Towards Robotic Mapping of a Honeybee Comb

Jiří Janota[1], Jan Blaha[1], Fatemeh Rekabi-Bana[2], Jiří Ulrich[1], Martin Stefanec[3]
Laurenz Fedotoff[3], Farshad Arvin[2], Thomas Schmickl[3] and Tomáš Krajník[1]

***Abstract*— Honeybees are irreplaceable pollinators with a direct impact on the global food supply. Researchers focus on understanding the dynamics of colonies to support their health and growth. In our project "RoboRoyale", we aim to strengthen the colony through miniature robots interacting with the honeybee queen. To assess the colony's health and the effect of the interactions, it is crucial to monitor the whole honeybee comb and its development. In this work, we introduce key components of a system capable of autonomously evaluating the state of the comb without any disturbance to the living colony. We evaluate several methods for visual mapping of the comb by a moving camera and several algorithms for detecting visible cells between occluding bees. By combining image stitching techniques with open cell detection and their localization, we show that it is possible to capture how the comb evolves over time. Our results lay the foundations for real-time monitoring of a honeybee comb, which could prove essential in honeybee and environmental research.**

I. Introduction

Western honeybees (*Apis mellifera*) play a pivotal role in our ecosystem, acting as indispensable pollinators, directly impacting the food web and contributing to the biodiversity of flowering plants. Thus, monitoring and maintaining healthy bee populations is essential for both ecological balance and food security.

Not all behavioral strategies and self-organization mechanisms of the bees are fully understood. Interactions with other species, parasites, or pathogens add further complexity. However, advancements in robotics and artificial intelligence have made it possible to develop systems for animal-robot interaction in their natural ecosystems. Recently, researchers presented the potential of using fixed-camera systems capable of long-term observation of honeybee queen activity [1] and tracking of individual bees [2], [3].

In the EU-funded project "RoboRoyale", we aim to support the efficiency and growth of honeybee colonies by using a robotic system to interact with and support the honeybee queen. As the central element of the colony, she is solely responsible for practically all egg-laying activities and, thus, for the growth of the colony. Hence, a robotic manipulator was designed for observing the colony and the honeybee queen itself with a high-resolution moving camera [4]. In the future, the robotic manipulator will also carry a mechanism that will interact with the honeybee queen.

Apart from the honeybee queen tracking, we can observe the rest of the hive by taking images at various levels of detail anywhere in the hive. Because the state of the whole hive is important to assess the colony's health and the queen's behavior, the system is also tasked with creating and updating a map capturing the contents of the underlying comb. The whole comb can be sequentially scanned with a grid of partially overlapping image tiles taken at different locations. Such scans can be directly utilized to create the map with the comb cells.

This work presents key steps towards autonomous mapping of the dynamic honeybee comb. Identifying all kinds of individual cells in the comb images can be a challenging task even for a human annotator, and thus, creating a dataset for direct object segmentation would be a complicated process. Instead, we use the fact that the state of the capped cells does not change until the cell is opened again; thus, we only need to identify the open ones. We developed a system for uncapped cell detection and mapping, which can track the evolution of individual cells over time. The proposed pipeline combines the odometry information from the motors with image registration to stitch the grid of overlapping image tiles, creating a snapshot of the entire comb. Then, we detect the uncapped cells not occluded by bees and determine their position on the comb, creating a map of uncapped cells. By analysing a sequence of such maps gathered over time, we capture the evolution of individual cells in Fig. 1. Such map dynamics can provide important indicators of colony health (e.g., growth rates) and act as a foundation for deeper understanding. In future work, we will develop algorithms for cell-type classification and integrate the gathered spatial maps into a semantic spatio-temporal representation of the whole comb. That will allow us to monitor the state of the living colony over time, assessing the results of the robot-queen interaction.

II. Related works

This section presents state-of-the-art research in relevant areas preceding our work. Later, we compare the methods for stitching a grid of partially overlapping images reported in the related works and show their limitations when applied in dynamic and cluttered environments with highly repetitive patterns. We also address the shortcomings of commonly used Circle Hough Transform for cell detection in comb images containing bees by employing deep learning methods.

[1] Faculty of Electrical Engineering, Czech Technical University in Prague, Czechia name.surname@fel.cvut.cz

[2] Durham University, Computer Science Department, United Kingdom name.surname@durham.ac.uk

[3] Artificial Life Lab, Department of Zoology, Institute of Biology, University of Graz; Graz, Austria name.surname@uni-graz.at

The code for the experiments is available at: https://gitlab.roboroyale.eu/janota/2024_marss_towards_mapping

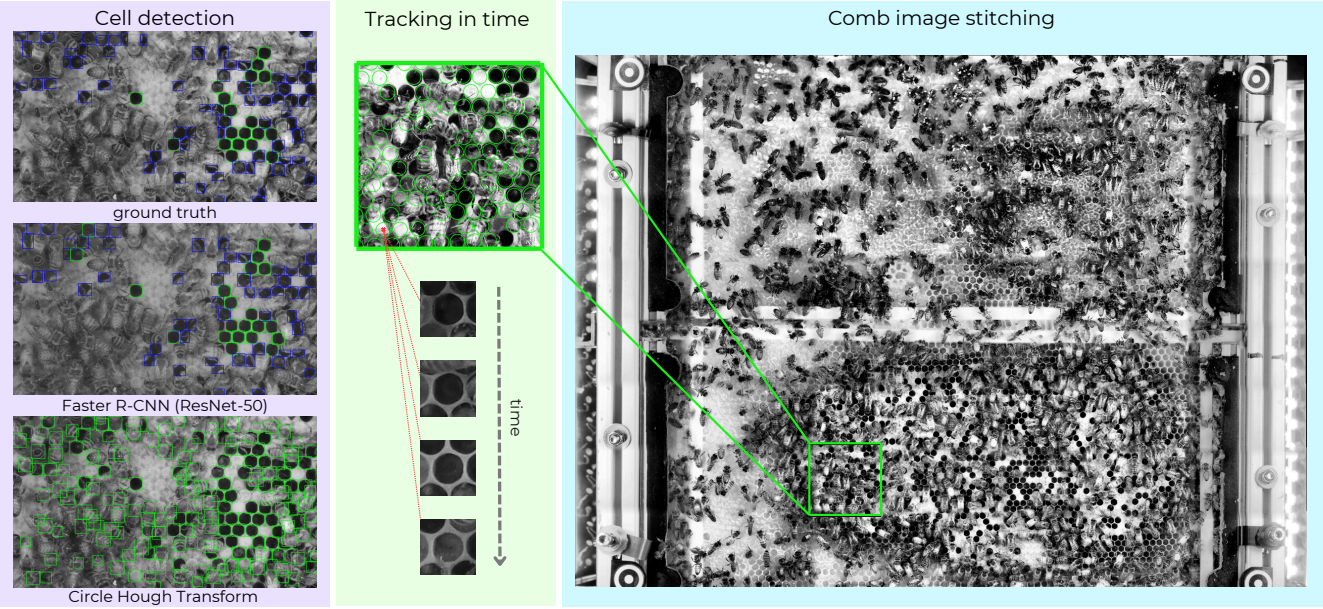

Fig. 1: Visualization of key components of our system: **(Left)** Cell detection with Circle Hough Transform and Faster R-CNN (ResNet-50 backbone) together with the ground truth (fully visible cells are highlighted in green, occluded cells in blue). **(Middle)** Tracking of cell detections in time with the development of a single cell, demonstrated on a single image tile from the whole stitched comb **(right)**.

A. *Image stitching*

Image stitching is a technique used to combine multiple images with overlapping fields of view into a single composite image. For a camera parallel to the observation plane, the full registration problem reduces to the estimation of horizontal and vertical translation. There are two prevalent approaches for image registration in the context of stitching grids of image tiles: the feature-based approach and the direct approach in a sliding window manner.

1) Feature-based image registration: The feature-based approach relies on a feature detector to identify key points of interest, which are then described using descriptors and matched between images based on their descriptor similarity. The final transformation is estimated from the matched correspondences with the Random Sample Consensus (RANSAC) algorithm as in [5]–[9].

A predominant choice for a feature detector in many studies [5], [6], [8], [9] is the Scale-Invariant Feature Transform (SIFT), introduced by Lowe [10]. However, the particular choice of feature detector depends on the specific application, and the features, in some cases, need to be adjusted for the task [7].

Lately, machine learning (ML)-based methods are often introduced. Particularly noteworthy is the work in [11], where authors introduce Superpoint, a CNN-based feature detector and descriptor. Additionally, [12] demonstrates the feasibility of using Graph Neural Networks for matching features. Researchers in [13] introduced Detector-Free Local Feature Matching with Transformers (LoFTR) to address the challenge of separate neural networks for feature detection and matching. Notably, [9] showcases the significant potential of using LoFTR instead of SIFT in the context of electron microscopy and image tile stitching.

2) Direct image registration: The direct approach to image registration aims to find the transformation by maximizing a similarity metric across all possible translations. Instead of using a computationally complex sliding-window approach, researchers often employ more efficient techniques like Phase Correlation (PC) [14]–[20] and evaluate a few highest peaks from the PC based on Normalized Cross-Correlation (NCC) values. In [15], the authors combine both and first applied PC on downscaled images and then obtained the precise solution using a sliding-window approach on the original unscaled images.

3) Global optimization: The methods introduced for image registration are designed to estimate transformations between individual pairs of images. However, the necessity for global optimisation methods arises when dealing with a grid of image tiles. These methods aim to minimize misalignment across all pairs of overlapping images, consequently mitigating the accumulation of image registration errors during the stitching process.

In the study by [5], the optimal rigid transformation is achieved by iteratively minimizing the square displacement of landmarks identified by SIFT. Another common strategy involves constructing a graph with image tiles as nodes and connecting edges between overlapping pairs of images. Many studies [18], [20], [21] employ the minimum spanning tree algorithm to identify the optimal subset of edges that connect all tiles and minimize the resulting misalignment. Additionally, [9] proposes to perform the optimization with graph-based 2D Simultaneous Localization and Mapping (SLAM)

method GraphSLAM. Another widely adopted technique is to construct an over-constrained system of linear equations and minimize the sum of all pairwise transfer errors through least squares [15], [16] or weighted least squares [14], [17], [19], [20], employing correlation values as weights.

B. Honeybee cell detection

Honeybee cell detection is an actively researched field, mainly motivated by automated health assessment and the threat posed by the *Varroa destructor* mite, an external parasite that significantly impacts bee colonies by causing malformation and weakening of the colony, as well as transmitting viruses. To address this, researchers explore alternative methods like assessing dead brood removal rates as an indicator of colony resistance to Varroa mites [22]. The hygienic behavior of bee colonies is traditionally examined manually by beekeepers, which is a labor-intensive process [23], [24]. In response, several studies aim to reduce the workload by automating the detection of uncapped cells or directly identifying brood cells. For instance, in [25], the authors utilize the Canny edge detector to identify contours and then employ a set of features for classifying uncapped cells, allowing them to track back the uncapping events in recordings. Another study, detailed in [26], opt for a convolution-based method with a circular mask of cell size to detect brood cells. Authors of [27] propose detecting the individual cells by image thresholding and partitioning the image into so-called superpixels.

Many researchers leverage the fact that the cells are of a circular shape and don't overlap. Typically, they start with preprocessing steps to suppress noise and normalize the image and then employ Circle Hough Transform (CHT) [28] for individual cell detection [29]–[34]. Additionally, [32] extends the CHT with the use of U-Net, a Convolutional Neural Network (CNN) for binary segmentation, to filter out cell detections outside the honeybee comb. In contrast to traditional computer vision methods for cell detection, the authors of [2] embrace a deep learning approach and use a U-Net neural network.

III. Methods

A. System Description

Investigating the honeybee colony and developing a spatially consistent map from the comb demands an autonomous observation mechanism to collect information from the hive continuously. In our work, we use a robotic system conceptually presented in [4], which allows moving the end effector parallel to the comb plane. Observations are captured with a camera, which is mounted on the end effector of the mechanism and provides high-resolution 1920 px $\times$ 1080 px images at a rate of 30 Hz with controllable zoom and focus. The system can select camera zoom to provide images with resolutions of 16 - 67 μm per pixel, which is sufficient for reliable detection of small objects like 1 mm $\times$ 0.3 mm honeybee eggs. The construction of the observation hive commonly used in research on honeybees is shown in Fig. 2.

Fig. 2: Observation hive, commonly used in biology to study honeybee behavior, courtesy of [35]. Standard combs are placed in a wooden construction behind a glass panel, and the hive is connected to the outside by a plastic tube entrance. All observations are done under near-infrared light to minimize disturbance of the colony.

B. Comb image stitching

In order to generate a metric map, the overlapping image tiles must be precisely aligned. While the robotic system provides odometry data for each image tile, any inaccuracies in this information can jeopardize the entire system's performance. Therefore, to enhance the odometry's accuracy, we evaluated the performance of both correlation-based and feature-based image registration, as each approach has different advantages and disadvantages. The collected images vary a lot in exposure and illumination, and the overlaps significantly differ in terms of content due to the movement of the bees, which can pose difficulties to the correlation. On the other hand, the feature-based approach may suffer from the lack of unique features, as the bees and the cells produce highly repeatable patterns.

1) Correlation-based approach: As mentioned in section II-A, the correlation-based method involves assessing all potential translations between images. To narrow down the search space, we leveraged the odometry as prior information and limited the deviation to 20 pixels in both the x and y axes. The similarity metric was then evaluated solely on the estimated overlapping parts of the images. Addressing challenges related to differing brightness and exposure, we employed normalization and histogram equalization, utilizing the Normalized Cross-Correlation Coefficient (NCC) as the similarity metric. The NCC for two images I_1 and I_2 with mean $\bar{I_1}$ and $\bar{I_2}$ is defined as:

$$NCC = \frac{\sum_{x,y}(I_1 - \bar{I_1})(I_2 - \bar{I_2})}{\sqrt{\sum_{x,y}(I_1 - \bar{I_1})^2}\sqrt{\sum_{x,y}(I_2 - \bar{I_2})^2}} \quad (1)$$

2) Feature-based approach: In the feature-based approach, we use SIFT [10] for keypoint detection and descrip-

tion. Similarly to the correlation-based approach, we apply normalization and histogram equalization on the images and then focus the feature detection solely on the estimated overlapping areas, constraining the potential deviation from the odometry to 20 pixels in both the x and y axes. Additionally, since the images do not vary in scale, we match two keypoints together only if the ratio of their scales is less than 1.5. We employ a simple histogram voting approach to determine the translation between images, where the transformation with the most votes is selected.

3) Global optimization: To address misalignment issues across all pairs of image tiles, we use error minimization using a least squares optimization. We construct a graph $G = (V, E)$, where V represents the positions of individual image tiles $\mathbf{p}_i = (x, y)_i$, and E is a set of edges connecting pairs of neighboring image tiles. To formulate an over-constrained system of linear equations, we utilize the estimated translations $\mathbf{d}_{ij}$ between image pairs and anchor the position of the first tile to origin $(0, 0)$. This gives us a problem of the form

$$\min_{\forall i:\, \mathbf{p}_i} \sum_{(i,j)\in E} \sum_{c\in\{x,y\}} \left((\mathbf{p}_{i,c} - \mathbf{p}_{j,c}) - \mathbf{d}_{ij,c}\right)^2 \quad (2)$$
$$\text{s.t.} \quad \mathbf{p}_1 = (0, 0).$$

While there is an option to weigh the contribution of individual estimated translations by confidence indicated, for example, by cross-correlation values, we chose not to employ this measure due to its sensitivity to the randomness of bees' motion in our case.

C. Cell detection

1) Circle Hough Transform: The Circle Hough Transform (CHT) [28] is an algorithm for detecting circles in images. As mentioned in section II-B, CHT has been commonly used by many researchers to detect uncapped honeybee cells, thus making it a reasonable baseline.

In the preprocessing stage, we applied Contrast Limited Adaptive Histogram Equalization (CLAHE) for image normalization and a Bilateral Filter for noise suppression. We then used the OpenCV [36] implementation of the CHT for circle detection. The parameters were determined experimentally and are summarized in Table I.

TABLE I: Parameters of the Circle Hough Transform

Parameter	**Value**
minimum distance between circles	40
minimum radius of the circles	30
maximum radius of the circles	45
Canny edge detector threshold	40
accumulator threshold	18

2) Object detection with neural networks: As usually reported, the CHT approach tends to produce false positives, which can then be further filtered out, for example, by restricting the area of the comb in the image [32]. Similarly, it could be possible to classify each of the circles detected by the CHT with classifiers, such as a compact neural network, in an R-CNN style, where the CHT would act as a region proposal algorithm. As this approach could significantly affect the possible reachable recall of the method already in the stage of the region proposal, we trained two different commonly used object detection models instead.

First, we adopted a two-stage object detection architecture, Faster R-CNN [37], implemented in Python TorchVision package [38]. We experimented with two backbones—ResNet-50 and ResNet-18 [39] both with additional feature pyramid network (FPN)—pre-trained on the COCO dataset [40]. All models were first trained with a frozen backbone and then fine-tuned in entirety, the details are in the provided code repository. For its faster inference, a single-stage YOLOv5s6 model from Python Ultralytics package [41] was then included as well with a similar training procedure.

To address the problem of varying exposure and illumination, we employed image augmentations such as brightness and contrast changes, defocus, and gamma correction. To prevent overfitting, we also used horizontal and vertical flips.

3) Cell accumulation: To be able to get the temporal evolution of individual cells, we first needed to be able to associate detections from different scans. Given the metric positions of the detections from the odometry and the correction by estimated translation, we establish correspondences between the new detections and the map using the RANSAC algorithm. We resort to a sampling algorithm, as for any given scan, a substantially different set of cells can be detected, which can lead to a small intersection between all of the map and all of the detections. In the algorithm, we consider candidate associations to be all map cells in the radius of 20 px from the center of the detection. If no match is found for the new detection in the map, it is entered in the map as a new cell.

IV. Datasets

In our experiments, we use a dataset gathered with the robotic gantry system in the observation season of 2023. The dataset consists of full comb scans, which form a regular rectangular grid of partially overlapping images on both sides with a camera resolution of 67 μm per pixel, each taken in a different location.

To evaluate the accuracy of both the odometry and the proposed image stitching methods, we created a dataset of randomly sampled image pairs along with corresponding translation information. The odometry served as prior information in determining these translations, which were then manually refined. Due to modifications in the robotic system during the experiments, two distinct datasets were created. The first dataset, *IS1*, consists of 82 comb image pairs with imprecise odometry information. The second dataset, *IS2*, comprising 176 pairs, has more precise odometry measurements (see section V).

For training the object detection models and determination of optimal parameters for the CHT, we annotated a dataset in a semi-automatic way comprising images with cells specified as bounding boxes. To ensure diversity in the dataset, the

image tiles were randomly selected from the collected data across different days and parts of the comb.

Initially, we applied the Segment Anything Model (SAM) [42] with the ViT-H model for image segmentation. Subsequently, we refined the resulting masks by implementing a simple filter based on the area and circularity of the individual masks, and finally, we manually annotated the pre-filtered segmentations. The cells were categorized into two classes, fully visible and partially occluded, to allow easier extension for subsequent analysis of the cells' type.

Using this annotation tool, we annotated a set of 260 honeybee comb images, which were split into training (200 images), validation (30 images) and testing (30 images) parts.

V. Results

A. Image stitching

The image stitching methods were evaluated on the datasets presented in the previous section. For each method, we calculated the mean error with standard deviation for the x and y axes. The results for the *IS1* dataset are summarized in Table II.

TABLE II: Image stitching results on *IS1* dataset

Method	**Glob. opt.**	**Error x-axis** [px]	**Error y-axis** [px]
odometry	-	10.24 ± 7.40	12.35 ± 7.94
NCC	-	**5.36** ± **5.72**	**4.33** ± **4.59**
NCC	least squares	7.23 ± 5.41	6.01 ± 5.09
SIFT	-	9.90 ± 9.08	10.63 ± 10.22
SIFT	least squares	8.72 ± 6.45	10.36 ± 8.63

It can be seen that the odometry in the *IS1* dataset is very imprecise. The best result was achieved with the direct approach with the Normalized Cross-Correlation Coefficient (NCC). Surprisingly, its performance was better without global optimization than with it. The feature-based approach performed worse, but the global optimization slightly improved its accuracy. Both of the methods, however, managed to decrease the odometry error.

The results for the image stitching evaluated on *IS2* dataset with more precise odometry are summarized in Table III. The table shows that in the case of precise odometry, the other methods could not achieve nearly as low an error as the odometry itself. However, the feature-based approach outperformed the correlation-based one. In both cases, global optimization helped achieve better error rates.

TABLE III: Image stitching results on *IS2* dataset

Method	**Glob. opt.**	**Error x-axis** [px]	**Error y-axis** [px]
odometry	-	**1.67** ± **4.04**	**1.85** ± **4.51**
NCC	-	9.59 ± 7.32	9.53 ± 7.09
NCC	least squares	7.67 ± 6.28	8.77 ± 5.30
SIFT	-	8.07 ± 5.23	8.97 ± 6.11
SIFT	least squares	**5.40** ± **4.27**	**5.69** ± **4.73**

In Figure 1, we show a stitched image of the honeybee comb from individual image tiles. We used only the odometry as it proved to be the most sufficient on the new setup. Furthermore, we implemented linear blending to make the transitions between the images smoother. All the image tiles underwent normalization and histogram equalization.

B. Cell detection

The cell detection methods were evaluated using the testing part of the dataset, which was introduced in the previous section. We applied class-agnostic Non-maximum Suppression (NMS) with an IoU threshold of 0.3 on the output of object detectors (where applicable) to account for the fact that each cell should be classified as either fully visible or occluded. We consider the detected circles from the CHT to be of class fully visible cells. For each method and class, we calculated the Average Precision (AP) metric averaged over IoU thresholds in interval $[0.5, 0.95]$. We also report precision (P) and recall (R) for an IoU threshold of 0.5 and a confidence threshold of 0.5. The results are for the class of fully visible cells summarized in Table IV and for the occluded cells in Table V. Moreover, the performance of the CHT and the Faster R-CNN with ground truth is demonstrated on a sample of test data in Fig. 1.

TABLE IV: Fully visible uncapped cell detection results

Method	**AP** [%]	**AP-50** [%]	**P** [%]	**R** [%]
CHT (Bilateral filter, CLAHE)	9.5	13.0	11.1	95.1
YOLOv5s6	86.6	94.0	85.1	93.9
Faster R-CNN (ResNet-18)	84.4	94.8	88.1	92.7
Faster R-CNN (ResNet-50)	**90.9**	**95.3**	**92.2**	**92.4**

TABLE V: Partially occluded uncapped cell detection results

Method	**AP** [%]	**AP-50** [%]	**P** [%]	**R** [%]
CHT (Bilateral filter, CLAHE)	-	-	-	-
YOLOv5s6	64.2	83.3	87.6	68.8
Faster R-CNN (ResNet-18)	64.5	87.7	83.6	80.0
Faster R-CNN (ResNet-50)	**79.2**	**92.0**	**87.4**	**84.9**

It can be seen that although the CHT detects almost all fully visible uncapped cells, it is not suitable for cell detection in natural living colony, as it produces a large number of false positives. The best results were clearly obtained with Faster R-CNN with ResNet-50 backbone, which outperformed all other methods. The Faster R-CNN with ResNet-18 backbone and YOLOv5s6 achieved similar results. The weakness of both models is the detection of partially occluded cells.

VI. Conclusion

In this work, we proposed a pipeline for autonomous mapping of the honeybee comb and tested several of its components. We compared multiple image registration techniques, evaluated the precision of system odometry, and assessed the performance of several uncapped cell detectors. The implemented methods for image registration were not as precise as the odometry in the system with improvements but did provide an edge when the original odometry was weak. For cell detection, the Faster R-CNN network achieved

promising performance even with a rather small dataset used for the training, where the classical CHT algorithm failed due to high levels of clutter caused by bees. By integrating the best-performing methods in a comb mapping pipeline, we were able to determine the positions of opened cells across the entire comb and track their evolution over time.

In future work, we will focus on the improvement of the map composition from detections and robustness to changes in the comb structures. We will extend the map by including semantic information about the detected cell contents, which will further improve the robustness of the map composition. We are confident that we are just a few steps away from having real-time monitoring of a living colony by an autonomous robot.

ACKNOWLEDGMENT

This research was conducted under the European Union's Horizon 2020 research and innovation program grant "RoboRoyale" (agreement no. 964492). JB and JU were also supported by the Grant Agency of the Czech Technical University in Prague, grant no. SGS22/168/OHK3/3T/13. TK was partially supported by EU project no. CZ.02.01.01/00/22_008/0004590. TS, MS and LF were supported by the Field of Excellence COLIBRI (Complexity of Life in Basic Research and Innovation) at the University of Graz.

REFERENCES

[1] K. Žampachů *et al.*, "A vision-based system for social insect tracking," in *2nd International Conference on Robotics, Automation and Artificial Intelligence (RAAI)*, 2022, pp. 277–283.

[2] K. Bozek *et al.*, "Markerless tracking of an entire honey bee colony," *Nature Communications*, vol. 12, no. 1, 2021.

[3] T. Gernat *et al.*, "Automated monitoring of honey bees with barcodes and artificial intelligence reveals two distinct social networks from a single affiliative behavior," *Scientific Reports*, vol. 13, no. 1, 2023.

[4] M. Stefanec *et al.*, "A Minimally Invasive Approach Towards "Ecosystem Hacking" With Honeybees," *Frontiers in Robotics and AI*, vol. 9, p. 791921, 2022.

[5] S. Saalfeld *et al.*, "As-rigid-as-possible mosaicking and serial section registration of large ssTEM datasets," *Bioinformatics*, vol. 26, no. 12, p. i57–i63, 2010.

[6] A. Cardona *et al.*, "Trakem2 software for neural circuit reconstruction," *PLOS ONE*, vol. 7, no. 6, pp. 1–8, 2012.

[7] K. Li and G. Ding, "A novel automatic image stitching algorithm for ceramic microscopic images," in *2018 International Conference on Audio, Language and Image Processing (ICALIP)*, 2018, pp. 17–21.

[8] G. Mahalingam *et al.*, "A scalable and modular automated pipeline for stitching of large electron microscopy datasets," *eLife*, vol. 11, 2022.

[9] P. Šilling, "Deep learning for image stitching," Master's thesis, Brno University of Technology, 2023.

[10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, p. 91–110, 2004.

[11] D. DeTone *et al.*, "Superpoint: Self-supervised interest point detection and description," *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshop*, 2018.

[12] P.-E. Sarlin *et al.*, "Superglue: Learning feature matching with graph neural networks," *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020.

[13] J. Sun *et al.*, "Loftr: Detector-free local feature matching with transformers," *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2021.

[14] D. Steckhan *et al.*, "Efficient large scale image stitching for virtual microscopy," in *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 2008, pp. 4019–4023.

[15] M. Emmenlauer *et al.*, "Xuvtools: Free, fast and reliable stitching of large 3d datasets," *Journal of Microscopy*, vol. 233, no. 1, p. 42–60, 2009.

[16] S. Preibisch *et al.*, "Globally optimal stitching of tiled 3d microscopic image acquisitions," *Bioinformatics*, vol. 25, no. 11, p. 1463–1465, 2009.

[17] T. Tasdizen *et al.*, "Automatic mosaicking and volume assembly for high-throughput serial-section transmission electron microscopy," *Journal of Neuroscience Methods*, 2010.

[18] J. Chalfoun *et al.*, "Mist: Accurate and scalable microscopy image stitching tool with stage modeling and error minimization," *Scientific Reports*, vol. 7, no. 1, 2017.

[19] D. Zukić *et al.*, "Itkmontage: A software module for image stitching," *Integrating Materials and Manufacturing Innovation*, vol. 10, no. 1, 2021.

[20] J. L. Muhlich *et al.*, "Stitching and registering highly multiplexed whole-slide images of tissues and tumors using ASHLAR," *Bioinformatics*, vol. 38, no. 19, pp. 4613–4621, 2022.

[21] S. K. Chow *et al.*, "Automated microscopy system for mosaic acquisition and processing," *Journal of Microscopy*, vol. 222, no. 2, p. 76–84, 2006.

[22] M. A. Palacio *et al.*, "Evaluation of the time of uncapping and removing dead brood from cells by hygienic and non-hygienic honey bees," *Genetics and Molecular Research*, vol. 4, no. 1, pp. 105–114, 2005.

[23] V. Dietemann *et al.*, "Standard methods for varroa research," *Journal of Apicultural Research*, vol. 52, no. 1, pp. 1–54, 2013.

[24] L. Jeker *et al.*, "Computer-assisted digital image analysis and evaluation of brood development in honey bee combs," *Journal of Apicultural Research*, vol. 51, no. 1, pp. 63–73, 2012.

[25] U. Knauer *et al.*, "Application of an adaptive background model for monitoring honeybees," 2005.

[26] P. Rodrigues *et al.*, "Geometric contrast feature for automatic visual counting of honey bee brood capped cells," *In EURBEE 2016: 7th European Conference of Apidology. Cluj-Napoca, Romania*, 2016.

[27] A. C. Sparavigna, "Analysis of a natural honeycomb by means of an image segmentation," *Philica*, vol. 2016, no. 897, 2016.

[28] J. Illingworth and J. Kittler, "The adaptive hough transform," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-9, no. 5, pp. 690–698, 1987.

[29] L. H. Liew *et al.*, "Cell detection for bee comb images using circular hough transformation," in *International Conference on Science and Social Research (CSSR)*, 2010.

[30] B. Höferlin and M. Höferlin, "Hiveanalyzer," https://hiveanalyzer.visionalytics.de.

[31] T. Colin *et al.*, "The development of honey bee colonies assessed using a new semi-automated brood counting method: Combcount," *PLOS ONE*, vol. 13, no. 10, 2018.

[32] T. S. Alves *et al.*, "Automatic detection and classification of honey bee comb cells using deep learning," *Computers and Electronics in Agriculture*, vol. 170, p. 105244, 2020.

[33] G. Paolillo *et al.*, "Automated image analysis to assess hygienic behaviour of honeybees," *PLOS ONE*, 2022.

[34] N. Rathore *et al.*, "Semi-automatic analysis of cells in honeybee comb images," in *2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*, 2023, pp. 1–4.

[35] K. Žampachů, "Visual analysis of beehive queen behaviour," Master's thesis, Czech Technical University in Prague, 2022.

[36] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[37] S. Ren *et al.*, "Faster r-cnn: Towards real-time object detection with region proposal networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017.

[38] TorchVision maintainers and contributors, "Torchvision: Pytorch's computer vision library," https://github.com/pytorch/vision, 2016.

[39] K. He *et al.*, "Deep residual learning for image recognition," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016.

[40] T. Lin *et al.*, "Microsoft COCO: common objects in context," *CoRR*, vol. abs/1405.0312, 2014.

[41] G. Jocher, "YOLOv5 by ultralytics," https://github.com/ultralytics/yolov5, 2020.

[42] A. Kirillov *et al.*, "Segment anything," *arXiv:2304.02643*, 2023.