}^{N} \text{AP}_i$$

Tan, Ren Jie. Breaking Down Mean Average Precision (mAP). Towards Data Science. 2019. https://towardsdatascience.com/breaking-down-mean-average-precision-map-ae462f623a52
Hui, Jonathan. mAP (mean Average Precision) for Object Detection. Medium.com. 2018. https://medium.com/@jonathan_hui/map-mean-average-precision-for-object-detection-45c121a31173