Inquiry: The University of Arkansas Undergraduate Research Journal

Volume 21 | Issue 1 Article 5

2022

# Artificial Intelligence System for Automatic Imaging, Quantification, and Identification of Arthropods in Leaf Litter and Pitfall Samples

Pierce Helton
*University of Arkansas, Fayetteville*

Khoa Luu
*University of Arkansas, Fayetteville*

Ashley Dowling
*University of Arkansas, Fayetteville*



## Recommended Citation

Helton, P., Luu, K., & Dowling, A. (2022). Artificial Intelligence System for Automatic Imaging, Quantification, and Identification of Arthropods in Leaf Litter and Pitfall Samples. *Inquiry: The University of Arkansas Undergraduate Research Journal, 21*(1). https://doi.org/10.54119/inquiry.2022.21101

# Artificial Intelligence System for Automatic Imaging, Quantification, and Identification of Arthropods in Leaf Litter and Pitfall Samples

**Pierce Helton**[1] (pchelton@uark.edu)
**Faculty Mentors – Khoa Luu**[1] (khoaluu@uark.edu) **and Ashley Dowling**[2] (adowling@uark.edu)

[1] *Department of Computer Science and Computer Engineering, College of Engineering, University of Arkansas, Fayetteville, AR 72701*
[2] *Department of Entomology and Plant Pathology, Dale Bumpers College Agricultural, Food, and Life Sciences, University of Arkansas, Fayetteville, AR 72701*

## Abstract

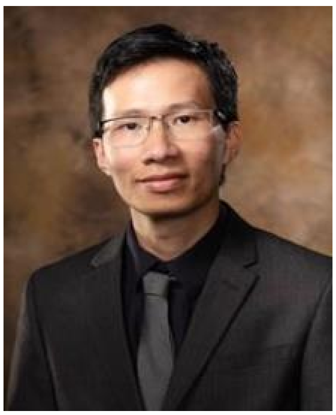
It is well known that arthropods are the most diverse and abundant eukaryotic organisms on the planet. Museum and research collections have huge insect accumulations from expeditions conducted over history that contain specimens of both temporal and spatial value, including hundreds of thousands of species. This biodiversity data is inaccessible to the research community, resulting in a vast amount of "dark data". The primary objective of this study is to develop an artificial intelligence-driven system for specimen identification that greatly minimizes the time and expertise required to identify specimens in atypical environments. Successful development will have profound impacts on both ecology and biodiversity sciences as it will increase the resolution for ecological studies and allow us to work through the backlog of insect collections, unlocking tremendous amounts of biodiversity data. Development of the system will address multiple challenges in deep learning, including problems associated with limited training data and moving from known domains into unknown. The cutting-edge AI solutions will be a final component in a smart specimen identification system scalable in multiple platforms and across geographic region.

## Keywords:

Adversarial Training – Training a model using a discriminator and generator. The generator generates images while the discriminator determines which images are not close to the ground truth.
Domain Adaptation – Applying the knowledge of a deep learning model to a different domain.
Semantic Segmentation – Assigning a label to every pixel in an image.
Semi-supervised Learning – Training a model on labeled data, then using unlabeled data to improve the training results.

## Profiles of the Authors

**Pierce Helton** is an undergraduate CSCE major at the University of Arkansas. He is working as a research assistant in Dr. Luu's CVIU lab. His research interests are Machine Learning, Deep Learning, and Computer Vision. Pierce has plans to work in the industry and potentially pursue a M.S. in Computer Science after earning his B.S.



**Dr. Khoa Luu** is an Assistant Professor and the Director of Computer Vision and Image Understanding (CVIU) Lab in Department of Computer Science and Computer Engineering (CSCE) at University of Arkansas, Fayetteville, USA. He is serving as an Associate Editor of IEEE Access Journal. He was the Research Project Director in Cylab Biometrics Center at Carnegie Mellon University (CMU). His research interests focus on various topics, including Biometrics, Face Recognition, Tracking, Human Behavior Understanding, Segmentation, Scene Understanding, Domain Adaptation. He has received four patents and two best paper awards and coauthored 100+ papers in conferences, technical reports, and journals. He was a vice-chair of Montreal Chapter IEEE SMCS in Canada. He is a co-organizer and a chair of CVPR Precognition Workshop in 2019, 2020, 2021; MICCAI Workshop in 2019, 2020 and ICCV Workshop in 2021. He is a PC member of AAAI, ICPRAI in 2020 and 2021.

**Dr. Ashley Dowling** is a Professor in the Department of Entomology and Plant Pathology at the University of Arkansas in Fayetteville, USA. His lab will provide expertise in insect biodiversity and identification. He will also participate in the intellectual development and training of the artificial intelligence system and design of the data collection system. Dr. Dowling's lab focuses on identification, ecology, and evolution of insects living in both terrestrial and aquatic habitats and has 70+ papers on arthropods. Dr. Dowling and his lab will conduct insect capturing, handling, and identification, which is critical for successfully training the computer to identify a diversity of insects. Dr. Dowling also has extensive experience imaging insects, both alive and dead, and will help develop the image capture system on the trap and assist with the integration of these components into field-ready traps.



https://scholarworks.uark.edu/inquiry/vol21/iss1/5
DOI: 10.54119/inquiry.2022.21101

## Introduction

The short-term goal of this project is to create and evaluate a model's performance with our current progress in imaging and labeling to verify the proof of concept of our long-term goals; this is the primary focus of this paper. The long-term goal of this project is to develop a novel AI-based technology to monitor species of insects and provide real-time agricultural recommendations to farmers. A smart machine that utilizes this AI technology would give farmers as well as crop surveyors access to a continuous flow of data and information; a prototype of this design is shown in Figure 1. The final product incorporating the research shared in this paper would be a similar machine to the one shown; it could calculate when to apply the necessary chemicals and pesticides, the proper dosages, and timing of application, likely catching outbreaks before they occur and reducing overall pesticide use while also providing information to the agricultural community to make informed decisions regarding insect management. All of which will save the US farm industry money and, through the reduction of pesticides entering the environment, ultimately, make the world a much safer place. The proposed solution stands to have an enormous impact on both the economy and environment.

The work in our lab is focused on creating an automated system that can detect and identify arthropod specimens. In order to identify each insect that these AI systems encounter, a model must be trained first. This work requires the imaging of tens of thousands of insects; these images will then be used to train an AI for later work. In order to verify the accuracy of the system, it will be tested on novel data gathered in the same locality of the existing insect dataset. Eventually, this work will be optimized for low-power deep learning in order to deploy it on low-cost devices. The final primary goal of the project is to submit and publish multiple papers for top AI conferences and journals. Right now, we are working on a funded project to develop a prototype system that can detect and capture photos of insects; this work will be used to develop the previously mentioned AI that can classify insect species. At the current stage of research, we have captured around 7,000 images but have labeled less than 5% percent of these images: examples of labeled images can be seen in Figure 3. The rate at which it takes to label these images inhibits the production of a model that can reliably identify images, as the model greatly benefits from being able to classify each part an image. The process of labeling an image requires carefully tracing the outline of each body part: this list includes the body, head, legs, and wings.

Figure 1: Testing a prototype of our developed AI-based insect detection and counting system.

By labeling each part of the insect, the model can distinguish between different species, increasing the reliability of properly identifying any given insect. In order for a human to classify an insect, one must inspect each part of the body, with the wings being the most important. If a human, or a computer model, knows what to look for in the wings of the insect, the chances that the insect is correctly identified greatly improves. The process of identifying each pixel in an image is referred to as semantic segmentation in computer vision, and annotation refers to the manual process of semantically identifying each part of an image.

Automatic semantic object understanding and segmentation are computer vision and machine intelligence operations that aim to assign each pixel in an image to a corresponding, predefined class; this data is then used to train AI models. Semantic segmentation has various practical applications such as medicine, agriculture, law enforcement, and transportation. When a model is needed to recognize details and structures of objects, semantic segmentation is a reliable approach. A typical supervised segmentation model is trained on datasets with labels. To train an AI to semantically segment images, a few learning approaches can be used, but all approaches require a sizeable dataset, and the more dependable ones require annotation. However, manually annotating insect images for the semantic segmentation task is costly and time-consuming. For example, our current progress of manual segmentation and annotation is not sustainable, as the time it takes to annotate an image (as shown in Figure 2) vastly outweighs the time it takes to capture an image. Selecting the right machine learning approach can help to alleviate this problem.

The three primary machine learning approaches are supervised, unsupervised, and semi-supervised learning. Supervised learning uses data that has already been labeled, and the model learns from this data, but this approach requires a large amount of labeled data. Unsupervised learning allows the AI to recognize patterns on its own without the aid of labeling, but the results are usually not as accurate when compared to supervised methods. Semi-supervised learning is a combination of both: the model learns on labeled data then uses the unlabeled data for further training. Domain Adaptation offers a solution to the problem of labeling while also maximizing model accuracy; it uses simulated data to train a model used for real-world applications. The simulated data, including ground truth labels, is used to train the model. A ground truth label refers to the annotated image used as the baseline for training. Once the model is trained, the knowledge is transferred and applied to the real-world data. This approach allows the model to use large amounts of labeled data, but there is usually a slight difference between the simulated data and the real data, referred to as the domain gap. This domain gap is then minimized by training the model on the unlabeled real-world data. Richter *et al* [13] pioneered one of the more well-known applications of Domain Adaptation by using information from the game GTA V provided by the game engine to create a new, virtual dataset. The model's knowledge is usually transferred and refined on real-world city images, discussed further in papers such as AdvEnt. The ability to maximize the efficiency of a model using the methods found in the AdvEnt paper[10] does not only apply to scenarios which require simulated and real-world data: any dataset of labeled and unlabeled images that requires generalization across a domain gap can benefits from Domain Adaptation. By treating the labeled insects as the simulated data and the unlabeled insects as the real-world data, we were able to produce a model that can segment images of unlabeled insects.



https://scholarworks.uark.edu/inquiry/vol21/iss1/5
DOI: 10.54119/inquiry.2022.21101

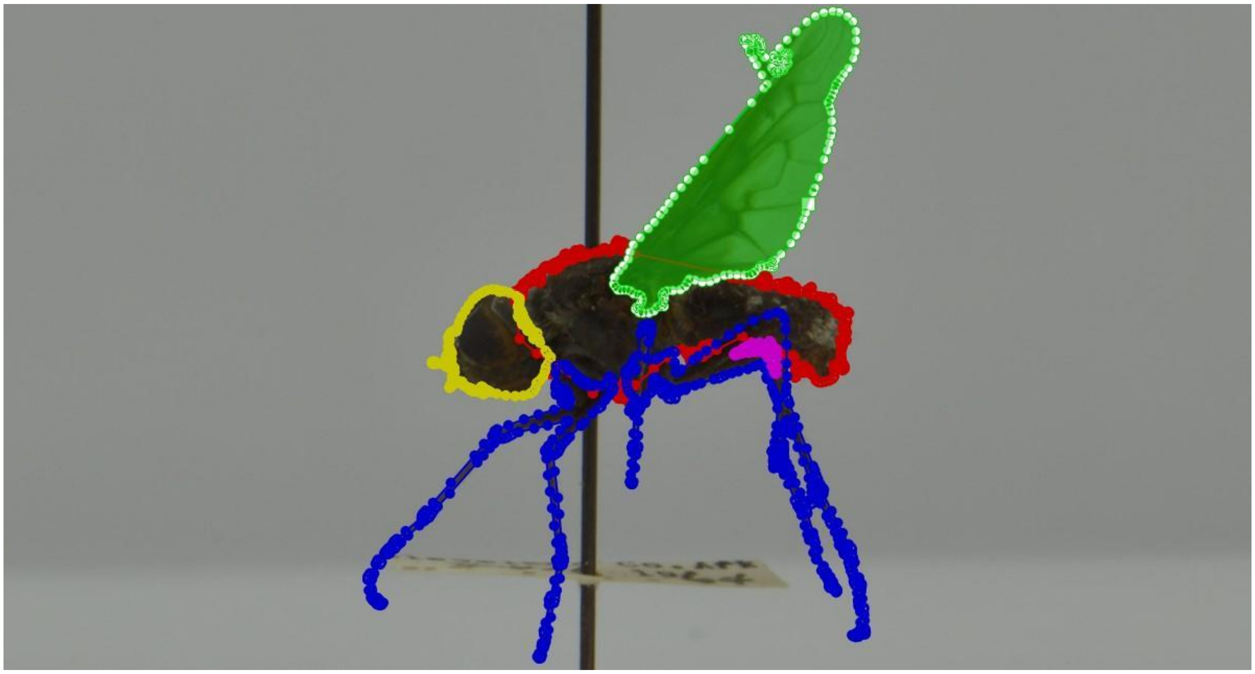


Figure 2: An example of an annotated image. Labeling an image requires tracing an outline around each part of the insect.

## Related Work

Unsupervised Domain Adaptation has recently seen a rise in popularity, prompting more research activity, and when it comes to semantic segmentation, many of the common approaches use generative adversarial training. Adversarial Networks focus on training a generator and discriminator. The generator creates results that mirror the training data, while the discriminator identifies the results that do not fit the training data. Together, the two continue improving results until the generator can create images, that either reach the desired output or convince the discriminator that they are authentic images. Work in our lab related to semantic segmentation implements a Bijective Maximum Likelihood loss to improve the results of scene segmentation. Truong *et al* [9] also used a Domain Score to measure the efficiency of a model on a new domain. The first GAN approach applied to domain adapted semantic segmentation was used by Hoffman *et al* [4]. Yang and Soatto improved their segmentation results by performing domain alignment to reduce the variance between the source and target domains. Cheng *et al* [2] implemented a dual path learning framework that allows two pipelines to interact and improve segmentation results. Ning *et al* [6] used a multi-anchor approach resulting in more representative labels in the target domain, as opposed to traditional centroid based UDA methods.

Insect detection and other agricultural applications of computer vision have been researched in the past. Insect detection in nature relies on small object detection and sufficient generalization, ensuring the methods can work in complex environments. Deng *et al* [3] focuses on using a biologically inspired detection method to identify insects. Kuzuhara *et al* [5] utilizes a two-stage method based on CNNs for detecting and identifying small insects. Research done by Rani uses computer vision techniques to determine whether agricultural crops are affected by pests by using an SVM classifier [5]. Wang *et al* [11] uses an Artificial Neural Network and a Support Vector Machine to improve accuracy results in identification. Chen *et al* [1] created an inexpensive system that uses various acoustic and optical sensors to classify insects. The research done by Samata and Ghosh uses correlation-based feature selection and incremental back propagation in an artificial neural network to detect insects and reduce their impact on crop production.

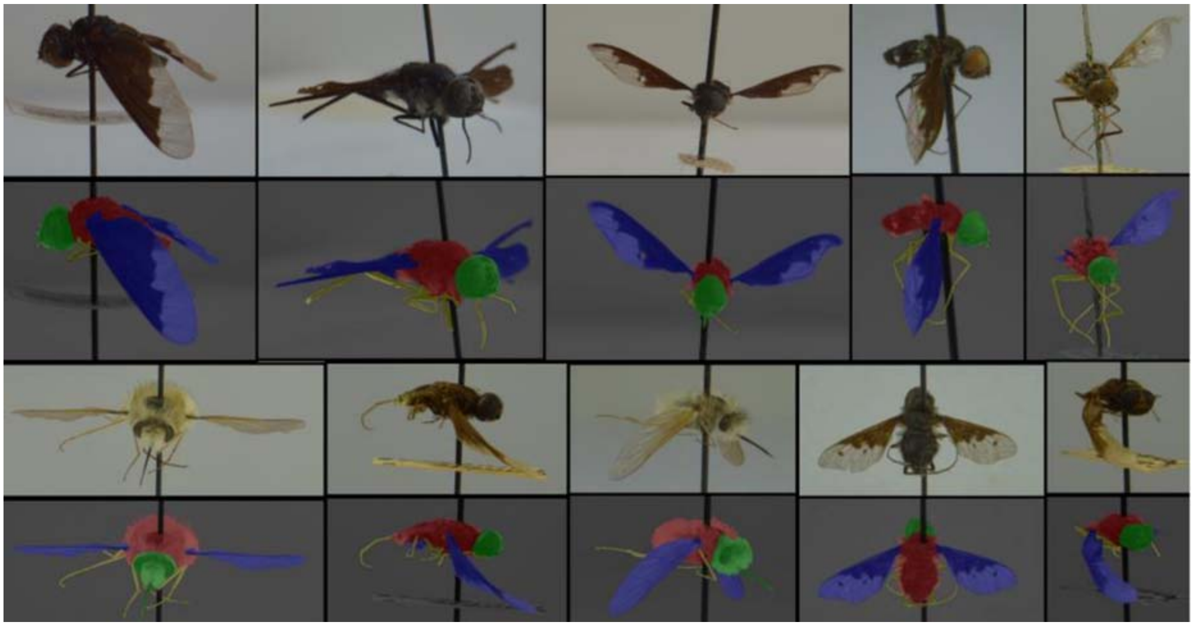


Figure 3: A dataset (7,000 samples) of captured insect photos and their detailed annotation used during training. Part of the research included collecting and labeling some of these samples.

## Methods

Creating a model that can attain the intended results is approachable in a few different ways. The first of these methods uses unsupervised training. At first glance, unsupervised learning may seem like the best approach for our problem: we have an abundance of unlabeled data that would take a significant amount of time to label. However, unsupervised learning has its limitations. The main concern is the risk of inaccurate results. Training a model to segment images is difficult enough, and segmenting small images is even more challenging. Expecting an unsupervised model to reach the desired results is more than hopeful. Validating the results is another concern with unsupervised methods. In order to improve the accuracy of unsupervised learning, human input is often required to ensure the model is approaching the problem correctly. The process of verifying the output of an unsupervised model requires a significant investment of time. Additionally, unsupervised methods usually require larger training sets, which can increase the computational complexity of the system. This also results in lengthier training times, slowing down the pace of improving the results. Overall, unsupervised training is a gamble when it comes to the intended outcome.

Supervised training, on the other hand, requires large amounts of labeled data. Currently, labeling an insect photograph takes around 15 minutes. Labeling the entire dataset of photographed arthropods would take a tremendous amount of time. The room for human error also increases with supervised learning. Improperly labeled images can skew the training results and cause problems for the model's accuracy. For more complex problems, the images may also need to be labeled by an expert in the field in order to properly identify each part of the image. Another problem with supervised training presents itself when the model is applied to the target domain. Usually, the model performs very well on the source domain, but supervised methods can have problems generalizing if the dataset does not include a wide variety of images. Training a supervised model can also take considerably more time than other methods, as each image and its corresponding label needs to be analyzed. Overall, supervised training usually produces better results as opposed to unsupervised methods, but the time and resources required to achieve these results can be inhibiting.

The semi-supervised method we used combines the benefits of the two styles of training while also maximizing the accuracy of the model. Semi-supervised methods train the model by using both labeled and unlabeled data. In the case of our research, a generative model was used. The training data provides ground truth labels and a baseline for the intended results. The unlabeled data allows the generator to learn how to segment the images while the discriminator continues improving the results. The semi-supervised method of training reduces the time it takes to train as well as the overall complexity of the model. Additionally, the amount of labeled data required does not compare to the supervised method, and the results are better than those of an unsupervised method. No human input is required to verify the results, either. A figure of the model is displayed below (Figure 4).

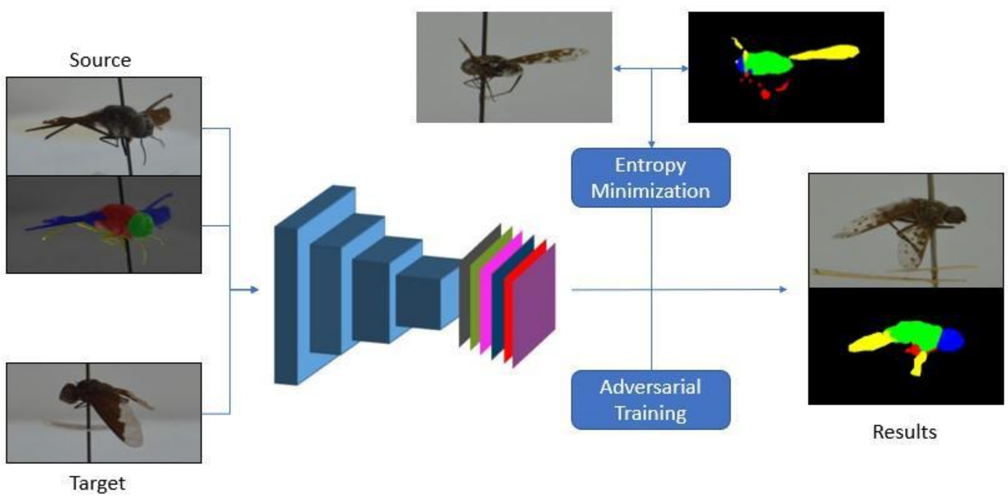



Figure 4: Graphical indication of the methods used to train the model: Domain Adaptation and Adversarial Training.

## Results

The DeepLabv2 model was used during training along with the AdvEnt domain adaptation methods. In order to run the experiment, the AdvEnt datasets needed to be modified. First, the insect label files were converted to the ground truth masks and moved into a repository along with the image files. Each class needed to be changed, too. At this point, the network was being used to segment insect body parts rather than city scenes. Once these steps were completed, the model was trained on the ground truth labels.

The unsegmented images were used during training to allow the model to holistically learn the structure of insects and the segmentation patterns. After the model was sufficiently trained, it produced several segmented images on the unlabeled data. The mean IoU achieved by the model was 38%. IoU, or intersection over union, refers to an algorithm that identifies the overlap and union between the ground truth label and the output produced by the model. The intersection of these two parts is divided by their union to produce the IoU value. Based on this value, one can evaluate the model's efficacy. An IoU value greater than 50% is considered good: because these are preliminary tests, the final number we achieved is promising. The final results of our experiments can be seen in Figure 5.

## Conclusions

One thing to note was the model's high accuracy when segmenting the side profile of insects and the accuracy in segmenting the head; the wings and body are not as accurate. The model cannot segment legs.

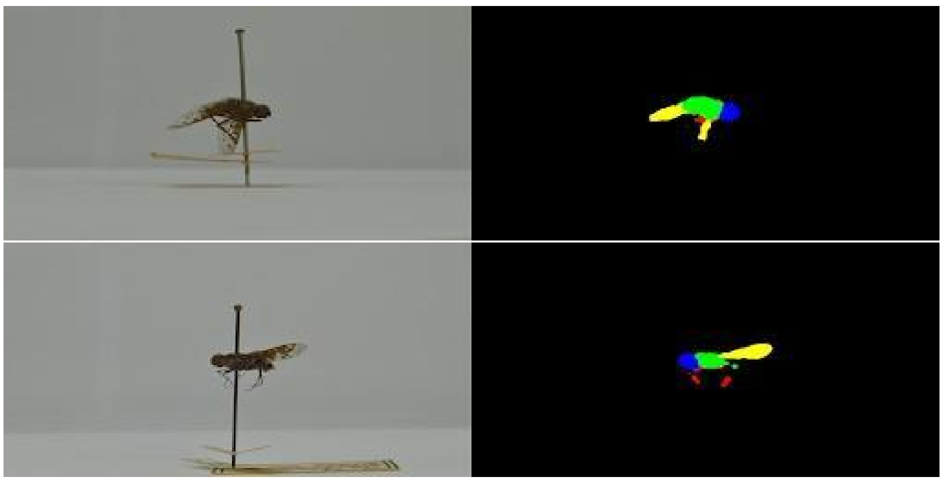


Figure 5: Segmented images produced by the model after training.

However, small objects cannot be reliably segmented by most models. Additionally, the model does not generalize well: insects with different sizes or colors cause issues. From the images, we can learn how to improve the results of the model. Including more variety in the insects used to train the model would allow it to segment different insect types more reliably. Incorporating other methods that segment small objects reliably and improve overall accuracy would likely improve results, as well.

## Acknowledgements

This work is supported by NSF Data Science, Data Analytics that are Robust and Trusted (DART), Chancellor's Innovation Fund, UAF and NSF Small Business Grant.

## References

[1]Chen, Y., Why, A., Batista, G., Mafra-Neto, A., & Keogh, E. (2014). Flying insect detection and classification with inexpensive sensors. Journal of Visualized Experiments.

[2]Cheng, Y., Wei, F., Bao, J., Chen, D., Wen, F., & Zhang, W. (2021). Dual path learning for domain adaptation of semantic segmentation. ArXiv:2108.06337 [Cs].

[3]Deng, L., Wang, Y., Han, Z., & Yu, R. (2018). Research on insect pest image detection and recognition based on bio-inspired methods. Biosystems Engineering, 169, 139–148.



https://scholarworks.uark.edu/inquiry/vol21/iss1/5
DOI: 10.54119/inquiry.2022.21101

[4]Hoffman, J., Tzeng, E., Park, T., Zhu, J.-Y., Isola, P., Saenko, K., Efros, A. A., & Darrell, T. (2017). Cycada: Cycle-consistent adversarial domain adaptation. ArXiv:1711.03213 [Cs].

[5]Kuzuhara, H., Takimoto, H., Sato, Y., & Kanagawa, A. (2020). Insect pest detection and identification method based on deep learning for realizing a pest control system. 2020 59th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE), 709–714.

[6]Ning, M., Lu, D., Wei, D., Bian, C., Yuan, C., Yu, S., Ma, K., & Zheng, Y. (2021). Multi-anchor active domain adaptation for semantic segmentation. ArXiv:2108.08012 [Cs].

[7]Rani, R., & Amsini, P. (2016). Pest Identification in Leaf Images using SVM Classifier. International Journal of Computational Intelligence and Informatics, Vol. 6(No. 1).

[8]Samanta, R., & Ghosh, I. (2012). Tea insect pests classification based on artificial neural networks. International Journal of Computer Engineering Science.

[9]Truong, T.-D., Duong, C. N., Le, N., Phung, S. L., Rainwater, C., & Luu, K. (2021). Bimal: Bijective maximum likelihood approach to domain adaptation in semantic scene segmentation. ArXiv:2108.03267 [Cs].

[10]Vu, T.-H., Jain, H., Bucher, M., Cord, M., & Pérez, P. (2019). Advent: Adversarial entropy minimization for domain adaptation in semantic segmentation. ArXiv:1811.12833 [Cs].

[11]Wang, J., Lin, C., Ji, L., & Liang., A. (2012). A new automatic identification system of insect images at the order level. Knowledge-Based Systems.

[12] Yang, Y., & Soatto, S. (2020). Fda: Fourier domain adaptation for semantic segmentation. ArXiv:2004.05498 [Cs].

[13]Richter, S., & Vineet, V., & Roth, S., & Koltun, V. (2016). Playing for Data: Ground Truth from Computer Games. ArXiv:1608.02192 [cs.CV].