

# Machine Learning and Computer Vision Techniques in Bee Monitoring Applications

Simon Bilik: [bilik@vut.cz](mailto:bilik@vut.cz)<sup>1</sup>, Ondrej Bostik: [bostik@vut.cz](mailto:bostik@vut.cz), Lukas Kratochvila: [kratochvila@vut.cz](mailto:kratochvila@vut.cz), Adam Ligocki: [ligocki@vut.cz](mailto:ligocki@vut.cz), Matej Poncak: [203325@vut.cz](mailto:203325@vut.cz), Tomas Zemcik: [zemcikt@vut.cz](mailto:zemcikt@vut.cz), Milos Richter: [richter@vut.cz](mailto:richter@vut.cz), Ilona Janakova: [janakova@vut.cz](mailto:janakova@vut.cz), Petr Honec: [honecp@vut.cz](mailto:honecp@vut.cz), Karel Horak: [horak@vut.cz](mailto:horak@vut.cz)

**Abstract:** Machine learning and computer vision are dynamically growing fields, which have proven to be able to solve very complex tasks. They could also be used for the monitoring of the honeybee colonies and for the inspection of their health state, which could identify potentially dangerous states before the situation is critical, or to better plan periodic bee colony inspections and therefore save significant costs. In this paper, we present an overview of the state-of-the-art computer vision and machine learning applications used for bee monitoring. We also demonstrate the potential of those methods as an example of an automated bee counter algorithm. The paper is aimed at veterinary and apidology professionals and experts, who might not be familiar with machine learning to introduce to them its possibilities, therefore each family of applications is opened by a brief theoretical introduction and motivation related to its base method. We hope that this paper will inspire other scientists to use the machine learning techniques for other applications in bee monitoring.

**Keywords:** Bee monitoring, Bee inspection, Machine Learning, Deep Learning, Computer Vision, Object detection, Varroasis detection, Bee counting

## 1. Introduction

Honeybee (*Apis mellifera*) is the most important pollinator worldwide and it covers a significant part of the floral visits, some of them exclusively [1]. Nevertheless, bee colonies worldwide face many problems connected with the parasites, diseases, and colony collapse, which might cause problems with the food supply chains and huge economic losses. This problem affects both wild and breeding pollinators with the expected reduction from 3 to 8% of the agricultural production in the case of the total absence of animal pollinators [2].

In recent years, modern machine learning techniques have proven to be very efficient for processing multidimensional data with a huge amount of information and dependencies. These techniques which are often used in computer vision applications made a breakthrough in previously hard-to-solve problems, such as classification, recognition, or inspection tasks, and they could be successfully applied also to the automated bee inspection methods. An exhaustive overview of those methods developed over the last century is presented in [3], but despite the high quality of this paper, the machine learning-based methods are mentioned only briefly. We aim to cover the existing gap with this paper, because a wider use of these techniques could

---

<sup>1</sup> Brno University of Technology

lead to time and economic savings in bee colony research, or to early recognition of the potentially dangerous situations, possibly allowing for a corrective action before dramatic action is required. Most of the described techniques could be easily used for the recognition of the pollen-bearing bees, bee counting, foreign insect detection, or the early diagnosis of various infections.

For the above-mentioned reasons, this paper presents the state-of-the-art methods for bee monitoring based on machine learning, computer vision or their combination in more detail. In the first part of the paper, we briefly describe the basic concepts of machine learning, its prerequisites, possibilities and limits. The second part contains an overview of applied methods from the selected machine learning families together with a short introduction to each technique and the last part describes opportunities of the mentioned application together with a short demonstration of a bee counting algorithm. All methods and techniques are then compared and discussed. We believe that this paper will aid the introduction of these techniques into everyday use and the presentation of their potential to researchers from non-technical fields.

## **2. Bee monitoring and artificial intelligence**

Machine learning is a part of the so-called artificial intelligence, which is a term covering a broad range scale of methods, models, and algorithms, which try to model human behavior in various topics, from computer science (computer vision), classification and recognition tasks, physics (modelling physical processes), pharmacology, and biology (discovering new drugs and molecules). The term artificial intelligence has many definitions, and it is often misused - for those reasons we will focus on the term machine learning.

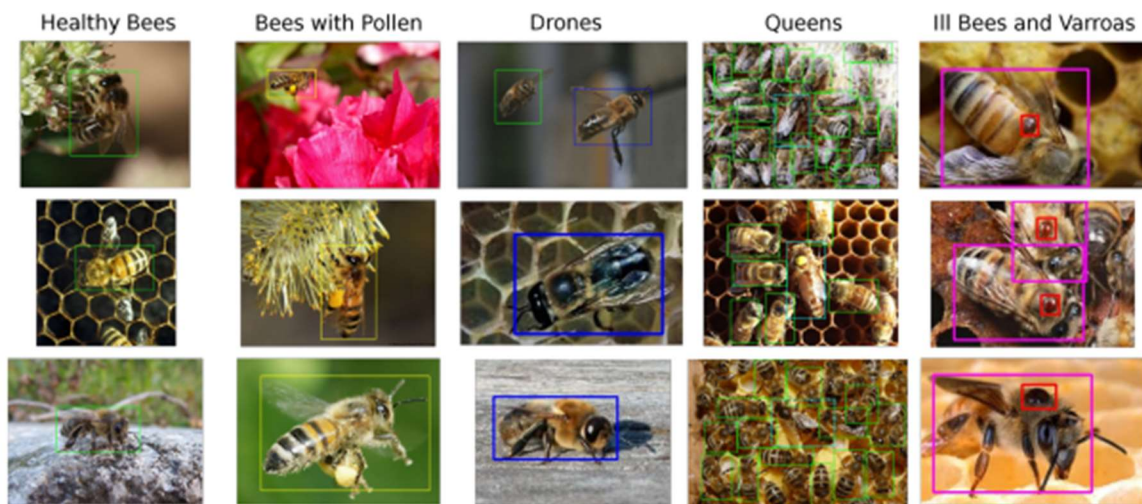
Machine learning itself can be defined as a section of artificial intelligence, which attempts to model human ability to learn independently from a given data and to find patterns or clusters in them. These approaches could be based on statistics, clustering, transformations or on deep learning, but in almost all cases, they are heavily dependent on the given data (so-called datasets described below) and successful learning on one problem does not ensure satisfactory results on another input data. A common division of the machine learning methods is according to the learning process to the supervised and unsupervised - it says whether the inputs have a known (labelled) content at the beginning of the learning process or not. If the input data could be labelled, the supervised approach requires more work in the preparation stage, but in comparison with the unsupervised methods, it could bring more satisfactory results. Nevertheless, input data might not always allow labelling and the unsupervised learning approach must be used in this condition [4].

For the purposes of this article, we will divide those methods into three groups - conventional classifiers, deep learning classifiers and the object detectors, which are all described in more detail below. The main difference between the conventional and deep learning classifiers is in the way, how the features to classify the input data are selected - if they have to be hand-picked by the programmer (for example the observed object size, color, or other various descriptors) in the conventional classifier case, or if they are obtained by the model during the learning process in the deep classifier case.

An advantage of the deep learning approach is that the model itself finds the most suitable features to describe the object, which could be often problematic to get by hand. On the other hand, those features might not be the best representatives from time to time (for example fragments of background or general biases in the data) and a larger amount of data is typically needed for the learning process. Object detectors are usually based on the deep learning approach, and they perform both classification and localization of the searched object [5].

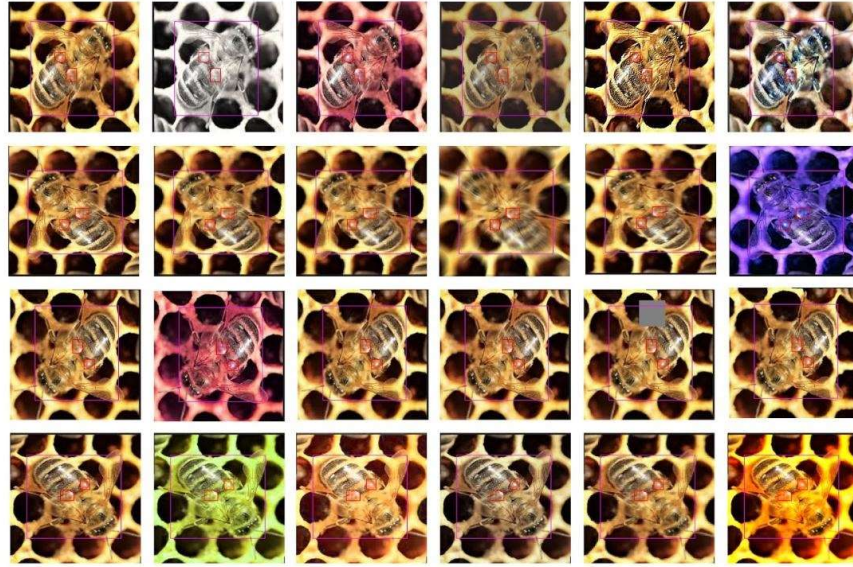
## 2.1. Available datasets

As mentioned above, machine learning techniques require a huge amount of input data called datasets. That data must be annotated by an expert human annotator before the learning and testing stage of the selected method. Preparation of a proper dataset is crucial for the results quality - for conventional machine learning algorithms, it should contain a similar number of all classes examined in the experiment in the required quality. It is often complicated to fulfil this requirement, because some classes could appear only rarely, or it could be difficult to acquire the necessary amount of them.



*Figure 1 Example of a bee dataset as presented in [6]*

For these reasons, data augmentation techniques are often used in order to increase the number of the input data using techniques such as rotation, blurring, scaling or various color filters, which could achieve better learning results and a greater robustness of the selected model. On the other hand, a sufficient amount of information about the observed data has to be already present in the original dataset [7].



*Figure 2 Example of a dataset augmentation as presented in [6]*

There are several datasets containing bee images. Dataset [8] contains high-resolution images of the pollen-bearing bees and is presented in [9]. The authors of [10] made available their dataset for the V.-mite detection at [11], which contains short videos of the healthy and infected bees in high resolution. Dataset [12] contains a high number of bees with extensive amount of additional information about their health-state, pollen carrying, or subspecies, but in a low resolution, and [13] is designed for the distinguishing between the bees, wasps and the other insect. The authors of [6] use a dataset created from the images available online, but it was not made public.



*Figure 3 Image sample from our experimental dataset*

For the purposes of our experiments, we created a dataset containing over 1000 annotated images containing bees incoming and leaving the beehive through narrow passways inspired by [14] and [15], which should mechanically separate the bees for their easier image processing, see Figure 3. To collect this dataset, we used our device described in [16]. We distinguish five classes - a complete bee heading in or out the beehive, the head part, the abdominal part and a cluster of several bees. Frames were captured with the frequency of 5 FPS, which ensures that every passing bee will be captured at least three times (head part - complete bee - abdominal part), and the annotations are in YOLO format. This dataset was used for the image-based bee counting experiments described below and in our future work, we plan to enhance it and make it public. Nevertheless, the dataset is available upon request.

## **2.2. Conventional computer vision and machine learning techniques**

Conventional computer vision (CV) tasks are commonly divided into image filtration, segmentation and recognition. The classical machine learning methods could be categorized as statistical methods, various feature space transformations, or the support vector-based classifiers. Speaking about the conventional approach, those methods do not use deep learning, but they rely on analytic and statistical approach, or classical signal processing. This leads to higher accuracy, inference times and better analytical description of those algorithms, but on the other hand, they are less universal than the CNN-based ones and some very complex problems (written text detection, speech recognition, etc.) might not be solved in a more universal way using them. Nevertheless, better analysis possibilities and readability of those methods remains their great advantage in comparison with the CNNs, which are more difficult and complex to describe. It must be noted that in the field of explainable AI advances are being made [17].

In the conventional approach, the common task like object detection or object classification is usually divided into several individual stages. Every stage is tuned by hand to obtain the data for the next stage. The first stage is commonly the computer vision in a role of feature extraction. The second stage is a suitable machine learning algorithm tuned for this extracted data to achieve the intended task. This approach is best suited for the industry-like environment, where controlled surroundings can be used for the simplification of the task. The main benefit is lower computational complexity.

A perfect illustration of this approach is presented in article [18] and the following [19], where the authors present two experimental setups to capture video of live bees in laboratory conditions. In the first setup, the bees are separated by narrow tunnels with contrasting backgrounds, which simplifies the segmentation and analysis. The device is equipped with artificial lighting to create proper recording conditions. No details are given about the light source used in the text but from the figures it seems that white LED strips producing diffused light are used. The second setup captures the bees in the open space with a camera placed above the measured surface, which speeds up the measurements process, but complicates further processing. The classification stage is designed to differentiate healthy bees without V.-mites and infected bees with V.-mites. Presented work compares various feature sets with three different classification models (Naive Bayes, SVM and Random Forest) resulting in accuracy and F-1 score over 0.8 for some combinations.

Another usage of conventional machine learning methods is presented as a part of [15], where the best light wavelength combination was sought by Linear Discriminant Analysis (LDA) to train and test a classifier looking for the resulting three-dimensional vector of wavelength components that is the most useful for separating bees and mites. The best results for bee mite separation are for light wavelengths above 700 nm. The chosen combination was 470-630-780 nm. In addition, the use of spectra above 700 nm will not impair the comfort of bees, because the bee's eye is not sensitive to higher wavelengths [20].

An interesting study using image processing techniques is presented in [21]. The authors explore the bee grooming process by observing the bees covered with baking flour. For the grooming process analysis conventional computer vision methods were used, such as color corrections and binary segmentation.

Another study is presented in [22], where the authors also recognize the pollen-bearing bees with the classical computer vision methods. They use a compact portable system based on the RaspberryPi 2 embedded computer, which is placed at the beehive's entrance. The study contains its own dataset in a low resolution. The method has achieved a test accuracy around 88%.

Besides the traditional methods for the V.-mite detection, as for example the bee detritus analysis, or the powder sugar-based monitoring, several contemporary approaches were developed. The authors of [23] use an IR camera to detect the temperature increase of the infected brood cells. They prove a slight temperature increase of the infected cells, the cause of which is still unclear. The measurements were performed in laboratory conditions and in the current form, they would be quite impractical for in-field measurements. The article nevertheless shows a certain potential for the thermal measurements approach.

An interesting study is described in [24]. The authors capture a video of an infected bee brood, where they track the V.-mite location, its movement and behavior. The experiment was performed under laboratory conditions and the authors use conventional image processing methods to subtract the background and to localize the V.-mite. This experiment might be further improved on, by using state-of-the-art object detection methods and we believe that it might be useful for future V.-mite behavior analysis.

Several studies also focus on the visual analysis of the brood cells. One of those studies [25] compares several conventional classification techniques in order to recognize the brood cells uncapping and shows the possibility of such in-field measurements, even with the presence of bees. Another study [26] covering the same topic is described in the following section focused on the CNN-based methods.

From the non-vision approaches, in [27] and [28], the authors use a gas sensor array to measure the quality of air in the beehive. The results are used as a feature vector, which is then fed to a k-NN classifier. The authors prove the potential of a machine learning approach for V.-mite detection with 17% detection error and also the potential of the devices for in-field measurements. A great advantage of this approach is that the measurement device doesn't require any significant intrusion to the beehive.

The authors of [29] use acoustic measurements for detection of early V.-mite infestation state based on the beehive acoustic behavior change after its stress exposure. The authors use frequency analysis in order to extract the significant features. Those features are processed by conventional classifiers, which are able to detect the difference when compared to the reference data. Similarly, as in [27] and [28], the system has a minimum negative impact on the beehive.



The study [14] presents a visual-based system for monitoring the in and out activity of the bees. The authors created a device with an infrared camera and illumination, which was placed in front of the beehive and allowed the bees to enter through narrow passways. A small sample of the bees was tagged by a visual identifier and their activity was recognized using Hough transformation and SVM classifier. Although the impractical and demanding tagging approach, the device design is well made and might be an inspiration for other authors.

## **2.3. CNN based classifiers**

A very specialized type of supervised Machine Learning techniques is the class of Deep Learning, and more specifically, the Deep Convolutional Neural Networks. Convolutional Neural Networks, first made practical in the 1980s in [30], were very promising but were prevented from further proliferation into more complicated tasks by their computationally demanding nature. Due to the increase in the availability of processing power in the 1990s and 2000s deeper and deeper networks could be implemented. Since the introduction of AlexNet in 2012, which decisively won the 2012 ImageNet contest, CNNs have become the de-facto state-of-the-art in many computer vision areas [31].

Convolutional neural networks have an advantage in the great complexity and huge number of parameters. This results in the ability to make all the necessary steps needed to classify the object in one step without the need for almost any preprocessing from computer vision algorithms. Convolution layers are automatically focused on sub-parts of the image, and if key objects are found, the object is classified into a learned class. All parameters, including convolution masks, are set automatically, but learning from the beginning can be very time-consuming. Therefore, networks taught for a general problem are used as a template and the networks are only re-trained for a specific task.

The main disadvantage of the deep learning approach is its high computational requirements for the real-time detection. However, in recent years, a large number of computing units adapted for mobile applications such as NVIDIA Jetson, or various accelerators applicable to RaspberryPi have emerged that are suitable for online classification. The training itself, which could be extremely computationally demanding, could be performed externally using specialized hardware. The description above is obviously very brief and simplified. Neural network techniques are a very complicated and fast-evolving field [5], it still stands for the set of mathematical summations and multiplications and is very distant from simulating natural brain processes.

One example of common use of convolutional neural networks in a role of classifier is presented in paper [10]. The authors utilize semantic segmentation approach using DeepLab3 and classification methods based on AlexNet and ResNet in order to separate healthy and infected bees. Because no exact mite images were presented, the authors created a new dataset [11] with 13500 annotated images with healthy and mite infected bees. This paper found good results with ResNet101 architecture pre-trained on ImageNet, or AlexNet architecture trained from scratch, but the best performing models use DeepLabV3 to identify pixels belonging to the parasite and therefore perform semantic segmentation with two variants of ResNet architecture

as encoding models. Per class accuracy of minimum 90% was reached during the experiments and the results prove the potential of the semantic segmentation approach.

Probably one of the most complex studies is a previously mentioned [15]. The authors describe a device suitable for in-field measurements, which is based on the findings presented in [19]. The device uses narrow tunnels to separate the bees together with multispectral illumination and a camera located beneath the tunnels. The entire device is designed to be located at the entrance of the beehive. For the individual bee detection, the authors use conventional computer vision techniques with custom CNN for the V.-mite detection and localization. The authors also suggest a methodology (ILE) for estimation of the infestation level of the whole beehive without a need to measure all bees based on the detected V.-mites and counted bees.

The research presented in [32] suggests a system for visual V.-mite detection and its termination using a swept laser beam controlled by a RaspberryPi 3 embedded computer. In future development, the authors plan to place a camera with the abovementioned laser system above the beehive's entrance and use a CNN classifier for the V.-mite recognition followed by its localization and precise laser beam hit. Although the experiment looks promising, it would be very difficult to achieve the planned results with the proposed equipment and not to harm the bee by an inaccurate hit. Furthermore, much development would be required if a practical mature product is to be developed.

An embedded device is presented in [33], where the authors use a RaspberryPi 3 embedded computer with a camera module to capture bee traffic on a beehive entrance. The camera is placed above the beehive entrance and captures the bee traffic in periodical intervals. From the collected data, the motion regions are detected and processed by a CNN based classifier in order to recognize and count individual bees. In the future, this is to be used for bee traffic estimation. The authors made their datasets publicly available.

The study [26], similarly as [25] focusing on the brood cells, extends the older one with deep learning techniques in order to detect seven possible states of the brood cell from a single image using common CNN classifiers and data augmentation techniques. The authors also provide their dataset and the analysis software.

Another application of CNNs similar to the [22] is shown in [9]. The authors created a portable system for in-field measurements of the beehive's entrance, which captures the entering bees on a defined background. The data is processed by conventional and CNN-based classifiers in order to recognize the pollen-bearing bees with the test accuracy over 95%. The authors also created and published their own dataset.

The research presented in [34] compares several CNN architectures on datasets presented in [10], [12] and [9]. The authors used the originally proposed classes of each dataset and they proved that the common CNN classifiers such as ResNet, AlexNet, or VGG are able to solve those tasks with a comparable accuracy to the original works, which used conventional classifiers or custom CNNs. The paper provides an inference times and performance analysis of all used models.

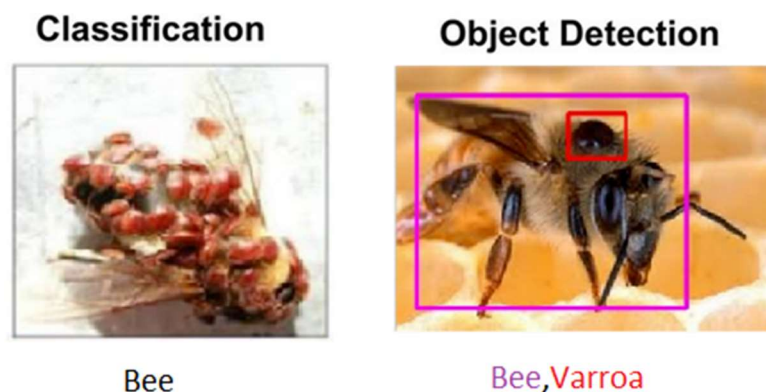


For completeness, we mention two non-English studies, which are therefore not described in great detail. The authors of the first paper [35] focus on distinguishing between the honeybees and other insect species. For this task, they use two publicly available datasets [12], [13] and a custom CNN-based classifier. Using this architecture, they were able to reach a precision of around 94% in the classifications task. The authors of [36] use several custom CNN classifiers to ascertain the health status of the observed bees.

## 2.4. CNN based object detectors

Convolutional neural networks are also great at object detection in real life environments. The convolutional neural network automatically generalizes the features of the searched object and is able to distinguish this described object from the background.

The convolution neural networks were introduced in the previous chapter. For a reminder, convolution neural networks are machine learning algorithms, which can classify an image in one of the categories. These algorithms are called classifiers. On the other side, an object detector can classify and also say, where the category appears in an image. The base of these algorithms are classifiers, but they return a bounding box, where the object exactly is. In advance, object detectors can return more objects in one image. An example of a difference is in Figure 4.



*Figure 4 Difference between classification and object detection*

The object detector model structure is similar to CNN but is enriched with grid and anchor boxes. Anchor boxes are hyperparameters and make it possible to find more objects in one image. A feature image from a convolution layer is divided into a grid and on every grid cell, we try to find anchor boxes. When the confidence is higher than a threshold, then the box is returned as detection.

An object detector approach for V.-mite detection is shown in the paper [6]. Unlike in the previously mentioned studies, the authors use CNN-based object detection techniques SSD and YOLO to classify and localize the V.-mite in a single step. The authors use a custom dataset with bees in various locations to train the networks in order to classify bees and V.-mite. The best results were obtained in the classification of the infected bees using the YOLO object detector. The main advantage of this approach, compared to the previous ones, is the potential of real-time measurement and analysis due to short algorithm inference times. Nevertheless, a practical evaluation of this experiment has not been performed yet.

## **3. Applications areas**

In this chapter, we discuss fields of applications and possibilities of the machine learning systems for automated bee monitoring. Using machine learning, it is possible to connect selected applications to one complex model. To demonstrate the possibilities of the ordinarily used neural network models and the speed of their development, we include a brief description and evaluation of our experimental bee counter.

### **3.1. Fields of application**

Considering the papers described in the previous chapter, we could define the following machine learning applications in the bee health monitoring field. The first application field is bee detection in general, where the experiments usually focus on distinguishing between the bees and other insects. This could be particularly useful for example in the Asian giant hornet and other intruder's detection, or bee anomalous appearance detection, when the results could lead to a fast response of the beekeeper and potential damage minimization.

This application field is very similar to the brood cell inspection field, where machine intelligence and computer vision techniques might be very time-saving and easily used for the inspection and research purposes. Another application field is pollen detection, where we can analyze a total count of pollen-bearing bees, the size of the pollen stockpiles, or with a sufficiently large and well-annotated dataset even a pollen species.

Another application field is the V.-mite detection, where the researchers aim to detect the presence of this parasitic mite usually with the image-based systems. Because this parasite causes great bee colony losses worldwide, development and deployment of this application could also lead to a great economic and bee savings, better-planned treatment, and continual inspection of the bee colony's health state. We also found a suggestion of a system capable of destroying the V.-mite directly on the body of the infested bee. Nevertheless, a problem of the visual-based systems is that most of the V.-mites occur on the bee brood, or that they are hidden between the bee abdominal segments and poorly visible. Therefore, an infestation level estimation methodology like the one presented in [15] has to be developed and extensively tested to allow a better use of this application.

The last application field defined in this paper is the bee counting, which might be very useful in the case of the bee losses early warning and analysis. Bee counting is nowadays solved by tagging the bees with the RFID chips, which is very time demanding, intrusive and it could not cover a large number of the flying bees. On the other hand, image-based systems are much less intrusive and if designed sufficiently, they could detect most of the flying bees. Bee counting systems are elaborately described in [3].

## 3.2. Resnet-50 based bee counter algorithm

The purpose of this experiment was to create a model based on the commonly used networks and classifiers, which could count the bees entering and leaving the beehive and which would demonstrate a rapid development possibility. The whole program was written in the Matlab environment and it is publicly available at [37].

### 3.2.1. Dataset and augmentation

For the purpose of this experiment, we used our experimental dataset described in chapter 2.1 with the structure shown in the Table 1. As the first step, the individual bee images were extracted using the known position of the separation passways. Because the classes *bee\_complete\_in* and *bee\_complete\_out* of the original dataset were strongly disbalanced, we changed the direction of a part *bee\_complete\_in* samples by 180 degrees, so the bees were heading the opposite direction. All images were further augmented by rotation around the X and Y axis with mirroring, which increased their number by four times. The train/test split ratio was set as 8:2.

Dataset classes	Original dataset	Balanced dataset
bee_head	284	284
bee_complete_in	490	287
bee_complete_out	84	287
bee_abdomen	358	358
bee_cluster	175	175

Table 1 Dataset structure

### 3.2.2. Algorithm description and evaluation

At the beginning of our experiment, we tried to directly use pre-trained networks such as Resnet-50, VGG-16, or DenseNet-201. Those models are usually very powerful, but we were not able to get satisfactory results in distinguishing incoming and leaving bees. We assume that this was caused by the great complexity of those networks, which caused too strong generalization of the input samples. Tuning of the network parameters or their fully connected layers did not bring a significant improvement.

Our final algorithm is therefore based on the Matlab example [38], which consists of two parts: the pre-trained Resnet-50 convolutional neural network without the fully connected layers used only for the feature extraction and the SVM classifier used for the actual classification from the extracted features. With this model, we were able to reach an average per-class accuracy of 85% and an average accuracy of 88% in classification of the *bee\_complete\_in* and *bee\_complete\_out* classes. The confusion matrix for the test dataset is shown in Figure 5.

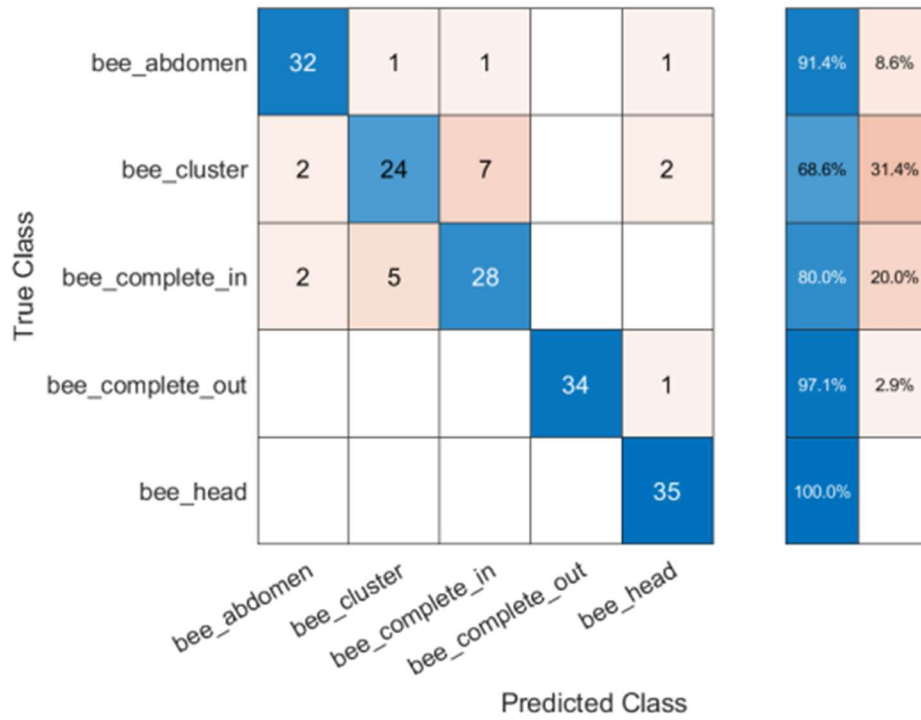


Figure 5 Confusion matrix for the test dataset

To spare computational time, we detect occupied tunnel segments by an average intensity thresholding and then we perform classification on the occupied tunnels only. Detections of the *bee\_complete\_in* and *bee\_complete\_out* classes are used to increment program counters and the other classes are used only to decrease the number of one bee multiple detection, which could otherwise distort the counters values.

### 3.2.3. Bee counter algorithm discussion

In the example of a bee counter algorithm, we show how easily we can design a functional bee monitoring system even with a simple dataset and basic augmentation techniques. With a sufficient dataset, a similar system could be designed for any of the applications described above. The whole algorithm is computationally low demand and it could be run on the embedded devices, such as NVIDIA Jetson, or rPi with the hardware accelerator. This allows a development of monitoring devices suitable for the in-field long term measurements with the online data processing.

## 4. Discussion

As we could see from the studies mentioned above, computer vision and machine learning techniques are often used also in bee monitoring applications. Those applications cover various topics from bee detection and counting, pollen and V.-mite detection, brood cell inspection and some others. An overview of the mentioned studies is shown in the Table 2.

		Application field				
		Bee detection	Pollen detection	V.-mite detection	Brood cell inspection	Bee counting
Technique	Conventional CV	[21]	[22]	[18], [19], [23], [24], [27], [28], [29]	[25]	[14]
	Deep learning	[34], [35]	[9]	[10], [15], [32], [36]	[26]	[33]
	Object detection			[6]		

*Table 2 Overview of the described techniques and applications*

Although the conventional CV and machine learning approach might be very efficient with fast inference times, the modern deep learning methods bring new possibilities and allow solving very complex tasks, which couldn't be successfully solved before. We believe that many studies solved with the conventional approach might be significantly improved using deep learning, or object detection approach, which also allows us to collect much more information from the given data using a single model.

Nevertheless, there remains a significant obstacle to the deep learning approach – the lack of a well-made and annotated dataset in a high resolution. One solution could be joining some datasets mentioned in chapter 2.1, which could be difficult due to the different environments, or to create a complex dataset made by a long-term in-field measurement containing sufficient amounts of the bees in various states (pollen bearing, V.-mite infestation...), drones, or even the bee queens. It could be interesting to cover also sensory data such as sound, in-hive temperature and humidity, beehive weight or others. This would allow development of the complex deep learning-based systems, which could be used for the long-term bee colony monitoring and to recognize potentially dangerous situations.

A great opportunity for those systems is a huge offer of the embedded devices, such as the rPi family. Those devices are widely used, relatively cheap and with plenty of accessories, such as cameras, or sensors. Its computational power is too weak for the efficient training of neural networks, but sufficient for the real-time image processing using selected pre-trained models. Together with the well accessible 3D print technologies, this allows rapid prototyping and brings good customization possibilities for various applications.

## Conclusion

In this paper, we bring an overview of the existing bee monitoring techniques based on computer vision and machine learning, which are divided into three groups - classical computer vision and machine learning techniques without the deep learning, then the CNN based classifiers and the CNN based object detectors. Every group is opened by a brief theoretical introduction and motivation.

We also discuss the problematic of the available bee datasets, which are crucial for the CNN based systems and we emphasize a need for more complex high-quality datasets for developing other applications in this field. Most of the presented papers could be also implemented in the embedded device, which could be used for monitoring, or research purposes. The systems based on the classical computer vision and machine learning techniques might be further enhanced by, the use of the CNN classifiers, or object detectors. In the end, we show how rapidly could be developed bee counter algorithm using a pre-trained model and simple dataset.

## Acknowledgement

The completion of this paper was made possible by the grant No. FEKT-S-20-6205 -"Research in Automation, Cybernetics and Artificial Intelligence within Industry 4.0" financially supported by the Internal science fund of Brno University of Technology.

# References

1. Hung K, Kingston J, Albrecht M, Holway D, Kohn J (2018) The worldwide importance of honey bees as pollinators in natural habitats. *Proceedings of the Royal Society B: Biological Sciences*. <https://doi.org/10.1098/rspb.2017.2140>
2. Aizen M, Garibaldi L, Cunningham S, Klein A (2009) How much does agriculture depend on pollinators? Lessons from long-term trends in crop production. *Annals of Botany* vol. 103::1579-1588. <https://doi.org/10.1093/aob/mcp076>
3. Odemer R (2022) Approaches, challenges and recent advances in automated bee counting devices: A review. *Annals of Applied Biology* vol. 180::73-89. <https://doi.org/10.1111/aab.12727>
4. Russell S, Norvig P (2009) *Artificial intelligence: a modern approach*, 3rd ed. Prentice Hall, Upper Saddle River
5. Horak K, Sablatnig R, Jiang X, Hwang J (2019) Deep learning concepts and datasets for image recognition: overview 2019. *Eleventh International Conference on Digital Image Processing (ICDIP 2019)*. <https://doi.org/10.1117/12.2539806>
6. Bilik S, Kratochvila L, Ligocki A, Bostik O, Zemcik T, Hybl M, Horak K, Zalud L (2021) Visual Diagnosis of the Varroa Destructor Parasitic Mite in Honeybees Using Object Detector Techniques. *Sensors*. <https://doi.org/10.3390/s21082764>
7. DeVries T, Taylor G (2017) Improved Regularization of Convolutional Neural Networks with Cutout
8. Rodriguez I, Mégret R, Acuña E, Agosto J, Giray T () PollenDataset. In: GitHub: <https://github.com/piperod/PollenDataset>
9. Rodriguez I, Megret R, Acuna E, Agosto-Rivera J, Giray T (2018) Recognition of Pollen-Bearing Bees from Video Using Convolutional Neural Network. *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*. <https://doi.org/10.1109/WACV.2018.00041>



10. Schurischuster S, Kampel M (2020) Image-based Classification of Honeybees. 2020 Tenth International Conference on Image Processing Theory, Tools and Applications (IPTA). <https://doi.org/10.1109/IPTA50016.2020.9286673>
11. Schurischuster S, Kampel M () VarroaDataset. In: Zenodo: <https://zenodo.org/record/4085044>
12. Yang J (2018) The BeeImage Dataset: Annotated Honey Bee Images: *Apis mellifera* with location, date, health, and more labels. In: kaggle: <https://www.kaggle.com/datasets/jenny18/honey-bee-annotated-images>. kaggle
13. Rey G (2020) Bee or wasp?: 19480 Hand curated photos of bees, wasps and other insects.. In: kaggle: <https://www.kaggle.com/datasets/jerzydziewierz/bee-vs-wasp>. kaggle
14. Chen C, Yang E, Jiang J, Lin T (2012) An imaging system for monitoring the in-and-out activity of honey bees. *Computers and Electronics in Agriculture* vol. 89::100-109. <https://doi.org/10.1016/j.compag.2012.08.006>
15. Bjerre K, Frigaard C, Mikkelsen P, Nielsen T, Misbik M, Kryger P (2019) A computer vision system to monitor the infestation level of *Varroa destructor* in a honeybee colony. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2019.104898>
16. Bilík Š, Zemčík T, Kratochvíla L, Horák K (2021) Bee Visual Inspector
17. Li X, Cao C, Shi Y, Bai W, Gao H, Qiu L, Wang C, Gao Y, Zhang S, Xue X, Chen L A Survey of Data-driven and Knowledge-aware eXplainable AI. *IEEE Transactions on Knowledge and Data Engineering*. <https://doi.org/10.1109/TKDE.2020.2983930>
18. Schurischuster S, Zambanini S, Kampel M, Lamp B (2016) Sensor Study for Monitoring Varroa Mites on Honey Bees (*Apis mellifera*)
19. Schurischuster S, Remeseiro B, Radeva P, Kampel M (2018) A Preliminary Study of Image Analysis for Parasite Detection on Honey Bees. In: *Image Analysis and Recognition*. SpringerLink, Póvoa de Varzim, Portugal, pp. 465-473

20. Hempel de Ibarra N, Vorobyev M, Menzel R (2014) Mechanisms, functions and ecology of colour vision in the honeybee. *Journal of Comparative Physiology A* vol. 200::411-433. <https://doi.org/10.1007/s00359-014-0915-1>
  
21. Giuffrè C, Lubkin S, Tarpy D (2017) Automated assay and differential model of western honey bee (*Apis mellifera*) autogrooming using digital image processing. *Computers and Electronics in Agriculture* vol. 135::338-344. <https://doi.org/10.1016/j.compag.2017.02.003>
  
22. Babic Z, Pilipovic R, Risojevic V, Mirjanic G (2016) POLLEN BEARING HONEY BEE DETECTION IN HIVE ENTRANCE VIDEO RECORDED BY REMOTE EMBEDDED SYSTEM FOR POLLINATION MONITORING. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*. pp. 51-57
  
23. Bauer D, Wegener J, Bienefeld K (2018) Recognition of mite-infested brood by honeybee (*Apis mellifera*) workers may involve thermal sensing. *Journal of Thermal Biology* vol. 74::311-316. <https://doi.org/10.1016/j.jtherbio.2018.04.012>
  
24. Elizondo V, Briceño J, Travieso C, Alonso J (2013) Video Monitoring of a Mite in Honeybee Cells. *Advanced Materials Research* vol. 664::1107-1113. <https://doi.org/10.4028/www.scientific.net/AMR.664.1107>
  
25. Knauer U, Bienefeld K, Zautke F (2007) A Comparison of Classifiers for Prescreening of Honeybee Brood Cells. *Proceedings of the 5th International Conference on Computer Vision Systems (ICVS 2007)*
  
26. Alves T, Pinto M, Ventura P, Neves C, Biron D, Junior A, De Paula Filho P, Rodrigues P (2020) Automatic detection and classification of honey bee comb cells using deep learning. *Computers and Electronics in Agriculture*. <https://doi.org/10.1016/j.compag.2020.105244>
  
27. Szczurek A, Maciejewska M, Bąk B, Wilk J, Wilde J, Siuda M (2019) Detection of Honeybee Disease: Varroasis using a Semiconductor Gas Sensor Array. In: *Proceedings of the 8th International Conference on Sensor Networks*. SCITEPRESS - Science and Technology Publications, pp. 58-66
  
28. Szczurek A, Maciejewska M (2021) Beehive Air Sampling and Sensing Device Operation in Apicultural Applications—Methodological and Technical Aspects. *Sensors*. <https://doi.org/10.3390/s21124019>

29. Burgess M (2015) Acoustics Australia. Acoustics Australia vol. 43::1-23. <https://doi.org/10.1007/s40857-015-0016-5>
30. LeCun Y, Boser B, Denker J, Henderson D, Howard R, Hubbard W, Jackel L (1989) Backpropagation Applied to Handwritten Zip Code Recognition. Neural Computation vol. 1::541-551. <https://doi.org/10.1162/neco.1989.1.4.541>
31. Krizhevsky A, Sutskever I, Hinton G (2017) ImageNet classification with deep convolutional neural networks. Communications of the ACM vol. 60::84-90. <https://doi.org/10.1145/3065386>
32. Chazette L, Becker M, Szczerbicka H (2016) Basic algorithms for bee hive monitoring and laser-based mite control. 2016 IEEE Symposium Series on Computational Intelligence (SSCI). <https://doi.org/10.1109/SSCI.2016.7850001>
33. Kulyukin V, Mukherjee S (2019) On Video Analysis of Omnidirectional Bee Traffic: Counting Bee Motions with Motion Detection and Image Classification. Applied Sciences. <https://doi.org/10.3390/app9183743>
34. Kaplan Berkaya S, Sora Gunal E, Gunal S (2021) Deep learning-based classification models for beehive monitoring. Ecological Informatics. <https://doi.org/10.1016/j.ecoinf.2021.101353>
35. Barros C, Freitas E, Braga A, Bomfim I, Gomes D (2021) Aplicando Redes Neurais Convolucionais em Imagens para Reconhecimento Automatizado de Abelhas Melíferas (*Apis mellifera* L.). Anais do XII Workshop de Computação Aplicada à Gestão do Meio Ambiente e Recursos Naturais (WCAMA 2021). <https://doi.org/10.5753/wcama.2021.15733>
36. Uzen H, Yeroglu C, Hanbay D (2019) Development of CNN architecture for Honey Bees Disease Condition. 2019 International Artificial Intelligence and Data Processing Symposium (IDAP). <https://doi.org/10.1109/IDAP.2019.8875886>
37. Poncak M (2022) bee-detection. In: GitHub: <https://github.com/999matej999/bee-detection>

38. (2022) Image Category Classification Using Deep Learning. In: MathWorks: <https://www.mathworks.com/help/vision/ug/image-category-classification-using-deep-learning.html>. Mathworks