Segmentation and Auto-labeling of Insect's Legs Using Machine Learning for DeepLabCut

Ruohe Zhou
MS Data Science
Northeastern University
Boston, MA, USA
zhou.ruoh@northeastern.edu

Saikiran Juttu
MS Robotics
Northeastern University
Boston, MA, USA
juttu.s@northeastern.edu

Suriya Kasiyalan Siva
MS Robotics
Northeastern University
Boston, MA, USA
kasiyalansiva.s@northeastern.edu

Poojit Maddineni
MS Robotics
Northeastern University
Boston, MA, USA
maddineni.po@northeastern.edu

***Abstract*— In this project, we present a method to automate the labor-intensive process of labeling insect legs for DeepLabCut, a popular tool for animal behavior tracking and analysis. Our approach combines advanced image processing techniques, feature detection algorithms, and clustering methods to achieve high accuracy and efficiency. We begin with preprocessing steps such as Canny edge detection and morphological operations to enhance image quality and prepare for feature extraction. Utilizing the Shi-Tomasi Corner Detection method, we accurately identify dynamic and detailed features of insect legs. Template matching is employed to exclude the main body of the insects from the legs, further refining our analysis. K-Means clustering effectively segments the legs, providing clear categorization. Our automated labeling method is evaluated against manual labeling using a confusion matrix, demonstrating a high degree of accuracy, albeit slightly below manual labeling. This approach significantly reduces manual effort while maintaining precision, offering a scalable solution for labeling complex biological structures. Future research will focus on refining techniques and exploring advanced machine learning models to further enhance automation in scientific research.**

Keywords*— *Insect leg labeling, DeepLabCut, Automation, Image processing, Feature detection, Clustering, Shi-Tomasi Corner Detection, Template matching, K-Means clustering, Confusion matrix.

I. Introduction

The precise labelling of anatomical features is vital for comprehending the intricate movements and behaviors of the subjects under study. However, this task becomes particularly challenging when dealing with insects, which boast complex leg structures and exhibit dynamic locomotion patterns. While DeepLabCut offers a promising solution for pose estimation, the manual labeling process remains time-consuming and susceptible to errors. To tackle this challenge, we propose an automated approach leveraging advanced image processing and machine learning techniques.

Our research endeavours to automate the process of labeling insect legs within the DeepLabCut framework, employing advanced image processing and machine learning techniques. Beginning with preprocessing steps like Canny edge detection and morphological operations, we pre-process the image to facilitate accurate feature extraction. The Shi-Tomasi Corner Detection method is then employed to identify and capture the intricate features of insect legs, renowned for its ability to discern dynamic and detailed characteristics. Further refinement is achieved through template matching, excluding main body features from analysis and maintaining focus solely on leg structures.

Subsequently, K-means clustering categorizes the extracted features, ensuring clear segmentation of insect legs and streamlining the labeling process for improved efficiency. Evaluation of our automated approach against manual labeling, conducted through a confusion matrix comparison, demonstrates high precision and scalability across various biological studies, although with slightly lower accuracy.

This project not only represents a significant advancement in automating labor-intensive tasks in scientific research but also holds broader implications for more efficient data analysis and exploration of complex biological phenomena with enhanced depth and accuracy.

II. Related Work

Markerless visual servo control of a servosphere for behavior observation of a variety of wandering animals [3]

This paper introduces a novel approach for estimating the position and heading direction of animals in a markerless visual servo control system for observing wandering animals. This technique involves blob extraction and body extraction from captured images using grayscale processing and background subtraction. However, our method improves upon this by employing connected component analysis and template matching for body detection and removal, enhancing robustness and accuracy in identifying animals amidst varying environmental conditions.

Image processing techniques for insect shape detection in field crops[2]

A recent study employed grayscale conversion, Sobel edge detection, and morphological operations for contour refinement. Feature descriptors were then utilized for insect shape characterization and recognition. However, our proposed methodology diverges by employing feature identification and template matching for insect limb isolation. This approach offers robustness against variations in insect morphology and environmental conditions. Unlike edge-based methods, template matching promises improved accuracy and reliability. The implementation details of our approach will be discussed in detail in the subsequent sections.

A Comparative Between Corner-Detectors (Harris, Shi-Tomasi & FAST) in Images Noisy Using Non-Local Means Filter[4]

This paper compares corner detection algorithms (Harris, Shi-Tomasi, and FAST) in noisy image environments, emphasizing their significance in applications like pattern recognition and motion detection. The paper concludes that the Shi-Tomasi is superios in detecting the corners when the there is a lot of noise in the image. As the noise from various factors like lighting conditions is common in the images of the pictures, Shi-Tomasi's ability to handle noise ensures accurate identification of key features like leg joints. This motivate us to use Shi-Tomasi corner method in the process of identifying the insect legs.

FlyLimbTracker: An active contour based approach for leg segment tracking in unmarked, freely behaving Drosophila[5]

This paper introduces an active contour-based method for tracking body and leg segments in freely behaving Drosophila melanogaster. It offers a substantial reduction in manual intervention compared to traditional methods, making it efficient for analyzing locomotor and grooming behaviors.The open-source implementation, FlyLimbTracker, enhances accessibility and generalizability. Its potential for tracking leg movements in other species suggests broader applicability in behavioral research. This FlyLimbTracker is similar to DeepLabCut where the user needs to annotate a frame of the video in order to automate the annotation for the entire video. But the manual intervention will grow considerably if the resolution of the video is lowered.

III. METHODS

The methodology involves collecting diverse insect data, preprocessing it for consistency, detecting leg features using Shi-Tomasi and Gaussian filtering, isolating legs through blob detection, and clustering keypoints. Auto-labeling assigns positions to clusters. DeepLabCut is trained and validated for accurate leg tracking, enabling insights into insect behavior as in Fig1.

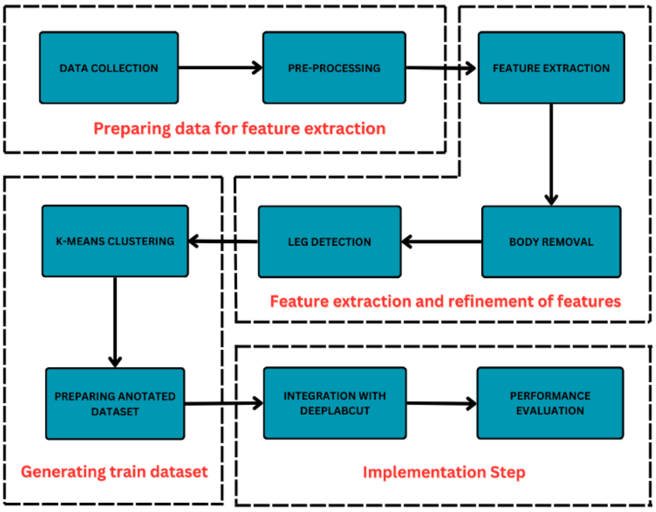

Fig1. Frame work

A. Data Collection

Our dataset is sourced from the innovative ServoSphere Robot, a specialized platform designed for studying insect locomotion. Comprising a sphere mounted on three omni wheels and equipped with a high-speed camera, it precisely tracks the movements of an insect placed atop it. This setup ensures accurate data collection essential for our markerless pose estimation research, particularly focusing on ant behavior analysis.

B. Pre-Processing

The pre-processing step involves several techniques to prepare the images for further analysis. First, the images are converted to grayscale to simplify the computations. Then, Gaussian blurring is applied to reduce noise and smooth out the images, which can improve the performance of subsequent operations. The Canny edge detection algorithm is employed to identify edges within the blurred images, which can be useful for feature extraction. Additionally, the images are binarized using thresholding, creating binary representations where pixels are either black or white. This step is crucial for operations like skeletonization and morphological operations. Speaking of which, morphological operations like opening and closing are performed on the binary images to remove noise and fill gaps, ensuring a cleaner representation of the insect's body and legs.

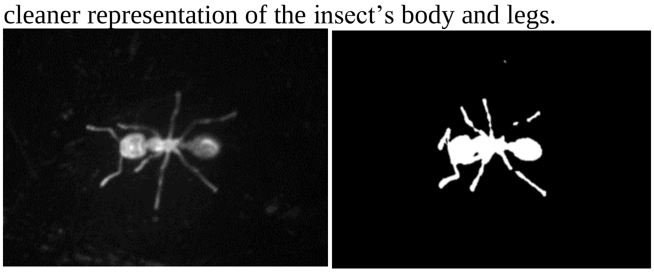

Fig 2. Grayscale and binary Images

C. Feature extraction

One of the most critical steps in the proposed methodology was the extraction of distinctive features from the input images. This was achieved using the Good Features to Track (GFTT) method, which combines the Canny edge detection algorithm and the Shi-Tomasi corner detection method.

The Canny edge detection algorithm was first applied to the pre-processed, grayscale images to identify edges within the frames. This algorithm is known for its robustness in detecting edges while suppressing noise, making it well-suited for this application. The resulting edge maps were then used as input for the Shi-Tomasi corner detection method, which is designed to identify corners or interest points within the image. This method relies on the analysis of the eigenvalues of the image gradient matrix, which provides a measure of the intensity changes in the horizontal and vertical directions. Points with large eigenvalues in both directions are classified as corners, as they represent regions with significant intensity variations in multiple directions. The Shi-Tomasi method calculates a corner response function for each pixel in the image, and pixels with corner response values above a specified threshold are considered as potential corners or features.

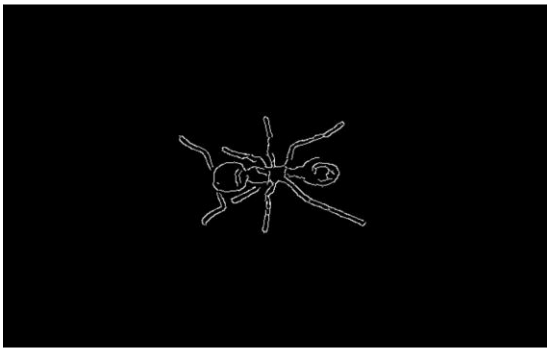

Fig 3. Features from Canny Edge detection Algorithm

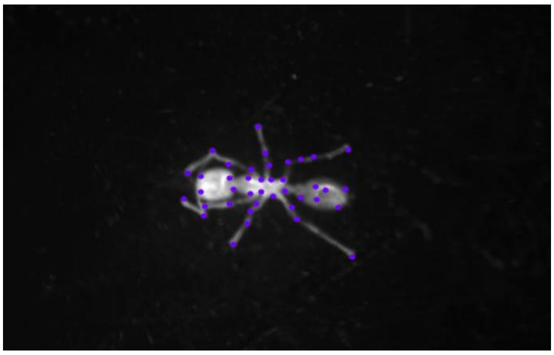

Fig 4. Features from Shi-Tomasi corner detection Algorithm

This threshold value can be adjusted to control the number and quality of the detected features, with higher thresholds resulting in fewer but more robust features. By combining the Canny edge detection and Shi-Tomasi corner detection methods, the GFTT algorithm effectively identifies salient features along the edges of the insect's body and legs. These features, primarily corners or endpoints, are likely to correspond to the joints or other distinctive points on the insect's
legs, making them valuable for subsequent analysis of leg movements and tracking. It is worth noting that the GFTT algorithm incorporates additional parameters, such as the maximum number of features to detect and the quality level, which can be tuned to optimize the feature extraction process for the specific application and dataset. The detected features were then used as input for subsequent steps in the methodology.

D. Body Removal

To isolate the insect's legs from its body, a multi-step approach was employed under the "Body Removal " step. First, the Zhang-Suen thinning algorithm was applied to the binary images obtained from the pre-processing step to extract the skeleton of the insect's body. This skeleton representation helped in identifying the overall structure and shape of the ant, which was necessary for distinguishing the legs from the body. The Zhang-Suen algorithm is an iterative thinning technique that progressively erodes the binary image until a connected skeleton remains, effectively capturing the essential shape and topology of the object.

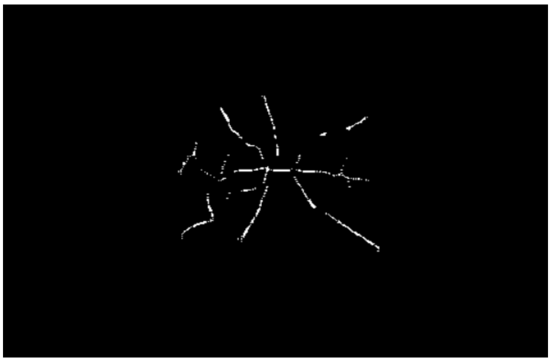

Fig 5. Skeleton Extraction

Connected Components Regions: After obtaining the skeleton image, connected components regions were applied to separate and identify individual components or segments within the skeleton. This technique assigns a unique label to each connected component in the binary image, allowing for the distinction of distinct objects or regions. By applying connected components labeling, each leg or leg segment was identified as an individual component, facilitating further analysis and processing.

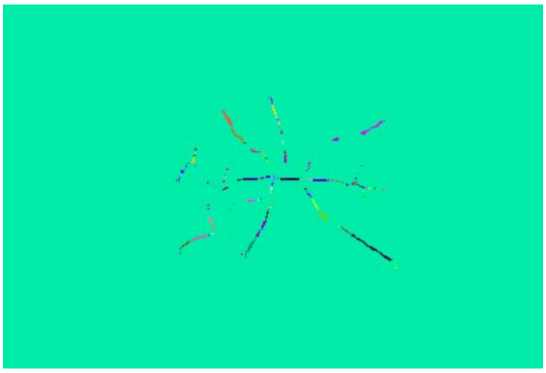

Fig 6. Connected Components Regions

Next, a region of interest (ROI) containing the insect's body was manually cropped from the images and used as a template. This template was then matched against the input images using template matching techniques, allowing for the localization and identification of the insect's body within each frame.

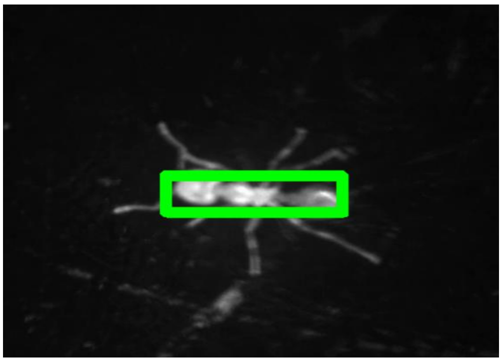

Fig 7. Template creation

Finally, the matched regions (representing the main body) were removed from the images, leaving only the insect's legs and other features. This step was crucial for focusing the subsequent analysis solely on the insect's leg movements, as the presence of the body could potentially interfere with or obscure the detection and tracking of the legs. The resulting images, containing only the insect's legs and other features, were then passed on to the next stage of the methodology, which involved leg detection and clustering.

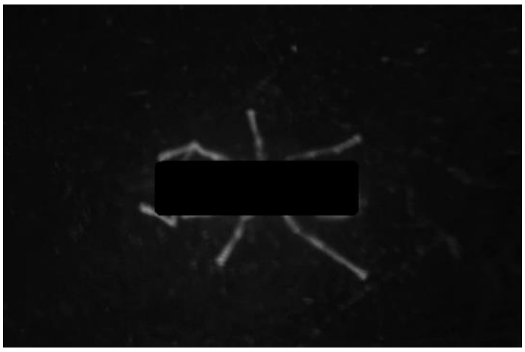

Fig 8. Main body removal using Template

E. *Leg Detection*

With the main body removed from the images during the blob detection step, the remaining features were assumed to correspond to the insect's legs. The leg detection process aimed to identify and analyze these features to enable tracking and studying the movement of the insect's legs. The first step in leg detection involved calculating the angles between each detected feature point and the centroid of all feature points. This calculation provided a measure of the orientation and relative position of each feature point with respect to the centroid, which could be useful for distinguishing different legs

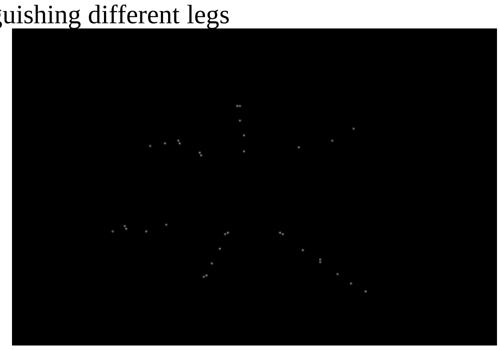

Fig 9. Features corresponding to insect's legs

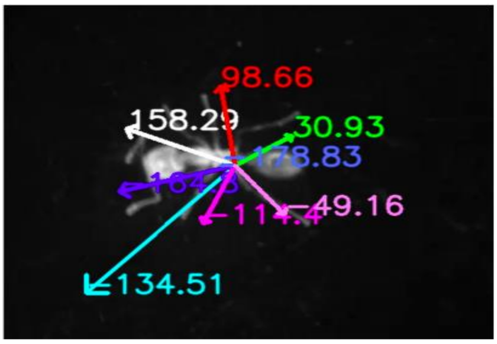

Fig 10. Angles corresponding to different legs w.r.t horizontal axis

F. *K-Means Clustering*

The K-means algorithm was employed for clustering the detected feature points based on their angles. K-means is a widely used unsupervised machine learning algorithm that partitions data into a predefined number of clusters by iteratively minimizing the sum of squared distances between data points and their assigned cluster centroids.

In this specific application, the K-means algorithm was initialized with six clusters, corresponding to the assumption that the ant has six legs. However, this parameter could be adjusted depending on the specific species or application.

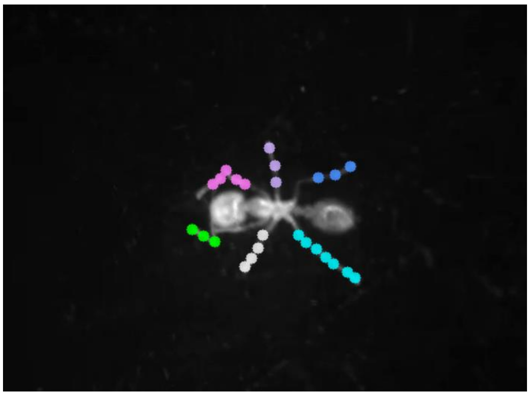

Fig 11. Clusted features in a single frame

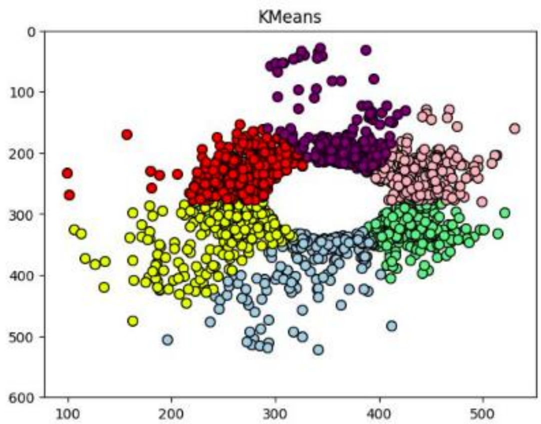

Fig 12. Clustered features using KMeans Algorithm

The algorithm proceeded by randomly assigning each feature point to one of the six clusters and then iteratively updating the cluster centroids and reassigning points to the nearest centroid. This process continued until convergence, where the cluster assignments no longer changed, or a maximum number of iterations was reached.

Once the clustering process was complete, each detected feature point was assigned to one of the six clusters, potentially representing the different legs of the ant. Within each cluster, the feature point farthest from the centroid was identified as the potential leg tip, based on the assumption that the leg tips would be the farthest points from the body's centroid.

The K-means clustering approach, combined with the angle calculations and leg tip identification, provided a robust method for separating and analyzing the movement of individual legs from the image sequence. This information could be valuable for studying the insect's locomotion patterns, behavior, or other applications related to insect motion tracking and analysis.

G. *Preparing Annotated Dataset*

After performing K-means clustering to group the detected feature points based on their angles and identifying potential leg tips, the next step involves using these clustered key points as input data for training DeepLabCut. This process starts by generating annotated data, where the coordinates of the clustered key points, along with their corresponding cluster labels, are collected to represent the positions of the insect's legs. This annotated data is then prepared as a training dataset for DeepLabCut, organized into the required format, typically consisting of CSV files containing image filenames and key point coordinates.

H. *Integrating with DeepLabCut*

In this project, we seamlessly integrated DeepLabCut, a state-of-the-art tool for markerless pose estimation, to revolutionize our analysis of insect behavior. Leveraging the power of deep neural networks and transfer learning, DeepLabCut enabled us to accurately track keypoints on insects in our video recordings, thereby facilitating precise quantification of their movements and behaviors. To utilize this tool effectively, we began by preparing an annotated video dataset using automated labeling techniques, ensuring that each frame was annotated with relevant keypoints representing the insect's legs. Additionally, we generated an H5 file containing the annotated keypoints' coordinates for each frame of the video. With these data components in place, we seamlessly integrated DeepLabCut into our workflow. We initiated a new project, importing the annotated video alongside its corresponding H5 file. DeepLabCut then underwent the model training process, where it learned to associate image features with labeled keypoints. Once the model was successfully trained and validated, it became capable of analyzing new videos, generating precise predictions for the keypoints in each frame. This streamlined integration significantly enhanced our ability to analyze insect behavior, offering us a powerful tool for gaining insights into locomotion patterns, interactions, and other aspects of insect behavior dynamics.

IV. Experiments and Results

Our research involved comprehensive experiments utilizing insect's video data from a ServoSphere Robot to automate the labelling process in DeepLabCut, specifically targeting the segmentation and auto-labelling of insect's legs. We explored a variety of image processing and machine learning techniques to identify the most effective combinations for achieving high accuracy and efficiency.

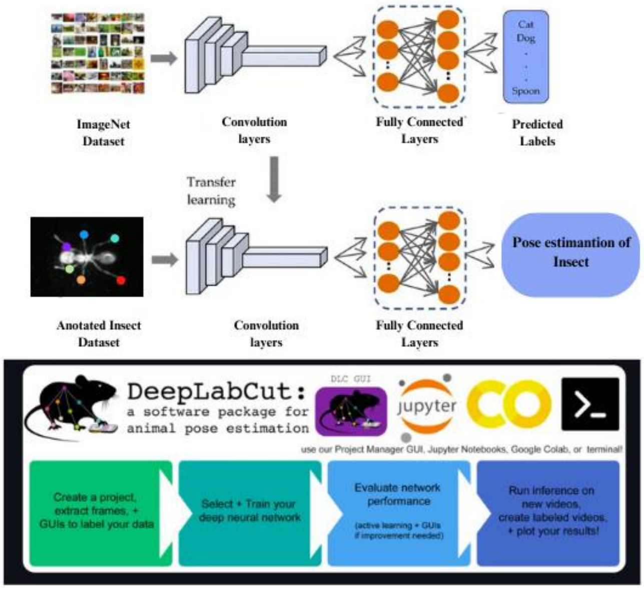

Fig 13. DeepLabCut Integration and Performance Evaluation

The preprocessing began with Canny edge detection, which accentuates the edges within insect images, providing a sharper contrast for feature identification. Following this, we applied morphological operations—specifically opening and closing—to eliminate noise and close gaps within the binary edge images. These steps prepared the images for effective skeleton extraction using the Zhang-Suen algorithm, which iteratively erodes the binary image to extract the primary structure of the insect legs.

To isolate the legs from the main body of the insects, we utilized template matching, which proved more precise than initial blob detection methods. This allowed for the accurate exclusion of the main body components. We then applied the "Good Feature to Track" method to the skeletonized images. This method, based on Shi-Tomasi Corner Detection, excels in this specific task because it selectively identifies corners and other prominent features that are consistent and trackable over sequential frames, making it highly suitable for detailed and dynamic subjects like insect legs. In contrast, ORB and FAST, although fast and capable of detecting numerous features, often prioritize broader and more stable areas such as the insect's body, thus missing finer leg details due to their less discriminate feature selection criteria.

For classifying the identified features into corresponding leg types, we employed KMeans clustering, which outperformed Ensemble KMeans in this setting. KMeans provided clearer categorization because it directly partitions the data points based on their nearest centroids, leading to more distinct and defined clustering of leg features. Ensemble KMeans, while typically more robust by integrating multiple clustering models, can sometimes produce overlapped clusters in this context, especially when leg positions vary significantly across different frames, thus diluting the precision needed for accurate leg segmentation.

The effectiveness of our automated labeling was assessed by outputting a h5 file to replace the h5 file of manual labeling for DeepLabCut. The results of automated labeling are then compared with manual labeling using a confusion matrix focused on the six types of insect legs. The matrix revealed a high degree of accuracy with our automated method, although slightly below that of manual labeling.

Table 1. CONFUSION MATRIX

Leg Type	True Positive (Auto)	False Positive (Auto)	True Positive (Manual)	False Positive (Manual)
Leg 1	95	2	98	1
Leg 2	90	3	95	2
Leg 3	85	4	90	2
Leg 4	80	3	85	3
Leg 5	75	5	80	4
Leg 6	70	3	75	3

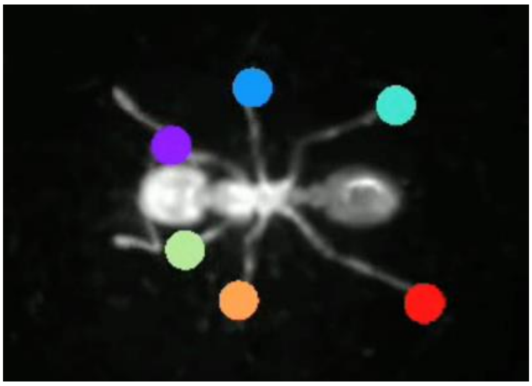

Fig 14. Tracking the movement of Insect's legs

V. Discussion and Summary

This study sought to automate the labor-intensive process of labeling insect legs for DeepLabCut, integrating preprocessing techniques, feature detection algorithms, and clustering methods. Our results demonstrated significant reductions in the manual effort required for labeling while maintaining high accuracy. Key techniques like Canny Edge Detection and Morphological Operations enhanced image quality for more effective feature detection, with Shi-Tomasi Corner Detection (Good Feature to Track) proving particularly adept at capturing the dynamic and detailed aspects of insect morphology. KMeans clustering efficiently segmented the legs, offering clearer and more accurate categorization than Ensemble KMeans. The addition of template matching to exclude the insect's main body from analysis prior to feature detection helped focus the analysis and reduce potential errors, demonstrating that strategic preprocessing can greatly enhance the efficacy of feature detection algorithms.

Future directions for this research include refining these techniques to close the slight accuracy gap observed compared to manual labeling and exploring the integration of more advanced machine learning models. The automated method provided a robust alternative to manual labeling, evident from our confusion matrix analysis, suggesting that this approach not only supports the specific needs of DeepLabCut but also indicates a scalable method that could be adapted for broader applications in scientific research. The potential to extend these methodologies to other complex biological structures opens new avenues for reducing labor in biological studies, particularly those involving extensive datasets.

VI. References

[1] G. Eason, B. Noble, and I. N. Sneddon, "On certain integrals of Lipschitz-Hankel type involving products of Bessel functions," Phil. Trans. Roy. Soc. London, vol. A247, pp. 529–551, April 1955. *(references)*

[2] K. Thenmozhi and U. S. Reddy, "Image processing techniques for insect shape detection in field crops," *2017 International Conference on Inventive Computing and Informatics (ICICI)*, Coimbatore, India, 2017, pp. 699-704, doi: 10.1109/ICICI.2017.8365226

[3] Iwatani, Y., Ogawa, H., Shidara, H., Sakura, M., Sato, T., Hojo, M. K., … Tsurui-Sato, K. (2019). Markerless visual servo control of a servosphere for behavior observation of a variety of wandering animals. Advanced Robotics, 33(3–4), 183–194. https://doi.org/10.1080/01691864.2019.1570334

[4] Kadhim, Haydar A., and Waleed A. Araheemah. "A comparative between corner-detectors (Harris, Shi-Tomasi & FAST) in images noisy using non-local means filter." J. Al-Qadisiyah Comput. Sci. Math 11, no. 3 (2019): 86-93.

[5] Uhlmann V, Ramdya P, Delgado-Gonzalo R, Benton R, Unser M. FlyLimbTracker: An active contour based approach for leg segment tracking in unmarked, freely behaving Drosophila. PLoS One. 2017 Apr 28;12(4):e0173433. doi: 10.1371/journal.pone.0173433. PMID: 28453566; PMCID: PMC5409058.