Intelligent Environments 2020
C.A. Iglesias et al. (Eds.)

A Computer Vision Approach to Monitoring the Activity and Well-Being of Honeybees

Sanket KACHOLE, Gordon HUNTER[1] and Olga DURAN[2]
Faculty of Science, Engineering and Computing, Kingston University, U.K.

Abstract. Honeybees, in their role as pollinators, are vital to both agriculture and the wider ecosystem. However, they have experienced a serious decline across much of the world over recent years. Monitoring their well-being, and taking appropriate action if that is in jeopardy, has thus become a matter of great importance. In this paper, we present an approach based on computer vision to monitor bee activity and motion in the vicinity of an entrance/exit to a hive, including identifying and counting the number of bees approaching or leaving the hive in a given image frame or sequence of image frames.

Keywords. honeybees; motion assessment; direction identification; computer vision; convolutional neural network, LSTN

1. Introduction

Bees are vital contributors both to the agricultural industry and the wider ecosystem, primarily due to their role as pollinators. However, populations of many bee species, including honeybees (*apis mellifera*) and many types of bumblebees (*bombus*), have been in serious decline over the last few decades, particularly in Europe and North America. Various factors, including insecticides, pollution, predators, parasites, diseases, and even microwave EM signals used in telecommunications have been proposed as causes for this decline. Nevertheless, whatever the reasons or it, obtaining a better understanding of the lives of bees in the modern environment, and being able to monitor and preserve the well-being of bee colonies is becoming a higher priority in order to try to stem their severe decline in numbers.

In this paper, we propose and evaluate a method based on computer vision techniques, to monitor honeybee activity close to the entrance/exit to a hive, and attempt to use this to count the number of bees approaching, and the number leaving the hive in a given time interval. This should provide a means of monitoring the level of bee activity in the vicinity of a hive, and to monitor whether the bees are behaving as expected, or whether substantially more bees leave a hive, e.g. due to a swarm occurring, during a day than return later the same day, etc. This should also allow early detection of situations where a hive is in serious decline, e.g. due to disease, parasites or predators, and hence contribute to the promotion of well-being of bee populations. It will complement our other sensor-based approaches to bee monitoring [1, 2]

[1] Corresponding Author, School of Computer Science and Mathematics, Kingston University, KT1 2EE, U.K.; E-mail: G.Hunter@kingston.ac.uk

[2] Corresponding Author, School of Engineering and the Environment, Kingston University, London SW15 3DW, U.K.; E-mail: O.Duran@kingston.ac.uk

2. Related Previous Work

Various previous authors have proposed several quite different approaches to monitor bee colonies remotely using electronic sensor technology of one form or another. Many of these have been reviewed in [5].

Although radio tracking – either using miniature transmitters attached to the body, or a passive RFID tag - has been used to monitor the movement of larger animals or birds for several decades, the smaller size and body mass of insects resulted in this technique not being used on insects until considerably later. One of the first applications of radio tracking to insect movement was Hayashi & Nakane's study of dobsonfly (*protohermes grandis*) larvae [6], using a transmitter weighting just 185 mg. This approach has subsequently been applied to tracking a wide variety of insects [7, 8], including bees [9, 10].

Acoustic monitoring of bee colonies has been carried out since at least the 1960s [11], and has more recently been applied to studying important colony phenomena such as swarming [12] and absence of a queen [13]. Other approaches of in-hive monitoring, including monitoring of brood temperature to predict swarming, have also been quite extensively used (e.g. [14, 15].

However, monitoring the motion of bees in flight is more complicated, and previous attempts to automatically observe and analysis bee activity using computer-vision based approaches have been very limited [16].

3. Data Used

The dataset used was a set of 12 videos captured in Summer 2019 using an AXIS P1346 Network Camera (https://www.axis.com/en-gb/products/axis-p1346) set up on a tripod approximately 1.5 metres from the entrance to a beehive which contained a thriving colony of honeybees. The frame rate used was 30 frames per second at a resolution of 1920 x 1080 pixels. The videos were each approximately 5 seconds (150 frames) long, giving a total of around 1800 image frames.

4. Methodology

4.1. Orientation Identification

Bees (including honeybees) are part of the hymenoptera order of insects and thus have wings attached to their thorax, or central part of the body. The head is smaller and lighter than the abdomen (or tail part) of the bee, and hence the bee's centre of gravity is behind its wings. Therefore, in flight, a bee's head will be raised relative to its abdomen, making a bee's direction of flight easy to determine based on the orientation of its body. The typical orientation of a bee's body during flight is shown in Figure 1. However, before the individual bees could be identified, it was necessary to carry out the following pre-processing steps : (1) Background identification, (2) Background subtraction, (3) Bee bounding box selection, and finally (4) Bee orientation calculation. The first two of these tasks are standard Digital Image/Video processing tasks – see e.g.[3] for standard methods for achieving these. Bespoke tools were written in Matlab to carry out these talks. See Kachole [4] for further details.

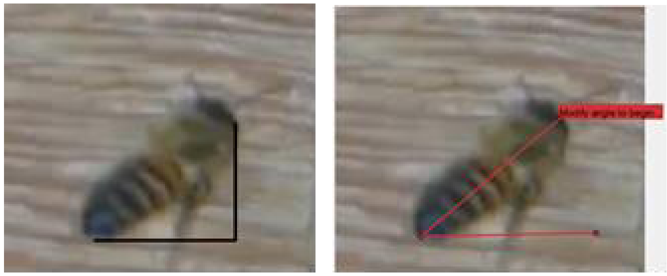

Figure 1 : Close-up image (from our videos) of a bee in flight, close to the wooden hive. The head is clearly above the abdomen, enabling the direction of the bee's motion to be determined from a single frame. The left-hand image shows the *x* (horizontal) and *y* (vertical) components of the bee's oriented dimensions. The right hand image indicates the axis of the bee relative to the horizontal direction. In this case, the bee is oriented in the "upper right" quadrant of the plane, indicating that the bee is moving to the right. Other bees could be oriented in the "upper left" quadrant, indicative of moving to the left,

4.2. Obtaining "Ground Truth" for Training Data

"Ground truth" data was also obtained using our bespoke Matlab labelling tool to manually place rectangular bounding boxes around each bee in a selected number of frames. For each box identified as containing a flying bee, a label was also added to indicate whether the bee was approach or moving away from the hive entrance. An example of a "marked-up" image frame, showing some bees in flight around the hive, is displayed in Figure 2.

4.3. Training a Convolutional Neural Network to Distinguish Bees from Background

Once the Ground Truth had been established, the 2-D U-Net convolutional neural network [17, 18] was trained and then used to preform automatic segmentation of each image. U-Net has various advantages over other types of convolutional neural networks for this task, in that its architecture allows highly accurate segmentations to be performed very quickly despite only requiring a relatively small number of training images. The architecture of a U-Net system has two distinct paths – a contracting path and an expanding path, arranged in a "U" shape which gives the architecture its name. The contracting path is similar to many longer- established convolutional neural networks, consisting of a cycle of two 3x3 unpadded convolutions, each followed by a "Rectified Linear Unit" and a 2x2 MaxPooling layer with stride 2 pixels.

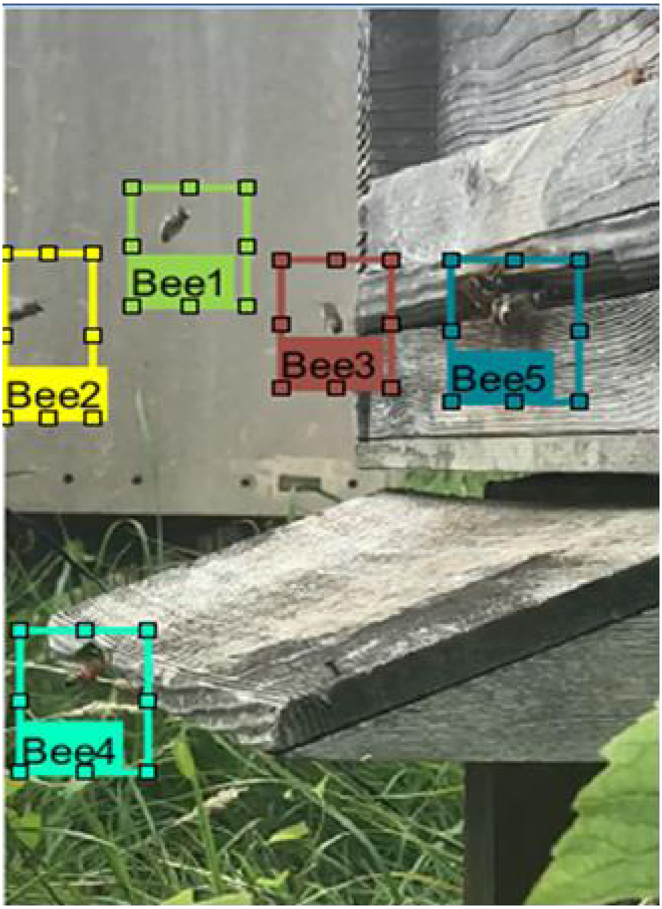

Figure 2 : Ground Truth for one frame – a rectangular bounding box is produced for each bee in the image by a human operator using our bespoke Matlab labelling tool. Note that “Bee1”, “Bee2” and “Bee4” are approaching the hive, whilst “Bee3” is moving away from it, and “Bee5” may actually be a cluster of several bees.

At each stage of this contracting path, the resulting “image” shrinks in size, but the number of feature channels increases. The expansive path performs a sequence of upsampling, 2x2 “up-convolutions”, a concatenation with the feature map at the same level of the contracting path and a pair or 2x2 convolutions, each followed by the application of a Rectified Linear Unit, with one final 1x1 convolution to map the output to any one of the distinct categories being considered. In total, U-Net contains 23 convolutional layers [17]. Its developers state that it is capable of providing precise segmentations despite requiring very few training images, and has recently proved very popular with other researchers. 3D U-Net [18] is a generalization of U-Net which permits the segmentation of volumetric images. A schematic representation of U-Net is shown in Figure 3.

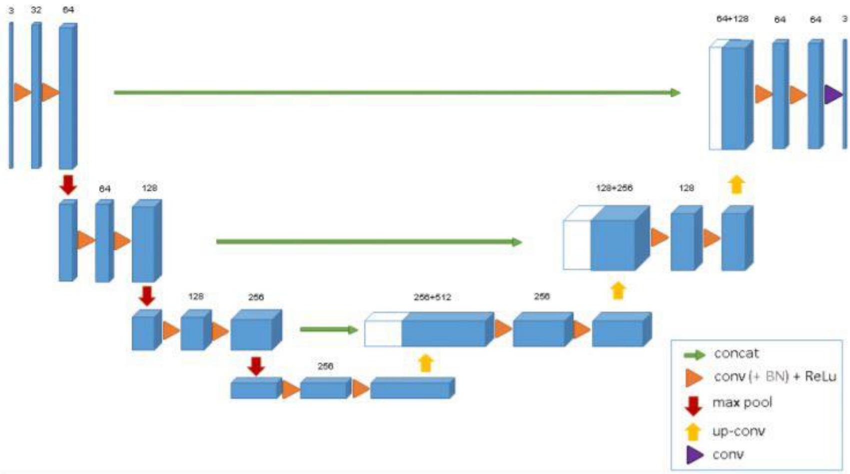

Figure 3 : A schematic representation of the 2D U-Net neural network architecture. “concat” means a concatenation operation, “conv (+BN) + ReLu” means “convolution with Batch Normalisation, followed by Rectified Linear Unit, “max pool” is Max Pooling, “up-conv” means up-convolutions, and “conv” is a conventional convolution. From [18].

In training our network, we monitored both the Root Mean Squared Error (RSME) across the training dataset and the “Dice loss” function [19] over all images in the dataset :

$$\text{Dice} = 2\,|A \cap B| / (|A| + |B|)$$

where A and B are the identified regions in the ground truth and processed images respectively. It takes the value 1 when these regions exactly coincide and 0 if they are disjoint.

During training, it was observed that both the RMSE and the dice loss function both stabilized, at about 20% and 5% of their early peak values respectively, after about 100 iteration cycles through the training data (see Figure 4). Further iterations through the training data only resulted in very modest additional reductions in the values of these metrics. Further details of the implementation can be found in Kachole [4].

In application, the trained U-Net network was used to segment the images into “bees” and “background” (the background could be further sub-divided into “hive” and “vegetation” for the images in our videos, but this sub-categorisation is not used in this study). The binary images, with pixels identified as “bees” displayed in white, and “background” in black, were then subjected to dilation and erosion operations [3] in order to produce a reasonable number of relatively large connected areas of pixels as “bee candidates” and to remove any noise. These been candidate areas could then be counted, have their orientations computed, and compared with the ground truth (where available) for evaluation purposes.

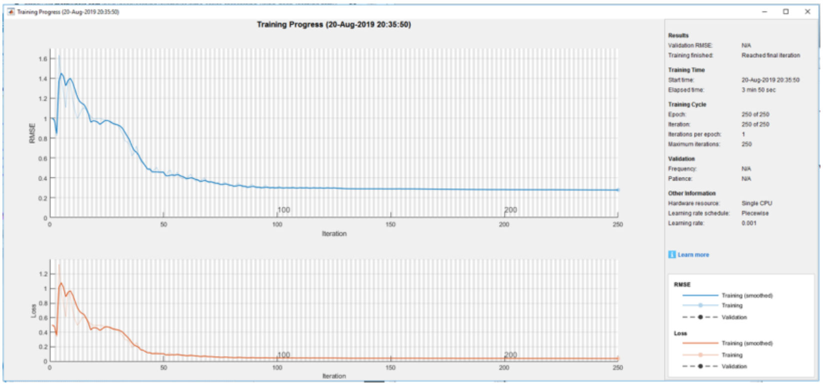

Figure 4 : Variation of Root Mean Square Error (RSME, upper graph) and Dice Loss (lower graph) with respect to number of iterations through the training data during training of the U-Net.

4.4. Monitoring and Evaluation of Bees "Approaching" or "Leaving" the Hive

Once each image frame had been segmented into "bees" and "background", the number of connected "bee" regions in each frame were counted. This is then taken as the computed number of bees in that frame. The orientation calculator (see section 4.1 above) was also applied to each such connected "bee" region to decide whether that bee was approaching or leaving the hive, or was stationary, and the number in each category in each frame was calculated.

In order to predict the number of bees in each category in future frames, a Long Short-Term Memory (LSTM) network was used [20] - a recurrent neural network that can cope with both short and long term dependencies and patterns within time series. LSTMs have proved highly successful in correctly identifying such patterns across many different domains. Further details of the implementation can be found in Kachole [4].

5. Results and Discussion

5.1. Segmentation of Image Frames into Bees and Background

After training, the U-Net network was applied to each image frame in term, to segment that frame into "bees" and "background"(see Figure 5 for an example). The result was compared with the ground truth for each frame, and the average Accuracy calculated :

Accuracy = (Number of pixels correctly classified) / (Total number of pixels in image).

Across the entire dataset, an accuracy of 83.47% was achieved. This compares quite well with other studies on bee classification – Campbell et al [16] achieved a precision of 94% and a recall of 79%, where these quantities are defined by :

Precision = (Number of True Positives) / (True Positives + False Positives)

Recall = (Number of True Positives) / (True Positives + False Negatives)

Thus, for the data of Campbell et al [16], the high precision indicates that there were very few False Positives, but the lower recall value shows that there were a substantial number of False Negatives. In contrast, accuracy compares the total number of True Positives and True Negatives (i.e. all correctly classified pixels) with the total number of all pixels in the image frame.

Figure 5 : (Left) Image of hive entrance, with flying bees, (Right) Binary image showing bees only, after background subtraction. The large "blob" in the binary image corresponds to a cluster of several bees in the original image. The orientation of each bee in flight can be used to determine whether it is approaching or leaving the hive.

5.2. Counting of Bees, including Numbers Approaching or Leaving the Hive

The total number of bees found, and the numbers approaching and leaving the hive, were computed for each frame, and these values noted over a sequence of successive frames. These counts were compared with the corresponding values obtained as predictions for each frame using the LSTM. A selection of these results are shown in Figures 6, 7, 8.

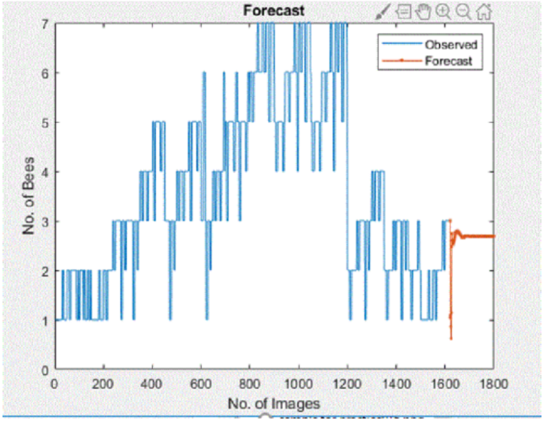

Figure 6 : Total number of bees in each frame against frame number, plus forecast for an additional 200 frames provided by the LSTM.

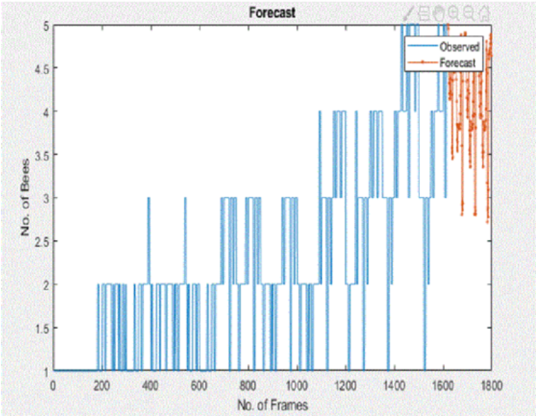

Figure 7 : Number of bees approaching the hive in each frame against frame number, plus forecast for an additional 200 frames provided by the LSTM.

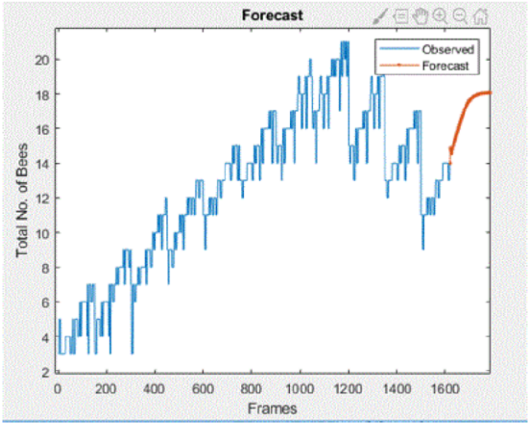

Figure 8 : Number of bees flying away from the hive in each frame against frame number, plus forecast for an additional 200 frames provided by the LSTM.

5.3. Discussion

The accuracy of the segmentation of image frames into bee and non-bee (background) regions seems satisfactory, and the performance is comparable to that obtained in other studies. The LSTM predictions of the number of bees flying in total, or towards or away from the hive have certain limitations – not least that the LSTM output is a continuous-valued variable, whereas in reality these counts will always be discrete (non-negative) integer values. This will therefore lead to errors. Assessment of the utility of the LSTM predictions is currently limited by the relatively small amount of data available. It is planned to extend this evaluation in the near future.

6. Conclusions and Future Work

We have developed and demonstrated a video-based system using a convolutional neural network which can be trained to distinguish between bees and a background consisting of a wooden hive and green plant vegetation. Furthermore, we are able to detect the orientation of bees in flight in order to distinguish bees which are approaching the hive from those leaving it. We have also developed a counting tool to count the total number of bees, the number approaching the hive, and the number leaving it in each image frame, and trained an LSTM recurrent network to make predictions of these quantities a few frames into the future.

The models we have developed so far have been trained on a rather limited dataset. In order to ensure that our system is robust, substantially more video data should be acquired for both training and testing the network.

A useful potential extension to this video-based work would be to use image frames to identify whether the bees were infested with parasites such as varroa mites (varroa destructor) or suffered from defects (e.g. malformed wings) which might be indicative of diseases being prevalent in the colony. Detection and identification of predator species from images would also be highly valuable to beekeepers. Whilst European common wasps (vespula vulgaris) will regularly try to raid beehives to steal honey, and superficially look similar to honeybees (workers of both common wasps and honeybees are around 12 – 17 mm in length) - identifying them might require higher-resolution images - they are not generally considered to be a major threat to bees. Other wasp species, particularly hornets, are a different matter – they are aggressive and may dismember bees in order to feed them to hornet lavae. The European hornet (vespa crabro) is normally significantly larger (workers being about 25 mm long) than honeybees or common wasps, and can threaten the well-being of a weak honeybee colony. The Asian hornet (vespa velutina) is considered a major threat to honeybees, and has been found in several Mediterranean countries and, in a few isolated cases, in Southern and Central England. Although shorter (at about 20 mm for workers) than the European hornet, the Asian hornet is larger than honeybee or common wasp workers and is quite dark – mainly brown – in colour, but has yellow face and legs [21]. Fortunately, the Asian giant hornet (vespa mandarinia) has not yet reached Europe, but is a major predator of bees elsewhere in the World.

The aim of our system is to provide a method for monitoring levels of bee activity to assess their well-being and productivity as pollinators. We are developing a project in collaboration with both academic institutions and NGO development agencies to monitor bees in West Africa to aid environmental protection and promote sustainable development through enhancing cash crop yields via improved levels of pollination of the crop plants' flowers. However, this work is currently just at the planning stage.

Acknowledgements

We would like to that the Biodiversity Team, particularly Sivi Sivanesam, and Colm Treacy (KU beekeeper) at Kingston University for allowing us to access and use the University beehives and for some funding towards the cost of the equipment. We are grateful to Alex Matovu and Dilshan Wettamuni for help recording the videos. We would also like to thanks the School of Education, Kingston University, for allowing us physical indoor space close to the hives, for storing some equipment and providing access to the hives.

References

[1] D. Howard, O. Duran and G. Hunter, “A Low-Cost Multi-Modal Sensor Network for the Monitoring of Honeybee Colonies/Hives” Proceedings of ISAPEP 2018, Rome, Italy, in “*Intelligent Environments 2018*”, pp. 69-78, IOS Press BV., Amsterdam, Netherlands, June 2018.

[2] G. Hunter, D. Howard, S. Gauvreau, O. Duran and R. Busquets, “Processing of Multi-Modal Signals from a ‘Smart’ Beehive”, *Proceedings of the Institute of Acoustics* (UK), Vol. 41(1), pp. 339-348, 2019.

[3] R.C. Gonzalez and R.E. Woods, “*Digital Image Processing*” (4th Edition), Pearson Publishing, 2018.

[4] S. Kachole “Computer vision to assess the wellbeing of Honeybees”, MSc project report, Kingston University, U.K., unpublished, 2019.

[5] R Scheiner et al, “Standard methods for behavioral studies of *Apis mellifera*”, *Journal of Apicultural Research*, Vol 52 (4), 2013

[6] F. Hayashi and M. Nakane, “Radio tracking and activity monitoring of the dobsonfly larva, *Protohermes grandis* (Megaloptera: Corydalidae)”, *Oecologia*, Vol 78, pp 468–472, 1989

[7] M.E. O’Neal et al, “Tracking Insects with Harmonic Radar: a Case Study”, *American Entomologist*, Vol. 50 (4), pp 212-218, 2004.

[8] W.D. Kissling, D.E. Pattemore, M. Hagen, “Challenges and prospects in the telemetry of insects”, *Biological Reviews*, Vol. 89 (3), pp 511- 530, 2014.

[9] S. Streit, F. Bock, C.W.W. Pirk, J. Tautz, “Automatic life-long monitoring of individual insect behaviour now possible”, *Zoology*, Vol. 106 (3) pp 169 – 171, 2003

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

[11] A.M. Wenner, “Sound Communication in Honeybees”, *Scientific American*, Vol. 210, pp 116-124, 1964

[12] S. Ferrari, M. Silva, M. Guarino & D. Berckmans, “Monitoring of swarming sounds in bee hives for early detection of the swarming period”, *Computers and Electronics in Agriculture*, Vol. 64, pp 72-77, 2008

[13] D. Howard, O. Duran, G. Hunter and S. Stebel, “Signal Processing the Acoustics of Honeybees (*Apis Mellifera*) to Identify the ‘Queenless’ State in Hives”, *Proceedings of the Institute of Acoustics*, Vol. 35 (1), pp 290 – 297, 2013

[14] D.S. Kridi et al, “Applications of Wireless Sensor Networks for Beehive Monitoring and In-Hive Thermal Patterns Detection”, *Computers and Electronics in Agriculture*, Vol. 127, pp 221 – 235, 2016

[15] A. Zacepins et al, “Remote Detection of the Swarming of Honey Bee Colonies by Single-Point Temperature Monitoring”, *Biosystems Engineering,* Vol. 148, pp 76 – 80, 2016

[16] J. Campbell, L. Mummert, R. Sukthankar, “Video monitoring of honey bee colonies at the hive entrance”, *Visual Observ. Anal. Animal Insect Behav*., *ICPR*, 8, pp 1-4, 2008

[17] O. Ronneberger, P. Fischer, T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation”, In: Navab N., Hornegger J., Wells W., Frangi A. (eds) “Medical Image Computing and Computer-Assisted Intervention” – MICCAI 2015, pp 234-241. *Lecture Notes in Computer Science*, Vol 9351. Springer, Cham, 2015

[18] Ö. Çiçek et al, “3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation”, In: Ourselin S., Joskowicz L., Sabuncu M., Unal G., Wells W. (eds) “Medical Image Computing and Computer-Assisted Intervention” – MICCAI 2016, pp 424 – 432. *Lecture Notes in Computer Science*, Vol. 9901, Springer, Cham, 2016

[19] L.R. Dice, “Measures of the Amount of Ecologic Association Between Species”, *Ecology*, Vol. 26 (3), pp 297–302, 1945.

[20] S. Hochreiter, J. Schmidhuber, “Long Short-Term Memory”, *Neural Computation*, Vol. 9 (8), pp 1735 – 1780, 1997.

[21] E. Osterloff, G. Broad “Why Asian hornets are bad news for British bees” https://www.nhm.ac.uk/discover/why-asian-hornets-are-bad-news-for-british-bees.html [Last accessed 3 May 2020]