**Using Machine Learning to Measure Changes in Immunity due to Climate Change in the Fall Armyworm, *Spodoptera frugiperda***

**Abstract:**

The fall armyworm (*Spodoptera frugiperda*) is a major agricultural pest causing millions in damage to farmlands across the globe. With the recent spread into Africa and Asia, the need for understanding the best form of biocontrol is at an all-time high. While viruses are often used in controlling insect pests, not much is known about how fall armyworms react with specialist pathogens and how that reaction may vary with changes in climate. Furthermore, current methods of measuring immune response require the manual counting of insect immune cells, requiring large amounts of time on the researcher’s part. Here we use machine learning and compare several object detection algorithms on their ability to automatically count hemocytes. We also rear fall armyworms generationally at two temperatures, 26° C and 31° C, and infect each generation with virus, allowing the survivors to reproduce. Hemocytes were extracted by cutting the fourth proleg and mixed with anticoagulant for use in photographing on a hemocytometer. Of the algorithms compared, YOLOv8 was the most accurate and the quickest to train. Counts tended to be fairly accurate and predicted that 26 populations had a higher immune response in control populations that were infected and in the uninfected coevolved population. The coevolved populations that were infected showed no significant difference in immune response when comparing temperatures. This new development will drastically speed up the process of measuring insect immune response and open the door for further research to be done on biocontrol methods in insects.

**Intro:**

The fall armyworm (*Spodoptera frugiperda*) is a major agricultural pest thought to originate from the tropical environments of Central and South America. The larvae feed primarily on corn, rice, and grasses, but have also been known to destroy crops like tobacco and cotton (Wang et al., 2022). In the last 10 years, fall armyworms have spread into Western and Central Africa causing massive agricultural damage to maize plots within the area (Goergan et al, 2016). Fall armyworms thrive in warmer climates and are most active in the late summers in North America (Ali et al., 1990). While pesticides have traditionally been used to help cull outbreaks, the emergence of resistance to commonly used pesticides like carbamates and organophosphates has made it increasingly challenging to control the populations (Tay et al, 2023). Biocontrol practices like the use of predators, parasitoids, and pathogens have been shown to be successful in reducing the number of larvae. Compared to pesticides, biocontrol methods have a unique ability to adapt and overcome host resistance, thereby leading to more long-term control options. Despite the potential aforementioned benefits, only a limited number of studies have investigated the use of biocontrol agents to control populations of the fall armyworm, with those available focusing primarily on parasitoids and pheromone traps (Allen et al., 2021; Varshney et al., 2021). Of those focusing on pathogens, many focus on the use of *Spodoptera frugiperda* multicapsid nucleopolyhedrovirus (SpfrMNPV), a specialist virus that has been shown to effectively reduce fall army worm populations, making it a candidate for further study (Behle & Popham, 2012).

An insect’s primary immune response against small viral particles is initiated through hemocytes. The primary function of hemocytes is to phagocytize or degrade invading small foreign objects through enzyme release. In lepidopteran species, like the fall armyworm, multiple uniquely shaped hemocytes have been found to aid in immune response (Lavine and Strand, 2002). Hence, hemocyte counting is a vital step in understanding how the immune system of the fall armyworm and other arthropod species responds to pathogens. When insects are infected with a pathogen, their hemocyte counts are lower than those of their healthy counterparts (Smilanich et al., 2018). Conversely, hosts infected with high viral concentrations display increased hemocyte counts when compared to healthy individuals and to those exposed to a low dose (Eslin & Prévost, 1998).

The current method for detecting and counting hemocytes requires manually tallying their presence by hand. This approach can be tedious and time-consuming since each image may take up to 60 seconds to process accurately. These manual counting methods may not always be reliable due to the potential for user error when counting out-of-focus hemocytes and other cells. To mitigate this problem, programs like ImageJ are commonly used for counting objects in images (Schneider et al., 2012). However, ImageJ often struggles to provide accurate counts due to the unique and diverse shape of hemocytes, the presence of other miscellaneous cells in samples, and the placement of cells on the hemocytometer grids. Additionally, hemocytes will cluster together after a period of time which can further compound these challenges. Moreover, camera quality and contrast can introduce additional inaccuracies in counts and skew the results. In particular, images taken in low light conditions result in reduced visibility of hemocytes and skew the analysis.

Object detection is a computer vision technique that allows for the localization and classification of objects contained within photos. Machine learning is crucial for object detection, allowing users to develop models to automatically detect objects within images. Many object detection approaches use convolutional neural networks (CNNs) to enable direct learning from photo data (Ren & Wang, 2022). Object detection algorithms are categorized based on the complexity of the model and the steps involved for detection. Single-stage detectors directly feed image data into a convolutional network to detect the location and classification of the objects. Examples of commonly used single-stage detectors include You Only Look Once (YOLO) and Single Shot Multibox Detector (SSD) (Redmon et al., 2016; Liu et al., 2016). In contrast, two-stage detectors incorporate an additional region proposal step during detection before classifying and drawing bounding boxes. Fast R-CNN and Faster R-CNN are both common examples of two-stage detectors (Girshick, 2015; S. Ren et al., 2016). In theory, two-stage detectors tend to be more accurate at the cost of additional time and computational resources due to the complexity of the network, while single-stage are faster to train, but less accurate overall and struggle more on small objects. For instance, a previous study found Faster-RCNN to the be the most accurate model compared to other single-stage detectors (Wang et al., 2022). However, other studies differentiating fish morphology and detecting cars from aerial images found single shot-detectors to be more accurate overall. Additionally, more recent developments in YOLO models (YOLO V8) suggest the potential to overcome their accuracy limitations. (Jocher et al., 2023).

Previous research has used deep learning methods to aid in the detection of cells other than hemocytes. For example, many studies have performed analysis of blood smears to detect infected white blood cells (Poostchi et al., 2018). However, these studies often aim to detect large objects occupying the majority of pixels in the photos. One current obstacle faced in objected detection is the difficulty in predicting small objects. This is in part due to the limited data for use in training and to the low resolution that often occurs due to the objects' small size (Nguyen et al., 2020). Additionally, many CNN-based architectures rely on hierarchy feature mapping in predicting objects. While this may work for large and medium sized objects, small objects can easily be lost in the latter layers of these models due to the down sampling required (Liu et al., 2021). Therefore, it is important to consider how to account for the small size of hemocyte cells and compare the accuracy of multiple models, including both single-stage and two-stage detectors.

To our knowledge, no studies have investigated the automation of hemocyte counts in insects driving the need for new methods. Using machine learning should lead to accurate hemocyte counts that take significantly less time when compared to manual counting. We expect single-stage detectors to be faster and more efficient during training, but we also expect Two-Stage detectors to be more accurate. For trial data on temperature comparisons and hemocytes, we expect temperature to influence hemocyte counts. Additionally, we anticipate that caterpillar colonies exposed to the virus to exhibit a stronger hemocyte response than those who have not. Finally, caterpillars exposed to virus when dissected should show a stronger hemocyte response than those that were not exposed.

**Materials and Methods:**

* **Rearing and Hemocyte Extraction**

The first generation of fall armyworms was purchased from Benzon Research Incorporated (cite?). Caterpillars were reared on a mix of corn-based diet until reaching fourth instars. For the first generation, caterpillars were split into two treatments: 26° C and 31° C. The 26° C treatment was kept at 26° C for 12 hours during the day and 16° C at night, while the 31° C treatment was kept at 31° C for 12 hours during the day and 21° C at night. For each treatment, a dose response was initially performed to calculate the lethal viral dose for 50% of the population (LD50). We isolated and starved 180 caterpillars for 24 hours to ensure consumption of SpfrMNPV. Caterpillars were then transferred to individual cups containing small cubes of diet laced with 3 **μl of** virus solutions. The dosage of virus ranged from 100 cells/μL of virus to 106 cells/μL of virus, increasing by a factor of 10 for each virus treatment. The control tray was dosed with DI water containing no virus to account for aversion to wet diet. We left caterpillars in cups with the virus (or water) laced cubes for another 24 hours to ensure full consumption of diet. Once the 24 hours had passed, caterpillars were transferred back into individual containers half full of the corn-based diet. The date of mortality or pupation was recorded for each individual for use in calculating the LD50 for the subsequent generation. Once all caterpillars had pupated or died, individuals infected with SpfrMNPV were processed to collect remaining viral particles. This processed virus was then used to infect the next generation of caterpillars. For individuals that survived infection and pupated, caterpillars were transferred to cages containing forty to sixty individuals to ensure reproduction. Eggs were collected every 24 hours and labeled as the next subsequent generation. Neonates that emerged from the new generation were isolated into 1 oz cups filled with ½ ounces of diet. Once the neonates grew into fourth instars, they were then starved, and another dose response was performed to calculate the LD50 for that generation. This was repeated every generation infecting each subsequent generation with virus collected from the previous generation.

Within each temperature treatment, 27 caterpillars were set aside as controls while an additional 27 caterpillars were dosed with enough virus to kill 95 percent of the population of virus. Nine caterpillars from the control and nine caterpillars from the infected treatment were taken for collection every 24 hours, until 72 hours had passed and all caterpillars were used. As per USDA protocol for extracting hemocytes, larvae were initially placed at 4°C for 30 minutes. Once the insects were mostly immobile, they were transferred on ice for extraction. The fourth abdominal prolegs were sterilized with ethanol and then amputated using surgical needles. Hemolymph from the incision was allowed to pool for collection. Three caterpillars per treatment were bled, and the total hemolymph of these caterpillars were collected. This was repeated two more times for both controls and infected, leading to three sets of photos. Once hemolymph was extracted, the solution was mixed with a chilled anti-coagulant saline solution in a ratio to ensure at least 10 µl of solution were present. Generally, a ratio of 1 hemolymph: 2 anticoagulants was used for 48 hour and 72 hour treatments, while a ratio of 1:3 hemolymph to anticoagulant was used for 24 hour treatments. The solution was then placed on a hemocytometer for counting under a 40x magnification lens. A Canon Powershot G10 14.7MP digital camera was used to photograph 5 of the 25 squares for estimating the total number of hemocytes.

* **Machine Learning**

To gather data used in training the model, the same approach was used for collecting and gathering hemolymph. Nine photos were taken on the hemocytometer grid. Additional photos were taken off of the grid if hemocytes exhibited a unique shape or was in a large cluster. Light intake and camera zoom were varied for each set of photos to account for differences when capturing photos. 571 photos were collected for use as training, validation, and test sets. Each photo was manually annotated using the program LabelImg (<https://github.com/qaprosoft/labelImg>) to ensure annotation accuracy. Halos of light present around hemocytes were included in boxes to increase the size of detection and give more context to the objects when training. Hemocytes out of focus or lacking definition were not included during annotation.

Photos were randomly separated into training, validation, and testing with 70% of photos entered as training data, 20% as validation, and 10% as test photos to ensure data was accurately read. Training data was used to train the models, the validation set was used during training to estimate the current model accuracy and help prevent overfitting, and the test set was used once training finished to measure the accuracy of each algorithm incounting hemocytes. After data was sorted, the photos labeled training were augmented to increase the size of the dataset. Data in training was then augmented to flip 180 degrees to account for any unusual orientation of hemocytes and to better account for changes in light. The contrast of photos was further exaggerated to ensure hemocyte detection could occur even with different levels of exposure. This tripled the size of our training data. Models were then trained twice, once with and once without the training data to determine if the additional computational time required to train more photos was beneficial to overall accuracy.

All models were trained using the programming language Python. Two versions of YOLO models (YOLOv5x and YOLOv8x) were used to represent single-shot detectors. YOLO models were allowed to run until no improvement had been made in the past 50 runs, up to a maximum of 200 epochs. Code for training YOLO models was sourced from Ultralytics for use in training on our own models (Jocher, 2020; Jocher et al., 2023). All models other than YOLO based models were trained using MMDetection and code available on Github due to their extensive model library and easy to configure models (Chen et al., 2019). The faster-RCNN model with a backbone of R-101-FPN was run until 200 steps were performed and no significant changes in loss rates were observed. The results from YOLO models were compared to all other models to determine the fastest and the most accurate algorithm.

All models were trained using two NVIDIA V100S GPUs, each with 32 GB of memory. Additionally, all photos were condensed to a size of 1280 x 1280 pixels to reduce the overall training time. Batch sizes were kept constant at 8 across all models to better measure training times. Models were validated during and after training to calculate precision and recall. Once training and validation of models had occurred, detection was performed on a subset of the data. This test data was then compared to the counts performed by hand to estimate the accuracy of our counting models. Additionally, models were characterized using the recall, precision, and the Mean Average Precision when predicted boxes overlapped more than 50% (mAP50) to determine the best model for use. mAP50-95, or the average precision across multiple overlap thresholds ranging from .5 to .95, was used to further classify the best model. Training time and detection time were also recorded to determine the most accessible model for hardware with lower power. The definition of precision and recall are shown below.

Once all models were trained, the model with the highest mAP50 was used for counting and further comparisons. Counting was compared by taking the best model’s outputs on detections and comparing them to manual annotations on the test data. Temperature treatments were initially compared using all photos regardless of infection status and a T-test was performed to test for significance. Once all photos were cropped down to the hemocytometer square, counts were remade, and another T-test was performed to compare hemocyte counts. Once comparisons between temperatures had been performed, the cropped photos were divided based on their temperature and infection status, then their counts were plotted. An analysis of covariance was performed on the data to determine its statistical significance.

**Results:**

Initial comparison of models showed YOLOv8 has the highest mAP50 at 98.5% on unaugmented with a precision of 95.4% and a recall of 94.2% (Figure 2). This model had a 0.04% higher mAP50 than the most complex YOLOv5 model and faster R-CNN models. Additionally, compared to the Faster-RCNN model, both YOLOv5 and YOLOv8 had overall shorter training times (Table 1). YOLO8 also showed a higher mAP50-95 than both models. Augmenting photos to increase training data had surprisingly little effect on accuracy and precision. YOLOv8 trained on augmented photos had a slightly higher mAP50 of 0.01, but a lower mAP50-95 compared to YOLOv8 models trained on unaugmented data. Other models followed similar trends where augmentation had little to no effect on the accuracy of the model. Due to having the highest accuracy, YOLOv8 was used for counting hemocytes and any further analysis present in photos. Compared to actual counts, counts generated from the model were mostly accurate, only missing one or two hemocytes for the most part. The accuracy of counts also tended to decrease as more hemocytes were present within the photo (Figure 3). False positives tended to be small objects on the edge of photos, or cells with a similar size and shape to hemocytes. False negatives tended to occur when a hemocyte was out of focus or clustered together with many other hemocytes. Counting by hand took on average XX number of seconds while computers were able to detect hemocytes on the same photos in less time.

When using uncropped photos and comparing all photos from the 26 °C to the 31 °C treatment, caterpillars at temperature treatment 26 °C showed a higher mean hemocyte count of 70.67. Caterpillars reared in the 31 °C treatment showed a mean count of 62.16 (Fig 4). A t-test of the two treatments yielded a p-value of 0.0486, meaning that the lower temperature populations showed an overall stronger response regardless of disease presence. When cropping the photos down to squares on the hemocytometer, the 26 °C treatments had an average of 32.04 hemocytes and the 31 °C had an average of 29.65 hemocytes (Fig. 5). However, there was no significant difference between hemocytes when focusing specifically on squares (p = 0.216).

Further breaking down temperature treatments into coevolved and control vs infected showed that coevolved colonies tended to have higher hemocyte counts regardless of the presence of virus. No significant difference in hemocyte counts was observed between the infected coevolved populations at 26 and 31 °C. Uninfected coevolved populations saw a significantly higher number of hemocytes in the 26 °C treatments compared to the 31°C (Figure 5). Infected control populations saw a significant difference in counts with a higher mean in the 26°C population. Finally, the uninfected control showed the opposite effect where the 31°C treatment had a higher average hemocyte count than the 26°C treatment.

**Discussion:**

Using object detection algorithms to count hemocytes in images was significantly faster than counting manually. Utilizing this method has the potential to reduce the amount of time and effort required for counting and analyzing large volumes of immune response data. Although the detection is not always completely accurate, the predictions consistently follow a trend close to a 1 to 1 line. Additionally, the predicted counts typically only differ by one or two hemocytes, thus making it an effective tool for predicting count trends. All models tested had a lower detection time than manual counts. Based on our results, unless computational power is a limiting factor, YOLOv8 is the most accurate and efficient model to use when detecting hemocytes. When compared to YOLOv5 and Faster-RCNN, YOLOv8 had higher accuracy and very short overall detection times. It is unsurprising that YOLOv8 is the most accurate compared to other single-stage models since it is the newest model in the YOLO family and is predicted to be more accurate on the COCO dataset than the other models (Jocher et al., 2023). It is surprising that YOLO outperformed Faster-RCNN, as Faster-RCNN is a two-stage detector while YOLO is a one-stage model. The performance difference could potentially be attributed to factors like background contrast or the fact that Faster-RCNN is an older model. More surprising, however, was how little augmenting the training data changed the overall accuracy of predictions. The addition of augmented photos has been shown to improve predictions in other studies (Shorten & Khoshgoftaar, 2019). This likely indicates that the level of data augmentation during training may not have been sufficient. It is possible that more extensive augmentation techniques beyond changing contrast and flipping photos horizontally may be required to see more substantial improvements. Additionally, reducing brightness may have led to loss of features in the photos, making it harder for the machine to accurately predict the location of hemocytes within the photos.

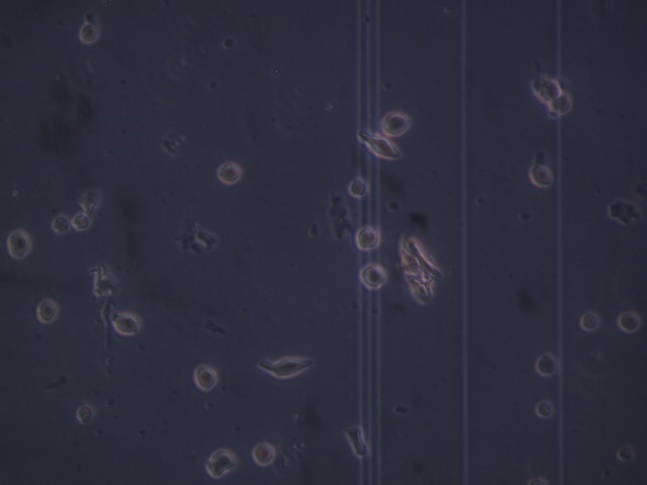
Uncropped photos grouped by temperature treatments showed a low significance indicating that caterpillars in the 26 °C treatment showed higher hemocyte count compared to the 31 °C. While previous studies have shown infected insects tend to prefer hotter temperatures, this could mean that caterpillars raised at the 31°C treatment may be expending more energy on temperature regulation rather than mounting an immune response (Catalán et al., 2012). In contrast, photos cropped to the hemocytometer squares showed no significant differences between temperature treatments. This suggests that cropping to the hemocytometer squares helped eliminate bias from camera zoom meaning that the significance found on uncropped photos could be an error due to the irregular area captured.

Splitting each temperature treatment into control vs infected and coevolved colonies vs no coevolution colonies yielded different results. All infected populations showed significantly more hemocytes than their uninfected counterparts. This is due to the need to mount a strong immune response inside the caterpillar to combat the infection. When comparing across temperatures, the 26°C and 31 °C infected coevolved showed no significant difference between treatments. This may suggest that populations of fall armyworms that coevolve with their virus are able to produce similar immune responses regardless of the temperature they are in. Other studies have shown insects raised at different temperatures exhibit varying susceptibilities to pathogens, but a wider spread of temperatures is often used, and results vary between hosts and viruses (Mastore et al., 2019). The uninfected coevolved showed higher hemocyte counts in the 26°C populations, indicating that coevolved caterpillars at colder temperatures may be better equipped at a given point to defend against pathogens. The control population showed higher hemocyte counts in the 31°C uninfected control, but the 26°C treatment showed higher hemocytes in the 26°C infected control. Additionally, these contrasting results may be attributed to the small sample size for the control treatments, or to errors when collecting hemocytes.

Overall, we successfully have used machine learning to count hemocytes and estimate immune response in fall armyworms. While the YOLOv8 model was the most successful, all models tested are viable options for detecting hemocytes and should be chosen based on needs and hardware available. Counts made by machine learning algorithms were mostly accurate and followed a 1:1 trend. Understanding how fall armyworms respond to and coevolve with their pathogens will allow us to better predict how biocontrol agents will act, and potentially save farmers from significant crop damage.

**Figures:**

A picture containing black

Description automatically generated 

**Figure 1.** Two examples of hemocytes placed on hemocytometer grids at different levels of brightness and saturation. A shows hemocytes on one square of the hemocytometer grid. B shows uniquely shaped hemocytes.

Chart, scatter chart

Description automatically generated

**Figure 2.** The comparison between the hemocyte counts obtained by hand and the hemocyte counts as per Yolov8. The solid line represents a perfect count.

**Table 1.** Machine learning algorithms and information on the average precision, speed of training using the training set, and detection times without data augmentation.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Architecture | Precision | Recall | Map50 | Map50-95 | Time to train | Detection time |
| YOLOv5 | 0.964 | 0.958 | 0.981 | 0.773 | 1.9 Hours |  |
| YOLOv8 | 0.954 | 0.963 | 0.985 | 0.785 | 1.14 Hours |  |
| Faster R-CNN | X | X | 0.976 | 0.764 | 2.024 Hours |  |
|  |  |  |  |  |  |  |

**Table 2.** Machine learning algorithms and information on the average precision, speed of training using the training set, and detection times with data augmentation. The training size increased to 1,371 after augmented photos were added.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Architecture | Precision | Recall | Map50 | Map50-95 | Time to train | Detection time |
| YOLOv5 | ..968 | 0.95 | 0.982 | 0.768 | 5.75 hours |  |
| YOLOv8 | .96 | .953 | 0.986 | 0.77 | 1.70 Hours |  |
| Faster R-CNN | X | X | 0.981 |  | Hours |  |
|  |  |  |  |  |  |  |

Chart, box and whisker chart

Description automatically generated

Fig 3. Average number of hemocytes in uncropped photos (1 is 26 °C, 2 is 31 °C).

Chart, box and whisker chart

Description automatically generated

Figure 4. The average number of hemocytes when focusing on one square of the hemocytometer. (1 is 26 °C, 2 is 31 °C)

Chart, box and whisker chart

Description automatically generated

Figure 5. Hemocyte counts at each temperature across treatments.

**Acknowledgements:**

We would like to thank Kale Costanza and Scott Grimmell for assisting in data collection. This study was funded by (Funding Source)

**Author Contributions:**

Nathaniel Haulk and Bret Elderd designed the ideas and experiments. Nathaniel Haulk collected data and ran analysis. Nathaniel Haulk and Bret Elderd wrote the manuscript.

**Conflict of Interest Statement:**

The authors have no conflicts of interest.

**Data Availability Statement:**

Photos and all code used are made available at <https://github.com/natom24/Machine-Learning>

**References:**

Ali, A., Luttrell, R. G., & Schneider, J. C. (1990). Effects of Temperature and Larval Diet on Development of the Fall Armyworm (Lepidoptera: Noctuidae). *Annals of the Entomological Society of America*, *83*(4), 725–733. <https://doi.org/10.1093/aesa/83.4.725>

Allen, T., Kenis, M., & Norgrove, L. (2021). Eiphosoma laphygmae, a classical solution for the biocontrol of the fall armyworm, Spodoptera frugiperda? *Journal of Plant Diseases and Protection*, *128*(5), 1141–1156. <https://doi.org/10.1007/s41348-021-00480-9>

Behle, R. W., & Popham, H. J. R. (2012). Laboratory and field evaluations of the efficacy of a fast-killing baculovirus isolate from Spodoptera frugiperda. *Journal of Invertebrate Pathology*, *109*(2), 194–200. <https://doi.org/10.1016/j.jip.2011.11.002>

Catalán, T. P., Wozniak, A., Niemeyer, H. M., Kalergis, A. M., & Bozinovic, F. (2012). Interplay between thermal and immune ecology: Effect of environmental temperature on insect immune response and energetic costs after an immune challenge. *Journal of Insect Physiology*, *58*(3), 310–317. <https://doi.org/10.1016/j.jinsphys.2011.10.001>

Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., Zhang, Z., Cheng, D., Zhu, C., Cheng, T., Zhao, Q., Li, B., Lu, X., Zhu, R., Wu, Y., … Lin, D. (2019). MMDetection: Open MMLab Detection Toolbox and Benchmark. *ArXiv Preprint ArXiv:1906.07155*.

Eslin, P., & Prévost, G. (1998). Hemocyte load and immune resistance to Asobara tabida are correlated in species of the Drosophila melanogaster subgroup. *Journal of Insect Physiology*, *44*(9), 807–816. <https://doi.org/10.1016/S0022-1910(98)00013-4>

Girshick, R. (2015). *Fast R-CNN* (arXiv:1504.08083). arXiv. <https://doi.org/10.48550/arXiv.1504.08083>

Goergen, G., Kumar, P. L., Sankung, S. B., Togola, A., & Tamò, M. (2016). First Report of Outbreaks of the Fall Armyworm Spodoptera frugiperda (J E Smith) (Lepidoptera, Noctuidae), a New Alien Invasive Pest in West and Central Africa. *PLOS ONE*, *11*(10), e0165632. <https://doi.org/10.1371/journal.pone.0165632>

Jocher, G. (2020). *YOLOv5 by Ultralytics* (7.0). <https://doi.org/10.5281/zenodo.3908559>

Jocher, G., Chaurasia, A., & Qiu, J. (2023). *YOLO by Ultralytics* (8.0.0). <https://github.com/ultralytics/ultralytics>

Lavine, M. D., & Strand, M. R. (2002). Insect hemocytes and their role in immunity. *Insect Biochemistry and Molecular Biology*, *32*(10), 1295–1309. <https://doi.org/10.1016/S0965-1748(02)00092-9>

Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). Ssd: Single shot multibox detector. *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*, 21–37.

Liu, Y., Sun, P., Wergeles, N., & Shang, Y. (2021). A survey and performance evaluation of deep learning methods for small object detection. *Expert Systems with Applications*, *172*, 114602. <https://doi.org/10.1016/j.eswa.2021.114602>

Mastore, M., Quadroni, S., Toscano, A., Mottadelli, N., & Brivio, M. F. (2019). Susceptibility to entomopathogens and modulation of basal immunity in two insect models at different temperatures. *Journal of Thermal Biology*, *79*, 15–23. <https://doi.org/10.1016/j.jtherbio.2018.11.006>

Nguyen, N.-D., Do, T., Ngo, T. D., & Le, D.-D. (2020). An Evaluation of Deep Learning Methods for Small Object Detection. *Journal of Electrical and Computer Engineering*, *2020*, e3189691. <https://doi.org/10.1155/2020/3189691>

Poostchi, M., Silamut, K., Maude, R. J., Jaeger, S., & Thoma, G. (2018). Image analysis and machine learning for detecting malaria. *Translational Research*, *194*, 36–55. <https://doi.org/10.1016/j.trsl.2017.12.004>

Prasetyo, E., Suciati, N., & Fatichah, C. (2020). A Comparison of YOLO and Mask R-CNN for Segmenting Head and Tail of Fish. *2020 4th International Conference on Informatics and Computational Sciences (ICICoS)*, 1–6. <https://doi.org/10.1109/ICICoS51170.2020.9299024>

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788. <https://doi.org/10.1109/CVPR.2016.91>

Ren, J., & Wang, Y. (2022). Overview of Object Detection Algorithms Using Convolutional Neural Networks. *Journal of Computer and Communications*, *10*(1), Article 1. <https://doi.org/10.4236/jcc.2022.101006>

Ren, S., He, K., Girshick, R., & Sun, J. (2016). *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks* (arXiv:1506.01497). arXiv. <https://doi.org/10.48550/arXiv.1506.01497>

Schneider, C. A., Rasband, W. S., & Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nature Methods*, *9*(7), 671–675.

Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. *Journal of Big Data*, *6*(1), 60. <https://doi.org/10.1186/s40537-019-0197-0>

Smilanich, A. M., Langus, T. C., Doan, L., Dyer, L. A., Harrison, J. G., Hsueh, J., & Teglas, M. B. (2018). Host plant associated enhancement of immunity and survival in virus infected caterpillars. *Journal of Invertebrate Pathology*, *151*, 102–112. <https://doi.org/10.1016/j.jip.2017.11.006>

Tay, W. T., Meagher, R. L., Czepak, C., & Groot, A. T. (2023). Spodoptera frugiperda: Ecology, Evolution, and Management Options of an Invasive Species. *Annual Review of Entomology*, *68*(1), 299–317. <https://doi.org/10.1146/annurev-ento-120220-102548>

Varshney, R., Poornesha, B., Raghavendra, A., Lalitha, Y., Apoorva, V., Ramanujam, B., Rangeshwaran, R., Subaharan, K., Shylesha, A. N., Bakthavatsalam, N., Chaudhary, M., & Pandit, V. (2021). Biocontrol-based management of fall armyworm, Spodoptera frugiperda (J E Smith) (Lepidoptera: Noctuidae) on Indian Maize. *Journal of Plant Diseases and Protection*, *128*(1), 87–95. <https://doi.org/10.1007/s41348-020-00357-3>

Wang, J., Mall, S., & Perez, L. (n.d.). *The Effectiveness of Data Augmentation in Image Classiﬁcation using Deep Learning*.

Wang, P., Meng, F., Donaldson, P., Horan, S., Panchy, N. L., Vischulis, E., Winship, E., Conner, J. K., Krysan, P. J., Shiu, S.-H., & Lehti-Shiu, M. D. (2022). High-throughput measurement of plant fitness traits with an object detection method using Faster R-CNN. *New Phytologist*, *234*(4), 1521–1533. <https://doi.org/10.1111/nph.18056>

Wang, W., He, P., Zhang, Y., Liu, T., Jing, X., & Zhang, S. (2020). The Population Growth of Spodoptera frugiperda on Six Cash Crop Species and Implications for Its Occurrence and Damage Potential in China. *Insects*, *11*(9), Article 9. <https://doi.org/10.3390/insects11090639>

Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). *Object Detection in 20 Years: A Survey* (arXiv:1905.05055). arXiv. <http://arxiv.org/abs/1905.05055>