Bee Counted – an Intelligent System for Counting Honeybees in Images and Videos

Gordon Hunter
School of Computer Science and Mathematics
Kingston University
Kingston upon Thames, U.K.
G.Hunter@kingston.ac.uk

Theo Quéméré-Kerleau
ENSI-CAEN
Caen, France
theo.quemere-kerleau@ecole.ensicaen.fr

John Futcher
School of Computer Science and Mathematics
Kingston University
Kingston upon Thames, U.K.
k1709382@kingston.ac.uk

Krzysztof Siatkowski
School of Computer Science and Mathematics
Kingston University
Kingston upon Thames, U.K.
krzysztof-siatkowski@wp.pl

***Abstract*— Honeybees and other pollinating insects are of vital importance to both the agricultural industry and the wider ecosystem, but they have been in serious decline over recent years. The spread of parasites and predators such as the varroa mite and the Asian hornet create additional threats to the beneficial species. Hence, monitoring the health and well-being of benign insects and the prevalence and location of pests becomes of great importance. In this paper, we describe the use of image processing and machine learning approaches to identify and count honeybees in both still images and videos, with a view to monitoring the activity and health of bees close to a hive. Possible extensions of the work to identify and monitor parasites, predators and other species are also discussed.**

Keywords — honeybees, pollinators, parasites, predators, activity monitoring, image processing, machine learning

I. Introduction

Pollinating insects are responsible for approximately 84 % of the World's food crops, especially fruit [1], and foremost among these are honeybees. However, many of these beneficial insect species have been in serious decline over recent decades [2]. Many possible causes – including pesticides, pollution, climate change and even radiation due to mobile telecommunications – have been proposed, but the effects of parasite and predator species have certainly had major impact. For honeybees, parasites include the varroa mite *Varroa destructor* [3] and tropilaelaps mites *Tropilaelaps mercedesae* [4]. Predators include the Asian hornet *Vespa velutina* [5] which has spread to Europe in recent years and caused the complete annihilation of many bee colonies in countries such as France. Remote monitoring of the health and well-being of bees and other beneficial insects, and of the presence and prevalence of pest species is therefore highly desirable. In this paper, we describe the use of image processing and machine learning approaches to identify and count honeybees in both still images and videos, with a view to monitoring the activity of bees close to a hive. Future potential extensions to the work, to identify and monitor other benign and harmless species, but also pests and predators are also discussed.

The remainder of the paper is structured as follows. The next section gives a brief overview of benign pollinator insects, including honeybees, and some of their major parasites and predators. Related work on remote monitoring of beneficial and detrimental insects and other invertebrates are reviewed in section III. The datasets and methodologies used in our work, and details of their results, are described in section IV. Finally, we discuss our results, present our conclusions and propose how this work can be extended in the future.

II. Pollinators, their Parasites, Pests and Predators

A. Pollinating Insects

Not only are pollinating insects vital to the agricultural industry, but they are also essential to the survival of most species of wildflowers and many types of trees. Although honeybees are the best-known and very prolific pollinators, many other species of bees, plus butterflies, moths, hoverflies and several species of flies and wasps also perform this role for many plants [1]. Some tropical flowering plants are even pollinated by hummingbirds or bats ! Many types of pollinating insects have been in serious decline over much of the world in recent decades [2]. This has been attributed to various factors, including increased use of insecticides, herbicides and other chemicals, monoculture agriculture, climate change, air pollution and increased prevalence of diseases, parasitic and predator species. Some of these factors are believed to be related to each other – for example, climate change may lead to new parasites and predators becoming common in various parts of the world.

Worker honeybees are between 10 and 15 mm in length and between 3 and 5 mm wide and high. Honeybee queens are somewhat longer, normally between 18 and 20 mm in length. Thus, for workers, the width to length ratio is normally between 0.2 and 0.5.

B. Parasistes, Pests and Predators of Honeybees

One of the major causes of honeybee decline are increased incidence and prevalence of parasites such as varroa [3] and tropilaelaps [4] mites, which feed on the body fat or haemolymph (blood) of both adult and juvenile (larvae and pupae) bees and are major vectors of bee diseases. Whilst some birds may eat modest numbers of bees, and mammals such as bears, badgers and rodents may raid bee colonies to steal honey, a major threat to honeybees is posed by other insects, including the small hive beetle (*Aethina tumida*) which eat bee brood, the wax moth (*Galleria mellonella*) and various wasps and hornets [6]. Notable amongst the latter are the Asian yellow-legged hornet (*Vespa velutina*) recently accidentally introduced to and causing major problems in North-Western Europe [5], the Oriental hornet (*Vespa orientalis*) now predating bees in the Mediterranean area and the Middle East, and the Northern giant hornet (*Vespa mandarinia*) recently causing concern in North America. However, many wasps and hornets also serve useful purposes to the environment and in some cases are themselves valuable pollinators. Furthermore, some other beneficial insect species, such as hoverflies, bear superficial resemblance to both bees and wasps, and non-expert eyes can easily get confused between benign and harmful insects. Thus, there is a need for image and video systems which can distinguish between such species, and also assess numbers and activity levels of both bees and their harmless insects and the parasites and predators, and distinguish between harmless and harmful flying insects.

III. Previous Related Work

There have been various different approaches used to monitor bee colony health and well-being. Some of these have used the sounds and/or vibrations produced by the bees [7, 8, 9, 10, 11, 12], whilst others have focused on other signals such as colony internal temperature [13, 14]. Although some authors have attempted using digital image processing and/or video analysis to monitor bee numbers and activity [15, 16] or levels of parasite infestation [17], this modality has been less popular since it tends to require more expensive equipment and large amounts of data storage and/or bandwidth for data transfer. However, with improvements in mobile phone camera over recent years, it is now feasible to use images and/or videos from such devices for serious monitoring and analysis. Nevertheless, Kachole et al [16] noted the difficulty of counting bees which were relatively tightly clustered together. Other approaches for monitoring the activity of bees or other insects include using micro-transmitters or RFID tags attached to individual insects [18], but this becomes highly impractical if it is desired to monitor larger numbers of insects.

IV. Datasets, Methodologies and Results

A. Dataset(s)

The primary dataset used in this project was the Kingston University Biodiversity Dataset, of images and videos captured around the University's campuses, including at the University's apiary of 4 beehives at a semi-rural suburban site in outer South-West London. This dataset can be obtained on request from its owners, via the corresponding author. Each image included various numbers (one or more) of bees, in various orientations, sometimes well-separated from each other, in other cases clumped together in clusters, and taken from various viewing angles. Since the images were acquired in a wide variety of lighting conditions, they were initially converted from RGB format to the HSL (Hue, Saturation, Lightness) colour encoding system [19, 20]. Although bees have distinctive colour patterns on their bodies, in many of the images the background was of similar colours to the bees, so contrast was not necessarily high and image segmentation was not always trivial.

For exemplars, we will illustrate the effects of our various approaches on the two images shown in Figure 1(a) and (b). In the former, a high proportion of the image, taken close-up, is occupied by bees whilst in the latter, taken from further away, only small parts of the image are taken up by bees.

(a) (b)

Figure 1 : Original colour images around a hive before processing, (a) close-up of bees around hive entrance, (b) image of hive entrance taken from further away and at an angle.

The YOLO models (see section B 4 below) were trained using the 80 segmented images from the DSLR Bee Instance Segmentation Dataset (https://universe.roboflow.com/dslr-fly-eggs/bees-b8adg), each containing at least 10 bees.

B. Methods and Results

1. Thresholding – For an initial approach, the images were first converted to greyscale using the standard ITU method [37] then converted to binary using a threshold selected using Otsu's algorithm [21] via its implementation in the scikit-image library in Python. This algorithm minimises the intra-class variance, which is defined as a weighted sum of the variances of the two separate classes. Although this approach yielded satisfactory segmentation of bees from the background in some images (e.g. Figure 2(a)), in others it also highlighted areas of high contrast (e.g. edges of parts of a beehive) where no bees were present (Figure 2(b)).

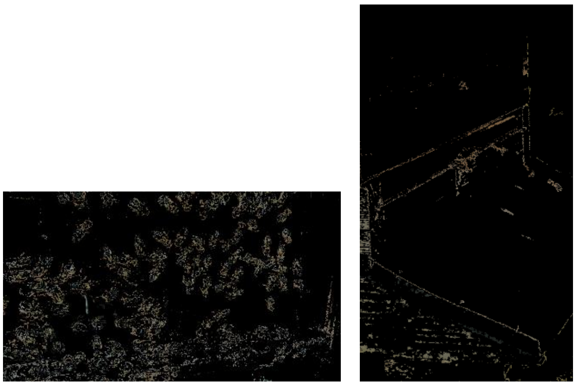
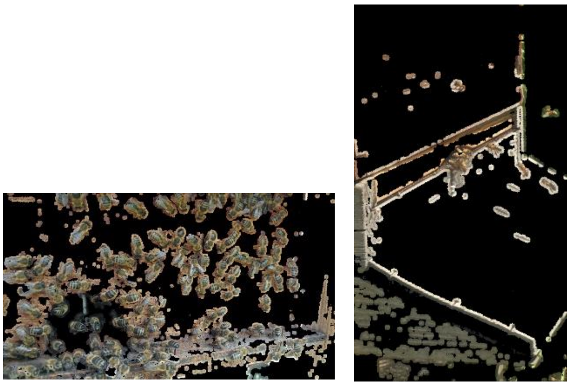


Figure 2 : Images of bees around a hive after thresholding, (a) close-up of bees around hive entrance, (b) image of hive entrance taken from further away and at an angle.

2. Contour Detection and Smoothing Filter – In an attempt to improve the results, we first detected contours in the binary images [22], then filled the interior of each contour with black or white pixels, as appropriate, according to the thresholding method (see Figure a, b for illustrative examples). There has been a clear improvement in the segmentation relative to the simpler approach. However, a few holes remain because the outer contours of the bee shapes are not always closed. Morphological transformations such as erosion and dilation [23] cannot solve this problem because they depend excessively on the distance between the camera and the bees, which can vary greatly from one example to another.

A bilateral smoothing filter was applied to the results of contour detection and filling. Example results are shown in Figure 3 (a), (b). Whilst in some ways the outcomes are substantially improved - in the sense that the bees are more successfully segmented, even for clusters of overlapping bees - there are still various high contrast non-bee objects – such as straight edges of parts of the beehive – highlighted (e.g. in Fig 3(b), which needed removal or being ignored in order for accurate count of bees present to be made.



(a) (b)

Figure 3 : Images of bees around a hive after thresholding, filling and application of bilateral smoothing filter, (a) close-up of bees around hive entrance, (b) image of hive entrance taken from further away and at an angle.

3. Bounding Rectangles or Ellipses – In order to remove the spuriously highlighted areas remaining after filtering, bee sized areas in the filtered thresholded images were identified using both "bounding rectangle" and "bounding ellipse" approaches – the former using a method based on the Hough transform [24] and the latter using the "fitEllipse" (using a least squares process) function in OpenCV [25]. Based on the expected dimensions of honeybees, it was anticipated that the length to width ratio of bounding rectangles or ellipses should be in the range 0.20 to 0.60, depending on orientation, and this was what was observed in our data – see Figure 4. Hence, only rectangles (or, respectively, ellipses) with length to width ratios in this range were retained.

An example of an image segmented this way, with retained bounding rectangles retained, is shown in Figure 5. It can be seen that this approach does not work well for cases where two or more bees are touching or overlapping in the image, since the restriction on the relative dimensions on the bounding rectangles will reject some containing multiple bees as being of the wrong shape, whilst accepting some which are either much too small or much too large. Very similar results were obtained using bounding ellipses instead of rectangles, so these approaches were both of limited value.

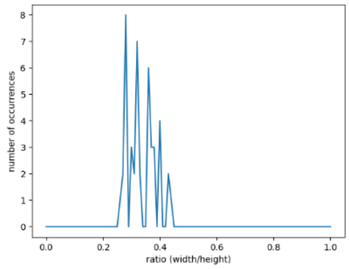


Figure 4 : Occurrence statistics (vertical axis) of width to height ratio (horizontal axis) for bounding rectangles of areas known to be of bees in a sample of our dataset, with values recorded to two decimal places. The "expected" range from apidological sources would be between 0.2 and 0.5, with median value around 0.33.

(a) (b)

Figure 5 : (a) Segmentation of bees with bounding rectangles in an image, (b) the same image, but with a restriction applied on the width to length ratio of the rectangles.

3. Use of Gabor Filtering and Canny Contour Detection

A further approach to improve the segmentation using traditional image processing techniques was carried out using Gabor filters [26] and contour edge detection using Canny's algorithm [27, 28]. The Gabor filters convolve the image array $I(x, y)$, where (x, y) is the position of a given pixel in the image, with a Kernel function $\psi(x, y; \theta, \nu)$, where θ is a rotation angle and ν is a spatial frequency :

$$\psi(x, y;\ \theta, \nu)\ =\ \frac{1}{2\pi\sigma^2}\ exp\left(-\ \frac{\xi^2\ +\ \upsilon^2}{2\sigma^2}\right) exp(i\ 2\pi\ \nu\ \xi)$$

where $\xi = x\cos(\theta) + y\sin(\theta)$ and $\upsilon = -x\sin(\theta) + y\cos(\theta)$ are the pixel's coordinates after the rotation through the angle θ, $i^2 = -1$ and σ is the "width" of the filter, with final output array $G(x, y;\ \theta, \nu) = I(x, y) * \psi(x, y;\ \theta, \nu)$, where $*$ represents convolution. The Gabor filters help to highlight textures, in this case textures characteristic of bees, and assists the subsequent application of Canny's algorithm to detect the contours of salient objects (bees) in the image. This resulted in substantially improved segmentation, but some unwanted objects were still highlighted in addition to the bees – see Figure 6.

Figure 6 : Example segmented image after application of Gabor filters and Canny contour detection.

4. Image Segmentation Based on Deep Learning – YOLO

YOLO ("You Only Look Once") is a Deep Learning-based approach to image segmentation and object detection [29]. Unlike some earlier models (e.g. [30]), YOLO can detect several objects in an image simultaneously. The latest version, YOLOv8, has implementations which are easy to use and, through being pre-trained on large datasets, can achieve very good results after just being "tuned" by re-training on a relatively small specialised dataset appropriate to the task of interest [31].

The YOLO model was pre-trained on the COCO dataset [32], and then re-trained using the 80 segmented images (each containing at least 10 bees) from the DSLR Bee Instance Segmentation Dataset (https://universe.roboflow.com/dslr-fly-eggs/bees-b8adg). Three variants of the YOLO segmentation model were used – the "small" (YOLOv8s-seg), "medium" (YOLOv8m-seg) and "large" (YOLOv8l-seg), each trained on 100 images (80 containing bees and 20 containing no bees) for 30 epochs, and tested on 20 previously unseen images. The progress of training the models are shown in Figure 7 and results of testing are shown as a set of confusion matrices in Table 1 below.

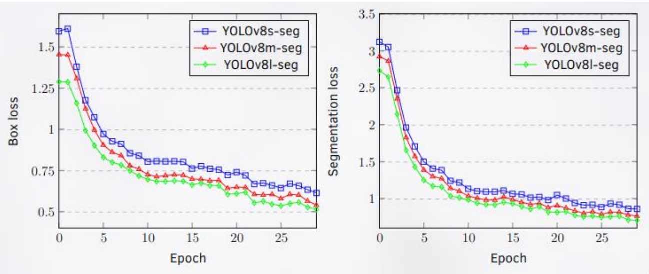


Figure 7 : Decline of "Box Loss" and "Segmentation Loss" whilst training the YOLO models. The Box Loss indicates how well the algorithm locates the centre of an object and how well the predicted bounding box covers that object, whilst the "Segmentation Loss" is based on the ratio of intersection over union of the true and predicted segmentation.

Table 1 : Confusion matrix when testing the YOLO models. PB = "Predicted to be Bee(s)", PN = "Predicted to have No bees", values as fractions between 0.00 and 1.00.

	YOLOv8s-seg		YOLOv8m-seg		YOLOv8l-seg	
	PB	PN	PB	PN	PB	PN
Bee	0.76	0.24	0.87	0.13	0.90	0.10
Non-Bee	0.29	0.71	0.24	0.76	0.19	0.81

Whilst the "large" model gave the best performance in testing, there is relatively little difference in performance between that and the "medium-sized" model, the latter being easier, and taking less time per epoch, to train. Example results from application of the YOLOv8m-seg model are shown in Figure 8. It can be seen that the segmentation is not perfect – with a few bees being missed, and a few areas of the hive surface being mis-identified as being bees, but generally the segmentation of the cluttered image is very good.

(a) (b)

Figure 8 : Example result of image segmentation using YOLOv8m-seg (a) with bounding rectangles included, (b) without bounding rectangles.

5. Video Object Segmentation (VOS) – Xmem and Pixellib

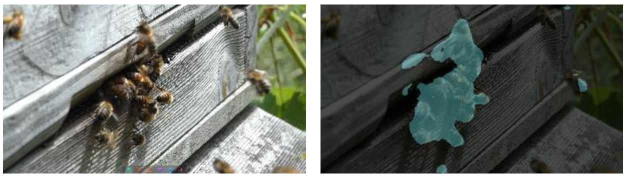
The various attempts at segmentation of the still images yielded results which still included parts of the background, which is static, whilst the bees move. Furthermore, one of the motives for this work is to monitor the level of activity of bees. Thus, it appears logical to investigate segmentation of objects in video sequences as a means to monitor bee activity. Two such Video Object Segmentation (VOS) models were used – Xmem [33] and Pixellib [34] to identify and track bees through video sequences.

Xmem [33] is a semi-supervised approach, based on the model of human visual memory of Atkinson and Shiffrin [35]. This model considers human memory to be made up of three components : sensory memory, short-term (or "working") memory and long-term memory. Visual stimuli are intialised in the sensory memory, and this interacts with the working memory, inserting a new item into the latter every *r* frames. The working memory has limited capacity, but transfers key features to the long-term memory as the former becomes full. The long-term memory slowly "forgets" features as they appear to become obsolete, i.e. no longer occurring in recent inputs. The long-term memory produces a consolidated model, retaining some features found in early observed images, but gradually reducing their influence if no longer found in more recent images, replacing them with features from the newer images. In practical application, a mask for the segmentation of the first image frame in a video is created using a colour map called an "indexed palette". This provides an initialisation for the "working memory component". Example results for the segmentation provided by Xmem for videos including a single bee, a small number (greater than one) of bees and many (around 60) bees are shown in Figure 9. It can be seen that excellent segmentation occurs for image frames containing small numbers of bees, clearly separated, in the image, but the segmentation is much less successful for frames containing many bees, particularly if the bees are densely clustered together. It was hypothesised that this could partly be due to bees moving rapidly from one image frame to another, but experiments involving varying the frame rate only yielded very similar results.

(a) (b) (c)

Figure 9 : Segmentation of video frames using Xmem with (a) a single bee, (b) a small number of bees, (c) many bees

Pixellib [34] is a segmentation model based on the Mask R-CNN hybrid Recurrent-Convolutional Neural Network [36]. We used a pre-trained model, Deeplab [30], trained using a broad variety of videos on different themes, containing many different objects of interest. Similar experiments, using video sequences containing a single bee, a small number (4 to 10) of bees and a large number (at least 40) bees, were carried out to those done using Xmem. The results were again very satisfactory for videos containing just one, two or three bees, but much less so if four or more bees were present. In addition to the same problems with large clusters of bees experienced using Xmem (see Figure 10), with Pixellib it was observed that segmentation could be discontinuous, with bees correctly segmented in some image frames not being identified in an immediate subsequent frame. However, unlike Xmem, Pixellib does not require initialisation using an image mask of the first frame of the sequence, so segmentation of bees which were not present in the first frame was better using Pixellib than with Xmem.



(a) (b)

Figure 10 : Example of poor image frame segmentation using Pixellib when the video contains a large number of clustered bees (a) original image frame, (b) after attempt at segmentation. The cluster of bees are identified as a single "blob", whilst some of the isolated bees present are missed.

V. Discussion, Conclusions and Future Work

The various models investigated here have proved very satisfactory at segmenting and, in the case of video sequences, tracking, bees in images and videos. However, the quality of segmentation tends to decline when larger numbers of bees are present – particularly if the bees are tightly clustered. This problem was previously identified by Kachole et al [2020] and also for the more general problem of segmentation of large numbers of objects by other authors on the GitHub repository forum for Xmem. Segmentation of bees not present in the first frame of any video failed with Xmem, due to the nature of the way the model was initialised, but this was not the case for the Pixellib model. This dependence of the Xmem model on the initial frame is likely to be a major drawback if trying to track bees through a longer video sequence, when bees are likely to enter, leave and possibly re-enter the field of view. This can be addressed by re-initialising the model at regular intervals, but this does present an additional complication, particularly if monitoring is required to be performed in real time.

Both video segmentation models, Xmem and Pixellib, could be re-trained on more appropriate video datasets, if sufficient data were available. This is an aim of ours for the future, for studies involving these and alternative models.

Our studies were based on limited data, using images and videos captured from close to hives, from a limited selection of angles, all in good lighting conditions. Any deviations from these conditions could lead to degradation of performance, so acquisition and use of images and videos from a wider variety of situations – including images of bees living in the wild and of bees with less uniform or predictable backgrounds – would be useful to test the robustness of our methods. Approaches to monitoring the activity levels of the bees – both inside and outside the hive – could also be interesting and useful, although the interior of the hive is normally kept in the dark, so an alternative imaging modality, such as infra-red imaging, might be necessary.

Other planned future work will be to identify species present other than honeybees. As noted in earlier sections, many other insects, both harmless and harmful, bear superficial similarity in appearance to honeybees and it would be of great benefit to beekeepers – and to the bees themselves - if harmful species (including predators such as hornets) could be promptly identified so that remedial action can be taken.

Acknowledgments

This work was in part supported by Innovate UK grant 45568 as part of the *BeeSmart* project, and by BBSRC grant BB/X011992/1 as part of the *ACCENT* project. Theo Quéméré-Kerleau was partly supported by the *ERASMUS* scheme of the EU, and Krzysztof Siatkowski by a Summer Internship awarded by Kingston University.

References

[1] Natural History Museum (UK), "Seven insect heroes of pollination" https://www.nhm.ac.uk/discover/insect-pollination.html

[2] D. Van Engelsdorp, J. Hayes, R.M. Underwood, J.S. Pettis, "A survey of honey bee colony losses in the United States, Fall 2008 to Spring 2009", Journal of Apicultural Research 2010, 49 (1), pp 7–14

[3] S. D. Ramsey et al (2019) "Varroa destructorfeeds primarily on honey bee fat body tissue and not hemolymph", Proceedings of the National Academy of Sciences, 116 (5) pp 1792–1801 https://www.pnas.org/doi/epdf/10.1073/pnas.1818371116

[4] P. Chantawannakul et al (2018) "Tropilaelaps mite: an emerging threat to European honey bee", Current Opinion in Insect Science, 26, pp 69-75, https://doi.org/10.1016/j.cois.2018.01.012

[5] K. Monceau, O. Bonnard, D. Thiéry (2014) "Vespa velutina : a new invasive predator of honeybees in Europe", Journal of Pest Science 87(1) pp. 1-16. doi : 10.1007/s10340-013-0537-3.

[6] Plant Health Australia (no date) "Pests" https://beeaware.org.au/pests/

[7] A.Qandour, I.Ahmad, D. Habibi, M. Leppard (2014) "Remote beehive monitoring using acoustic signals", Acoustics Australia, 42 (3) pp 204 – 209

[8] M. Bencsik et al (2011) "Identification of the honey bee swarming process by analysing the time course of hive vibrations", Computers and Electronics in Agriculture, 76 (1) pp 44 – 50

[9] A. Robles-Guerrero et al (2019) "Analysis of a multiclass classification problem by Lasso Logistic Regression and Singular Value Decomposition to identify sound patterns in queen-less bee colonies", Computers and Electronics in Agriculture, 159, pp 69–74. https://doi.org/10.1016/j.compag.2019.02.024

[10] S. Ruvinga, G. Hunter, J.C. Nebel, O. Duran, (2022) "Prediction of honeybee swarms using audio signals and convolutional neural networks", Proceedings of Workshop on Edge AI for Smart Agriculture (EAISA 2022); 20-23 Jun 2022, Biarritz, France.

[11] S. Ruvinga, G. Hunter, O. Duran, J.C. Nebel (2023) "Identifying queenlessness in honeybee hives from audio signals using machine learning", Electronics, 12 (7), p. 1627, https://doi.org/10.3390/electronics12071627

[12] D. Kanelis et al (2023) "Decoding the Behavior of a Queenless Colony Using Sound Signals", Biology , 12 (11): 1392. https://doi.org/10.3390/biology12111392

[13] G. Hunter, D. Howard, S. Gauvreau, O. Duran, R. Busquets (2019) "Processing of multi-modal environmental signals recorded from a 'smart' beehive", Proceedings of the Institute of Acoustics, 41(1), pp. 339-348

[14] A. Gradišek, J. Bizjak, A. Popovski, J. Grad (2023) "Bumble bee nest thermoregulation: a field study", Journal of Apicultural Research 62 (3), pp 634-642, DOI: 10.1080/00218839.2022.2164651

[15] J. Campbell, L. Mummert, R. Sukthankar (2008) "Video monitoring of honey bee colonies at the hive entrance", Visual Observation and Analysis of Animal and Insect Behavior, 19th ICPR, 8, pp 1-4, https://homepages.inf.ed.ac.uk/rbf/VAIB08PAPERS/vaib9_mummert.pdf

[16] S. Kachole, G. Hunter, O. Duran, “A computer vision approach to monitoring the activity and well-being of honeybees”, Proceedings of 4th International Workshop on Intelligen Systems for Agriculture Production and Environment Protection (ISAPEP’20), , http://dx.doi.org/10.3233/AISE200036

[17] J. Futcher, I. Choudhury, G. Hunter, “Comparison of Machine Learning Techniques to Determine Bee Hive Mite Infestation Level from Mobile Phone Digital Images”, Proceedings of 2nd International Workshop on Edge AI for Smart Agriculture (EAISA 2023), June 2023, Flic en Flac, Mauritius, http://doi.org/10.3233/AISE230023

[18] P. Nunes-Silva et al, “Applications of RFID technology on the study of bees”, Insectes Sociaux, Vol. 66, pp 15-24, 2019

[19] G. Valensi (1939) “Procédé de télévision en couleurs”, French Patent FR841335A, filed 6 February 1939, granted 17 May 1939

[20] G.H. Joblove, D. Greenberg (1978) “Color spaces for computer graphics”, Computer Graphics. 12 (3): 20–25. doi:10.1145/965139.807362

[21] N. Otsu (1979). "A threshold selection method from gray-level histograms". IEEE Transactions on Systems, Man, and Cybernetics. 9 (1): 62–66. doi:10.1109/TSMC.1979.4310076

[22] S. Suzuki, A. Keiichi (1985) “Topological structural analysis of digitized binary images by border following”, Computer Vision, Graphics, and Image Processing, 30 (1), pp 32-46,

[23] R.C. Gonzalez, R. E. Woods (2018) “Digital Image Processing” (4th Edition), Pearson Education, New York.

[24] C.R. Jung, R. Schramm (2004) "Rectangle detection based on a windowed Hough transform," Proceedings. 17th Brazilian Symposium on Computer Graphics and Image Processing, Curitiba, Brazil, 2004, pp. 113-120, doi: 10.1109/SIBGRA.2004.1352951

[25] OpenCV “Creating Bounding rotated boxes and ellipses for contours ” https://docs.opencv.org/4.x/de/d62/tutorial_bounding_rotated_ellipses.html

[26] G.H. Granlund (1978) “In search of a general picture processing operator”. In : Computer Graphics and Image Processing 8.2 (1978), pp 155-173. doi : https://doi.org/10.1016/0146-664X(78)90047-3

[27] Zhao Xu, Xu Baojie, Wu Guoxin (2017) “Canny edge detection based on Open CV”. In : 13th IEEE International Conference on Electronic Measurement Instruments (ICEMI), pp 53-56, doi : 10.1109/ICEMI.2017.8265710.

[28] J. Canny (1968) “A Computational Approach To Edge Detection, IEEE Transactions on Pattern Analysis and Machine Intelligence, 8(6):679–698

[29] J. Redmon, S. Divvala, R. Girshick, A. Farhadi A, (2016) “You Only Look Once: Unified, Real-Time Object Detection”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2016), Las Vegas, NV, USA, 2016, pp. 779-788, doi: 10.1109/CVPR.2016.91.

[30] L-C Chen et al (2017) “Deeplab : Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 40(4), pp 834-848

[31] D. Reis et al (2023) “Real-Time Flying Object Detecttion with YOLOv8”, arXiv : 2305.09972

[32] T-Y. Lin et al (2014) “Microsoft *COCO* – Common Objects in Context”, in Proceedings of Euorpean Conference on Computer Vision (ECCV 2014), pp 740 – 755.

[33] H-K Cheng, A. G. Schwing (2022) “XMem : Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model”, in Proceedings of the European Conference on Computer Vision (ECCV 2022)

[34] A. Olafanwa (2020) “Pixellib : A Library for Pixel-level Image and Video Segmentation”. https://pixellib.readthedocs.io/en/latest/

[35] R.C. Atkinson, R.M. Shiffrin(1968) “Human Memory : A Proposed System and its Control Processes” in *Psychology of Learning and Motivation* (Editors : K. W. Spence et J. Taylor Spence), Academic Press, pp 89-195.

[36] K. He et al (2017) “Mask R-CNN”, arXiv doi : 10.48550/ARXIV.1703.06870. url : https://arxiv.org/abs/1703.06870

[37] International Telecommunications Union (ITU), Recommendation ITU-R BT.470-6*, Conventional Television Systems, 1998 (http://www.itu.int/dms_pubrec/itu-r/rec/bt/R-REC-BT.470-6-199811-S!!PDF-E.pdf)