Recognition of Pollen-Carrying Bees Using Convolutional Neural Networks and Digital Image Processing Techniques

Josué Emmanuel Pat-Cetina[1,2] , Mauricio Gabriel Orozco-del-Castillo[1(✉)] ,
Karime Alejandra López-Puerto[1] , Carlos Bermejo-Sabbagh[1] ,
and Nora Leticia Cuevas-Cuevas[1]

[1] Departamento de Sistemas y Computación, Tecnológico Nacional de México/IT de
Mérida, Mérida, Yucatán, Mexico
{mg16081527,mauricio.od,karime.lp,carlos.bs,nora.cc}@merida.tecnm.mx,
mauricio.orozco@itmerida.edu.mx
[2] AAAI Chapter of Mexico (AAAIMX), Mérida, Mexico

Abstract. The accurate classification of bees into pollen-carrying and
pollen-free categories plays a crucial role in various aspects of bee
research and management, avoiding economic losses in the bee-keeping
sector and a reduction in pollinations of ecosystems. Although beekeepers
can identify when pollen-free bees enter other hives and initiate possible
looting, the task involves a lot of resources. In this paper, we propose
a method for classifying images of bees based on their pollen-carrying
status. We present two approaches: the first method utilizes a convolu-
tional neural network (CNN) to classify original RGB images, while the
second method enhances pollen regions in the images using digital image
processing techniques before training the CNN. The results of the clas-
sification metrics demonstrate the effectiveness of both methods, with
the second method achieving higher accuracy values and reduced loss
compared to the first one. Moreover, the image enhancement techniques
employed in the second method, including thresholding, morphological
operations, and circularity ratio calculation, contribute to improved clas-
sification performance. Additionally, we discuss the CNN architecture,
training parameters, and the significance of incorporating deep learning
techniques in bee image analysis. The proposed method exhibits poten-
tial for application across different bee species, and future work may
explore the extension of this approach to detect looting behavior.

Keywords: Convolutional neural networks · Artificial neural
networks · Digital image processing · Bee management

1 Introduction

The honey bee (*Apis mellifera L.*) is a species native to Africa, western Asia,
and Europe. It was taken to other continents for the production of honey and

pollination of crops, except in Antarctica, and it is the most used bee for those activities [1]. Beekeeping in Mexico is present in various agricultural areas of the country, and has been practiced since pre-Columbian times. Mexico is in third place of production, behind China and Argentina [13]. In Mexico there are approximately 44 thousand producers who generate direct and indirect jobs [13]. In 2013, there was an income of \$164.3 million dollars from the production of honey, generating approximately 2.2 million working days and \$24.9 million dollars for salary payments [13].

The attack on bee colonies by other insects, such as the Asian hornet, or even their own species, can cause economic losses in the beekeeping sector, and a reduction in pollination of ecosystems [11]. It is important to monitor the spread of the species that invade bee colonies in order to plan the actions and activities that lead to stopping their expansion. Some activities carried out to monitor invasive species include direct observations of hornets in hives and flowers, as well as the use of traps such as bottles, funnels, and sticky traps [11]. Although beekeepers can identify when pollen-depleted bees enter other hives and initiate possible looting, the task is expensive, both in terms of time and money [17]. With this in mind, technologies to automate these processes have come into focus in recent years. For instance, recognition of bees has been attempted by the detection of their buzz from audio recordings [8], detection and tracking in three dimensions of bees using real-time stereo vision [4], as well for monitoring the activity of entering and leaving the hive of honey bees [3].

Due to the growing amount of available data, data science and artificial intelligence (AI), particularly machine learning (ML), applications have become much more frequent. Digital image processing (DIP) and computer vision techniques jointly with methods of statistical analysis and inference such as regression, have managed to recognize 98.7% of bees correctly [26]. Some other more powerful ML methods have also been used for bees-related applications, such as artificial neural networks (ANNs). ANNs (described in more detail in Sect. 2.2) are models inspired by the nervous system of living beings, composed as a set of processing units called artificial neurons [19], interconnected through connections called artificial synapses [14]. A specific kind of ANNs particularly useful when dealing with images are convolutional neural networks (CNNs). This model has the capacity to imitate the way in which the visual cortex of the brain processes and recognizes images [10].

In terms of identifying attacks on bees colonies, CNNs have been particularly applied to classify images of bees into two categories: those which carry pollen and those which do not, therefore helping in the identification task automatically. The classification of bees with pollen and without pollen has been carried out using various approaches, for example shallow and deep CNNs were used to detect bees carrying pollen using a manually annotated dataset [21].

This work describes the implementation of a CNN to classify images of bees with pollen and without pollen. Using the dataset from [21], we implement our own CNN model to classify the images and compare its performance with a

second model in which the original dataset was enhanced after applying different DIP techniques, particularly related to mathematical morphology.

This paper is structured as follows. In Sect. 2 we describe the computational methods employed in this work, particularly mathematical morphology in Sect. 2.1 and ML in Sect. 2.2, with a particular emphasis on ANNs and CNNs. In Sect. 3 we present the development process, particularly the data acquisition and the two separate methods, the first one using just the CNN model, and the second one using DIP techniques to emphasize aspects to enhance classification by the model. In Sect. 4 we discuss our results and present our final remarks. Finally, in Sect. 5 future lines of work are explored.

2   Materials and Methods

2.1   Mathematical Morphology

Mathematical morphology describes a set of tools used to extract components from an image, which are used to represent and describe the shape of an object [7]. The analysis of objects is based on set theory, lattice theory, functions, among others [24]. In mathematical morphology objects in an image are represented by sets, for example, in a binary image the set of all white pixels is a morphological description of the original image. Each set in the binary image is a member of a two-dimensional space Z. In mathematical morphology the so-called *structural elements* are used, which are subsets or small images that are used to find properties of interest in an image [7]. By applying structural elements, features such as edges, fills, holes, corners, portions, and cracks can be extracted [24]. The structural elements can be rectangular arrays, but they can also be disc-shaped, rhombus-shaped, etc. [7].

There are several mathematical morphology operations. One of the most commonly-known is erosion [7], which can be thought of as a contraction [9]. Given two sets A and B in a space Z, the set A represents the objects of the image that will be operated by the erosion with the structural element B. The result is a set made up of the points Z, such that the set B translated by Z, is included in the set A [7]. Dilation, on the other hand, can be described as an expansion [9]. The result of the dilation is a set made up of the displacements Z, such that sets B and A overlap by at least one element [7].

Two other mathematical morphology operations are opening and closing, processes that manipulate erosion and dilation operations to improve an image [16]. Opening is a morphological operation used to soften the contour of an object, break up narrow extensions, and remove thin bulges [7]. In contrast, the closing operation serves to enhance the smoothness of contour sections, connect narrow gaps, bridge elongated and thin separations [7], eliminate small holes, and remove internal bulges [24].

Circularity is one of the approaches to describe image regions, and it is used to measure the compactness of a region. The ratio of circularity is the ratio between the area of a region and the area of a circle that has the same perimeter, where the area of a region is given by the number of pixels in the region and

the perimeter of a region is the length of its boundary. Therefore, the circularity ratio R_c is given by:

$$R_c = \frac{4\pi A}{P^2},$$

where A is the area of the region and P is the length of its perimeter. A value of 1 for R_c corresponds to a circular region, while a value of $\frac{\pi}{4}$, corresponds to a square. The circularity ratio is a measure that does not change with uniform scale or orientation changes, such as errors that can occur when resizing and rotating a digital region [7]. If the perimeter of the surface increases and the area does not change, then the ratio of circularity decreases; since irregularities appear on the surface, if the area decreases and the perimeter does not change, as in the case of an ellipse from a circle, then the circularity ratio also decreases [25]. Other measures of the region descriptors used are the average and median intensity levels, the minimum and maximum intensity values, and the number of pixels with values above and below the average [7].

2.2   Machine Learning

ML is a branch of AI that focuses on constructing models capable of extracting knowledge from data and classify, predict, or make decisions without being explicitly programmed to do so. It commonly involves the analysis of large datasets and the extraction of patterns and insights that can be used to improve future performance. ML techniques can be broadly classified into three categories: supervised, unsupervised, and reinforcement learning. Supervised learning algorithms learn from labeled examples, while unsupervised learning algorithms identify patterns and structures in unlabeled data. Deep learning (DL), a subfield of ML, utilizes ANNs with multiple layers to process and learn complex patterns from large amounts of data. By linking inputs and outputs through interconnected layers of artificial neurons, deep learning models are able to automatically extract hierarchical representations and achieve remarkable performance in tasks such as image recognition, natural language processing, and speech recognition. DL has made significant progress in solving AI problems, working with high-dimensional data, and has been applied in various fields such as science, business, and government, encompassing diverse tasks [12].

Artificial Neural Networks. An ANN is a supervised ML method [15]. It is a system made up of many simple elements called neurons, which are interconnected to process information and respond dynamically to external stimuli [18]. The structure of ANNs was developed based on known models of the nervous system and the human brain [14], aiming to emulate their behavior characterized by learning from experience and acquiring general knowledge from a dataset [6]. A neuron conducts electrical impulses generated by chemical reactions under specific operating conditions and consists of three parts: 1) dendrites, which continuously acquire stimuli from several connected neurons; 2) the cell body, or soma, which processes the information from the dendrites to produce a potential

activation indicating whether the neuron sends an electrical impulse along its axon; and 3) the axon, which terminates with branches called *synaptic terminals*, forming connections that transmit impulses from one neuron's axon to the dendrites of other neurons. Since there is no physical contact between the neurons at the synaptic junction, neurotransmitters released in the junction weigh the transmission from one neuron to another [14].

Convolutional Neural Networks. CNNs were developed in the 1980s but were forgotten due to their impracticality in real-world applications. They have since been revived and gained prominence since 2012. CNNs are a type of deep ANN used for image recognition and computer vision tasks [10]. They mimic the way the visual cortex of the brain processes and recognizes images [10]. CNNs process data composed of multiple arrays, such as images composed of three 2D arrays representing the intensity of the three RGB color channels for each pixel [12]. A CNN is broadly composed of two stages: the first one extracts features from an input image, such as gradients, edges, or spots [27], and the second one classifies the image based on the extracted features [10].

3 Development and Results

In this section we describe the methodology employed from the data acquisition to the development and analysis of two methods to classify the images of bees that carry pollen and those that do not: 1) a CNN model trained with the original images, and 2) a CNN model trained with images enhanced with a set of DIP techniques.

3.1 Data Acquisition

A set of images of bees that correspond to the species *Apis Mellifera* was obtained from [21]; data were originally collected from videos captured at a hive entrance. The set consists of 714 RGB images, of which 369 images correspond to bees with pollen and 345 without pollen, measuring 180×300 pixels each. Along with the dataset, a csv file containing the name of each image and the category to which it corresponds was also downloaded. A set of these images is shown in Fig. 1.

Method 1: CNN with Original Images. In the first method, we developed a CNN to classify images of bees based on whether they are carrying pollen or not. The images used in the study were in the standard RGB format, represented as 8-bit integers. In this format, the color values range from 0 (darkest color) to 255 (lightest color). To prepare the image data for CNN training, it was necessary to label each image with the corresponding category from the dataset. Since the images were not in the same order as the dataset, categorical variables were used for labeling. Additionally, to ensure compatibility with the CNN model,
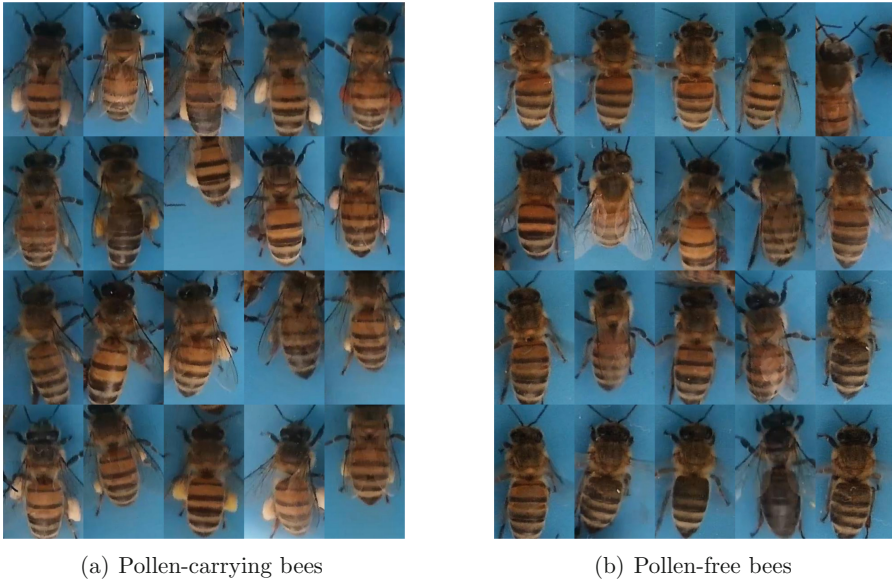
(a) Pollen-carrying bees   (b) Pollen-free bees

Fig. 1. Samples of the dataset retrieved from [21] showing 20 images of (a) pollen-carrying bees, and of (b) pollen-free bees. Notice the amber-colored lumps on the images on the left.

the color values were scaled to a range between 0 and 1 by dividing each value by 255.

The dataset was divided into three subsets: training, validation, and test sets. Random sampling was used to create these subsets, with 64% of the images assigned to the training set, 16% to the validation set, and 20% for testing the trained CNN model. The CNN architecture consisted of five convolutional modules, each using the Rectified Linear Unit (ReLU) activation function. ReLU was chosen as it allows the preservation of positive values, which helps extract image features effectively in subsequent convolutional layers, while setting negative values to zero [2]. This choice was made to avoid the limitations of other activation functions such as tanh or sigmoid, which would restrict the output values to a small range [23].

The configuration details of the convolutional modules used in the CNN are presented in Table 1. These modules play a crucial role in feature extraction and representation learning, enabling the CNN to capture relevant patterns for distinguishing pollen-carrying bees from pollen-free ones.

Within each convolutional module, a max pooling layer with a size of 2×2 is applied after the convolution layers. This pooling layer reduces the spatial dimensions of the filtered images by half, effectively downsampling them. The resulting feature maps are then flattened and connected to a fully connected layer comprising 32 neurons with linear activation function. Subsequently, the

Table 1. Configuration of convolution modules.

Module	Convolution Layers	Kernel	Filters
1	5	5×5	16
2	3	3×3	16
3	3	3×3	32
4	3	3×3	32
5	2	3×3	32

output is passed to a layer consisting of two neurons, responsible for calculating the probabilities corresponding to each category using the softmax function [5]. The softmax function ensures that the outputs represent the probabilities of the image belonging to either category. The training of the CNN model involves 65 epochs, a learning rate of 0.005, and a batch size of 18 images. These parameters were empirically determined to improve performance. The architecture of the CNN is illustrated in Fig. 2, showcasing the arrangement and connections of its different layers.

The accuracy and loss values corresponding to the training stage of the model are shown in Fig. 3a. Initially the accuracy is 0.5065 and the loss is 0.6914. By epoch 13, the accuracy improves to 0.6600, and the loss decreases to 0.6712, with a marginal difference between the two. Notably, by epoch 19, the accuracy significantly rises to 0.9035 while the loss decreases to 0.2713 compared to epoch 13, indicating substantial improvement. Remarkably, in epoch 65, the accuracy reaches its peak close to 1, indicating successful classification of the training data, while the loss drops to 0.0002. Unsurprisingly, validation shows a less stable behavior. In Fig. 3b, the accuracy and loss values during the model validation stage are presented. The accuracy and loss are, respectively, 0.5391 and 0.6914. By epoch 13, the accuracy increases to 0.8347, and the loss decreases to 0.6638, with relatively similar values. Notably, by epoch 65, the accuracy reaches 0.9652, while the loss further reduces to 0.2544 compared to epoch 19, indicating improved accuracy and decreased loss.

Method 2: CNN with Processed Images. In the second method, a DIP technique is employed to enhance the images with the intention to support and improve the CNN classification. This technique involves the application of segmentation techniques, morphological operations, and the calculation of circularity ratio to emphasize relevant features in the images (Fig. 4). To initiate the image enhancement process, a thresholding was proposed for an initial image segmentation. Manually, a set of pixels representing the pollen regions of the images was obtained. Using the average values of this set of RGB color samples the computed average values for the RGB colors are as follows: 178.12 for red, 151.13 for green, and 120.38 for blue. To achieve a wider range of segmentation for the images, a proposal was made to utilize two thresholds for each RGB color channel, derived from the calculation of their respective standard
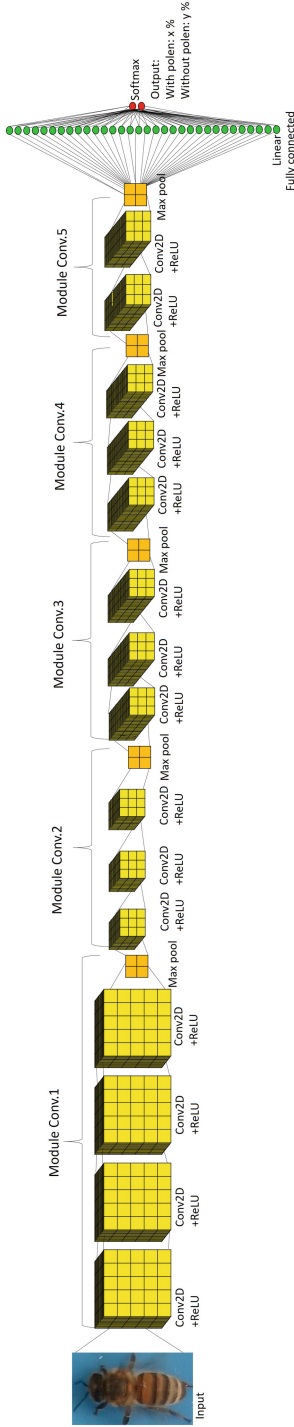
Fig. 2. The architecture of the CNN employed for RGB image classification is depicted. The input image, representing either a pollen-carrying or pollen-free bee, undergoes processing by the CNN for classification. The CNN architecture comprises five convolution modules (represented by yellow grids) followed by max pooling layers (represented by orange grids). These modules extract essential features from the input image. The flattened feature maps are subsequently connected to a fully connected layer (depicted by green neurons). Finally, the fully connected layer is linked to two output neurons employing the softmax function, facilitating the determination of the input image's class. (Color figure online)
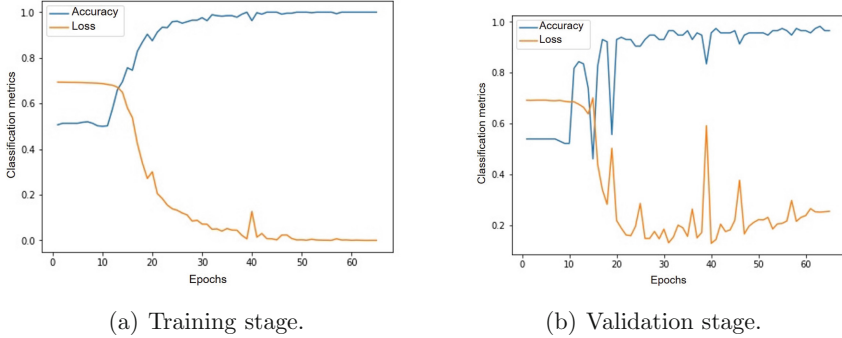
(a) Training stage.

(b) Validation stage.

Fig. 3. Classification metrics in the training and validation stage. Both graphs show the accuracy values in orange, and loss in blue. The x axis shows the epochs and the y axis shows the values of the classification metrics. (a) Accuracy and loss values in the training, (b) accuracy and loss values in the validation of the first method. (Color figure online)

deviations. The calculated standard deviation values are 27.84 for red, 27.45 for green, and 29.86 for blue. For each RGB color channel, the first threshold, denoted as T_1, is computed as $T_1 = \mu_\mathbf{C} - 2\sigma_\mathbf{C}$, *i.e.*, the average minus two times its standard deviation, while the second threshold, denoted as T_2, is computed as $T_2 = \mu_\mathbf{C} + 2\sigma_\mathbf{C}$, *i.e.*, the average plus two times its standard deviation, where μ denotes the mean and σ the standard deviation of a sample of RGB values for a given $\mathbf{C}$ channel. This approach was adopted to enable effective segmentation of multiple pollen regions within the images, as using a single standard deviation value yielded undersegmentation, and using more than two, oversegmentation of the pollen bodies. These values are shown in Table 2.

Table 2. Statistical parameters of the RGB values of the pollen samples and the calculated threshold values.

Color channel	Mean	Standard Deviation	T_1	T_2
Red	178.12	27.84	122.43	233.80
Green	151.13	27.45	96.21	206.04
Blue	120.38	29.86	60.65	180.11

For all three color channels, the respective segmented image $g(x,y)$ is given by:

$$g(x, y) = \begin{cases} 1, & \text{if } T_1 < f(x,y) < T_2 \\ 0, & \text{if } T_1 > f(x,y) \text{ or } f(x,y) > T_2 \end{cases}. \tag{1}$$

The result of this operation makes it possible to segment those areas with values similar to those of the mean color of the pollen. When performing the thresholding, it was observed that there are segmented regions in the image that do not correspond to pollen, particularly the bees' stripes, since their color is similar to that of pollen. The second column of Fig. 4a shows the result of a thresholded image of a pollen-carrying bee, and that of Fig. 4b shows the result of a thresholded image of a pollen-free bee.

To eliminate the areas that do not correspond to the pollen, the morphological opening operation was applied, using a circular structural element of radius 6. The result of this operation allows removing the thin areas of the image, such as the stripes of the bee's body, and soften the outline of the pollen. Most of the stripes on the bee's body were removed, but some still remained. The third column of Fig. 4a shows the result of opening the image of a pollen-carrying bee, and that of Fig. 4b shows the result of opening an image of a pollen-free bee. The pollen area has a nearly circular shape, while the non-pollen regions have a more elongated and not quite circular shape.

To eliminate the areas which do not correspond to pollen, it was proposed to calculate the circularity ratio R_c, and thus establish a value to discard those non-circular areas of the binary images. The value of R_c was determined empirically, trying to find regions that are somewhat circular. The value set for R_c was 0.6, therefore regions of the binary image that have R_c greater than 0.6 are kept in the image, and regions that have R_c less than 0.6 are discarded. The result of applying this operation makes it possible to remove the areas that are not circular, such as the fringes of the bee's body, and to preserve the rounder areas that correspond to the bee's pollen. The fourth column of Fig. 4a shows a pollen-carrying bee, whose areas were not discarded because the R_c value was not less than 0.6, and the Fig. 4b shows how these zones disappeared for the pollen-free bee, since the value of R_c was less than 0.6.

To emphasize the images of the bees, the average of the RGB colors that was calculated from the pollen (Table 2) is used in such a way that the position (x, y) of the binary image that has a value equal to 1, will have the average value of the RGB colors of the pollen in the original image (last column of Fig. 4a), otherwise the color of the original image is preserved (last column of Fig. 4b). The described process (Fig. 4) was carried out with the entire set of images to emphasize the pollen on those bees carrying it.

To prepare the image data for the CNN model, the color values are normalized to be between 0 and 1 by dividing them by 255. The dataset is then divided into three subsets: 64% for training, 16% for validation, and 20% for testing. These subsets are randomly selected to ensure representative samples. The CNN architecture used in this method is the same as the one employed in the first method. It consists of five convolution modules, each utilizing the ReLU activation function. After the convolution layers, a 2×2 max pooling layer is applied to downsample the images. The filtered images are then flattened and connected to a layer of 32 fully connected neurons with the linear activation function. This layer is further connected to a final layer of two neurons using
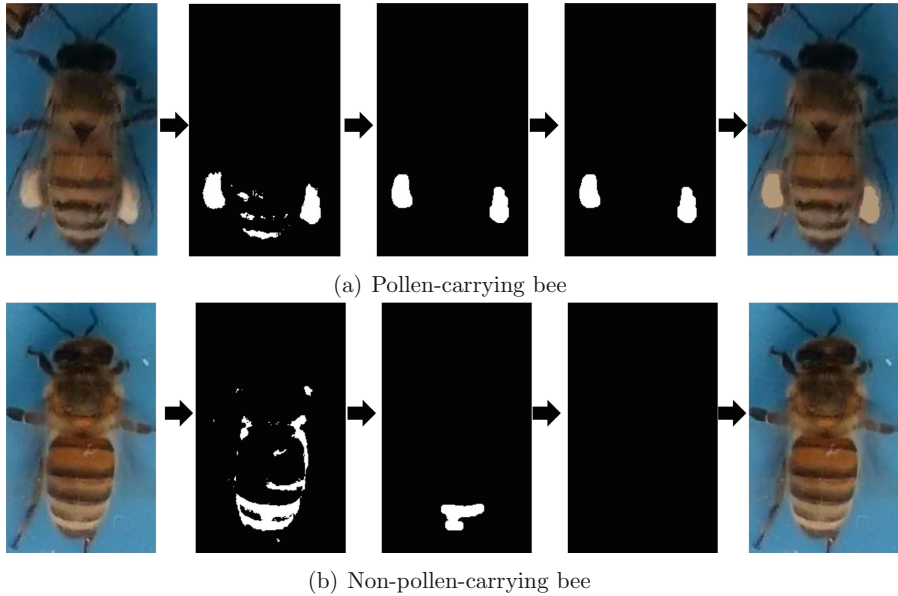
(a) Pollen-carrying bee

(b) Non-pollen-carrying bee

Fig. 4. The proposed process for image enhancement and segmentation of bee images, both for (a) pollen-carrying bees, and (b) pollen-free bees. From left to right, the original image of a bee, the image after thresholding, which separates the foreground (pollen-carrying regions) from the background, the image after applying the opening operation to remove noise and smooth the regions, the image after filtering to remove areas with circularity below 0.6, refining the pollen-carrying regions, and finally the enhanced image obtained by combining the binary areas from the filtering step with the original image, highlighting the pollen-carrying regions. The process effectively enhances the visibility and segmentation of the pollen-carrying areas in bee images.

the softmax function, which calculates the probabilities of an image corresponding to a bee with or without pollen. The CNN is trained for 50 epochs using a learning rate of 0.005 and a batch size of 18 images. The CNN architecture is depicted in Fig. 5. The specific parameters of the CNN were determined through empirical experimentation.

Figure 6a displays the accuracy and loss values throughout the training phase of the model. In the initial epoch, the accuracy value is 0.5153, accompanied by a loss value of 0.6940. By epoch 13, the accuracy increases to 0.7258, and the loss decreases to 0.6216. Notably, in epoch 18, the accuracy significantly improves to 0.9188 compared to epoch 13, while the loss decreases to 0.2397. Eventually, in epoch 50, the model achieves a perfect accuracy of 1, signifying successful classification of the training data.

With respect to the validation stage (Fig. 6b), initially, the accuracy is 0.5652, and the loss is 0.6892. A notable improvement is evident by epoch 13, when accuracy reaches 0.8608 and loss 0.5658. Further progress is observed in epoch 18, with the accuracy climbing to 0.8869 and the loss declining to 0.2051. By
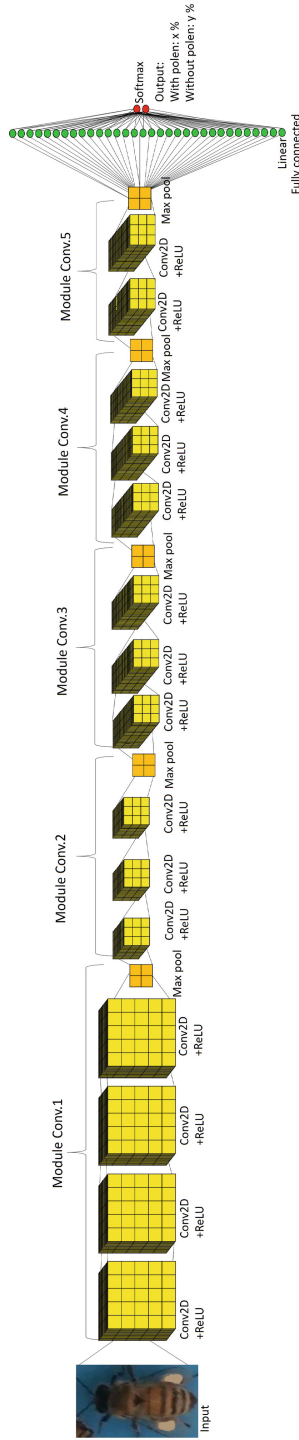
Fig. 5. The CNN architecture employed for classifying processed bee images. The image depicts a bee with pollen, which is fed into the CNN for classification into pollen-carrying or pollen-free categories. The architecture consists of five convolution modules represented by yellow grids, followed by max pooling layers depicted by orange grids. Subsequently, a fully connected layer is illustrated with green neurons, which are further connected to two output neurons utilizing the softmax function. This configuration enables the determination of the image's corresponding class. (Color figure online)

epoch 50, the accuracy achieves a steady state at 0.9478, accompanied by a loss value of 0.2204. Overall, the accuracy and loss values indicate a strong performance.
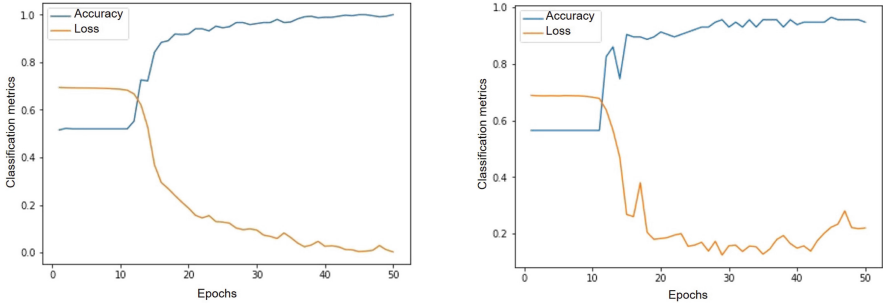
(a) Training stage.   (b) Validation stage.

Fig. 6. Classification metrics in the training and validation stage. Both graphs show the accuracy values in orange, and loss in blue. The x axis shows the epochs and the y axis shows the values of the classification metrics. (a) Accuracy and loss values in the training, (b) accuracy and loss values in the validation of the second method. (Color figure online)

4   Discussion and Conclusions

This work presents a comprehensive process for effectively classifying images of pollen-carrying bees and pollen-free bees. Initially, we utilized a dataset collected from videos captured at a hive entrance [20,21] as the foundation for developing two distinct methods to classify these bee images. A CNN was created and trained using images of pollen-carrying and pollen-free bees. Both methods utilized this trained CNN to compare their classification performance and determine the more effective approach. The classification metrics, including accuracy and loss, are presented in Table 3.

In the first method, we directly used the original RGB images of pollen-carrying and pollen-free bees for classification. The CNN yielded promising results, achieving an accuracy of 0.9230 and a loss value of 0.6861 during the testing stage.

The second method proposed a novel approach to enhance the identification of pollen-carrying bee images. The process involved several image enhancement techniques, including pollen area segmentation using the thresholding technique, application of morphological operations to eliminate irrelevant areas, and utilization of the circularity ratio to refine pollen detection. These emphasized images were then used to train the CNN. In the testing stage, the second method demonstrated improved performance, with an accuracy value of 0.9510 and a reduced loss value of 0.5707. The increased accuracy in the second method can

be attributed to the effectiveness of the DIP techniques in emphasizing pollen regions and subsequently enhancing the CNN's ability to correctly identify and classify them.

Comparing the two methods, we observed that method 1 performed better for the validation set, while method 2 outperformed it in the test set, increasing accuracy from 0.9230 to 0.9510 and decreasing loss from 0.6861 to 0.5707. This suggests that the application of DIP techniques could be highly beneficial in improving the overall accuracy of CNN models.

Our findings demonstrate the capability of CNNs to effectively classify images of pollen-carrying and pollen-free bees. Furthermore, by incorporating DIP techniques to emphasize relevant features, such as pollen zones, the accuracy of the classification process can potentially be significantly enhanced. Overall, this study provides a solid foundation for future research and opens up opportunities to advance the field of bee image analysis.

Table 3. Classification metrics of the test stage of the two methods, particularly accuracy and loss test values.

Method	Accuracy	Loss	Stage
CNN	1	0.002	training
CNN	0.9652	0.2544	validation
CNN	0.9230	0.6861	test
DIP+CNN	1	0.0024	training
DIP+CNN	0.9478	0.2204	validation
DIP+CNN	0.9510	0.5707	test

5   Future Work

Future work in bee image analysis should consider the advantages and disadvantages of using CNNs and DIP techniques. While the proposed method primarily focuses on honey bees, its applicability can be tested with other bee species, such as melipona bees. To further enhance the classification performance, improvements can be made to the image enhancement techniques by adjusting parameters, including the type and size of structural elements used in operations like opening. Fine-tuning the CNN's batch size and training epochs could also lead to improved classification results.

Although this work employed a CNN for image classification, exploring alternative ML methods for feature extraction and classification, such as nearest neighbor or Bayesian classifiers, presents an intriguing avenue for future research [22]. Comparing the performance of these alternative methods with the CNN approach would provide valuable insights into their effectiveness in the context of bee image classification.

While the proposed method does not directly address the detection of looting in bee colonies, the accurate classification of bees into pollen-carrying and pollen-free categories opens up possibilities for automatic counting. This data can be utilized in future research to develop methods for looting detection [22]. Investigating correlations between pollen-carrying behavior and other factors associated with looting could pave the way for identifying and analyzing looting events.

Future work should focus on expanding the applicability of the proposed method to other bee species, refining the image enhancement techniques, exploring alternative classification algorithms, and investigating the potential link between pollen-carrying classification and looting detection. These endeavors will contribute to advancing the field of bee image analysis and provide valuable insights for various research applications. Additionally, addressing concerns regarding the generalizability of results by incorporating a wider range of bee species and populations in the dataset is crucial. Furthermore, accounting for variations in image quality and potential artifacts that may impact classification accuracy should be carefully considered. It is important to recognize the subjectivity inherent in image enhancement techniques and address potential variations in the enhancement process and their impact on classification results.

References

1. Agüero, J.I., Rollin, O., Torretta, J.P., Aizen, M.A., Requier, F., Garibaldi, L.A.: Impactos de la abeja melífera sobre plantas y abejas silvestres en hábitats naturales. Ecosistemas **27**(2), 60–69 (2018). https://doi.org/10.7818/ECOS.1365. https://revistaecosistemas.net/index.php/ecosistemas/article/view/1365

2. Barrera, G.M., Ordiales, H.: Transferencia de estilo entre audios mediante redes neuronales. Revista Digital del Departamento de Ingeniería e Investigaciones Tecnológicas de la Universidad Nacional de la Matanza **4**(1), 12 (2019). https://www.researchgate.net/publication/336221120_TRANSFERENCIA_DE_ESTILO_ENTRE_AUDIOS_MEDIANTE_REDES_NEURONALES_STYLE_TRANSFER_BETWEEN_AUDIOS_USING_NEURAL_NETWORKS

3. Chen, C., Yang, E.C., Jiang, J.A., Lin, T.T.: An imaging system for monitoring the in-and-out activity of honey bees. Comput. Electron. Agric. **89**, 100–109 (2012). https://doi.org/10.1016/j.compag.2012.08.006. https://www.sciencedirect.com/science/article/abs/pii/S0168169912002074

4. Chiron, G., Gomez-Krämer, P., Ménard, M.: Detecting and tracking honeybees in 3D at the beehive entrance using stereo vision. EURASIP J. Image Video Process. (59), 1–17 (2013). https://doi.org/10.1186/1687-5281-2013-59. https://jivp-eurasipjournals.springeropen.com/articles/10.1186/1687-5281-2013-59

5. Cortés, E., Sánchez, S.: Deep learning transfer with AlexNet for chest X-ray COVID-19 recognition. IEEE Lat. Am. Trans. **100** (2020). https://latamt.ieeer9.org/index.php/transactions/article/view/4336

6. Flórez López, R., Fernandez, J.M., Fernández Fernández, J.M.: Las redes neuronales artificiales. Netbiblo, S.L., León, 1 edn. (2008). https://books.google.com.mx/books?id=X0uLwi1Ap4QC&dq=redes+neuronales+artificiales&lr=&hl=es&source=gbs_navlinks_s https://books.google.com.ec/books?id=X0uLwi1Ap4QC&printsec=frontcover&dq=las+redes+neuronales+tienen+una+importante+capacidad+para+detectar+

7. Gonzalez, R.C., Woods, R.E.: Digital Image Processing, 3rd edn. Pearson Prentice Hall, New Jersey (2007). https://dl.acm.org/citation.cfm?id=1076432

8. Heise, D., Miller-Struttmann, N., Galen, C., Schul, J.: Acoustic detection of bees in the field using CASA with focal templates. In: Proceedings of the 2017 IEEE Sensors Applications Symposium, SAS 2017, p. 5. Institute of Electrical and Electronics Engineers Inc., Glassboro (2017). https://doi.org/10.1109/SAS.2017.7894089

9. Khosravy, M., Gupta, N., Marina, N., Sethi, I.K., Asharif, M.R.: Morphological filters: an inspiration from natural geometrical erosion and dilation. In: Patnaik, S., Yang, X.-S., Nakamatsu, K. (eds.) Nature-Inspired Computing and Optimization. MOST, vol. 10, pp. 349–379. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-50920-4_14

10. Kim, P.: MATLAB deep learning. In: MATLAB Deep Learning, chap. 6, 1 edn., pp. 121,147. Apress, Berkeley (2017). https://doi.org/10.1007/978-1-4842-2845-6

11. Laurino, D., Lioy, S., Carisio, L., Manino, A., Porporato, M.: Vespa velutina: an alien driver of honey bee colony losses. Diversity **12**(1) (2020). https://doi.org/10.3390/D12010005

12. Lecun, Y., Bengio, Y., Hinton, G.: Deep learning (2015). https://doi.org/10.1038/nature14539. https://www.nature.com/articles/nature14539

13. Magaña, M.Á.M., Sanginés García, J.R., Lara y Lara, P.E., De Lourdes Salazar Barrientos, L., Morales, C.E.L.: Competitividad y participación de la miel mexicana en el mercado mundial. Revista Mexicana De Ciencias Pecuarias **8**(1), 43–52 (2017). https://doi.org/10.22319/rmcp.v8i1.4304

14. Nunes da Silva, I., Hernane Spatti, D., Andrade Flauzino, R., Liboni, Bartocci, L.H., dos Reis Alves, S.F.: Artificial Neural Networks, vol. 6, , 1 edn. Springer, Sao Paulo (2018). https://doi.org/10.1016/B978-0-444-53632-7.01101-1

15. Padilla Ospina, A.M., Medina Vásquez, J.E., Ospina Holguín, J.H.: Métodos de aprendizaje automático en los estudios prospectivos desde un ejemplo de la financiación de la innovación en Colombia. Revista de Investigación, Desarrollo e Innovación **11**(1), 9–21 (2020). https://doi.org/10.19053/20278306.v11.n1.2020.11676. https://revistas.uptc.edu.co/index.php/investigacion_duitama/article/view/11676

16. Perlis, U.M., Jambek, A.B., Anuar, K., Said, M., Sulaiman, N.: A study of image processing using morphological opening and closing processes. Int. J. Control Theory Appl. **9**(31), 15–21 (2016). https://www.researchgate.net/publication/314154399

17. Uroš, P., Rand'ić, S., Stamenković, Z., Pešović, U., Rand'ić, S., Stamenković, Z.: Design and implementation of hardware platform for monitoring honeybee activity. In: International Conference on Electrical, Electronics and Computing Engineering (October), pp. 5–8 (2017). https://www.etran.rs/common/pages/proceedings/IcETRAN2017/RTI/IcETRAN2017_paper_RTI1_5.pdf

18. Pino Díez, R., Gómez Gómez, A., de Abajo Martínez, N.: Introducción a la Inteligencia Artificial: sistemas expertos, redes neuronales artificiales y computación evolutiva. Universidad de Oviedo, Oviedo, 1 edn. (2001). https://books.google.com.mx/books?id=RKqLMCw3IUkC&dq=redes+neuronales+artificiales&lr=&hl=es&source=gbs_navlinks_s https://books.google.com.co/books?id=RKqLM

Cw3IUkC&lpg=PP1&dq=sistemasexpertosdefinicion&pg=PA10#v=onepage&q=
sistemasexpertosdefinicion&f

19. Rivera, J.S., Gagné, M.O., Tu, S., Barka, N., Nadeau, F., Ouafi, A.E.: Quality
classification model with machine learning for porosity prediction in laser welding
aluminum alloys. J. Laser Appl. **35**(2) (2023)

20. Rodriguez, I.F., Branson, K., Acuña, E., Agosto-Rivera, J.L., Giray, T., Mégret,
R.: Honeybee detection and pose estimation using convolutional neural networks.
Congres Reconnaissance des Formes, Image, Apprentissage et Perception (RFIAP),
pp. 1–3 (2018). https://par.nsf.gov/biblio/10095766

21. Rodriguez, I.F., Megret, R., Acuna, E., Agosto-Rivera, J.L., Giray, T.: Recognition
of pollen-bearing bees from video using convolutional neural network. In: Proceed-
ings of the 2018 IEEE Winter Conference on Applications of Computer Vision,
WACV 2018, pp. 314–322. Institute of Electrical and Electronics Engineers Inc.,
Lake Tahoe (2018). https://doi.org/10.1109/WACV.2018.00041

22. Salah, M.: A survey of modern classification techniques in remote sensing for
improved image classification. J. Geom. **11**(1), 20 (2017). https://isgindia.org/
wp-content/uploads/2017/04/016.pdf

23. Salas, R.: Redes Neuronales Artificiales-Rodrigo Salas. Uni-
versidad de Valparaíso, Valparaíso, Technical report (2005).
https://d1wqtxts1xzle7.cloudfront.net/50358783/Redes_Neuronales_
Artificiales.pdf?1479332205=&response-content-disposition=inline
%3B+filename%3DRedes_Neuronales_Artificiales.pdf&Expires=1619480693&
Signature=csYakiIQj8LjAQrIPQU1Nw44Yt7z~juzjQlzOq1AZdchPy5wH

24. Shih, F.Y.: Image Processing and Mathematical Morphology. CRC Press, Boca
Raton (2017). https://doi.org/10.1201/9781420089448. https://www.taylorfrancis.
com/books/mono/10.1201/9781420089448/image-processing-mathematical-
morphology-frank-shih

25. Takashimizu, Y., Iiyoshi, M.: New parameter of roundness R: circular-
ity corrected by aspect ratio. Progress Earth Planet. Sci. **3**(1), 1–16
(2016). https://doi.org/10.1186/s40645-015-0078-x. https://progearthplanetsci.
springeropen.com/articles/10.1186/s40645-015-0078-x

26. Tu, G.J., Hansen, M.K., Kryger, P., Ahrendt, P.: Automatic behaviour analysis
system for honeybees using computer vision. Comput. Electron. Agric. **122**, 10–18
(2016). https://doi.org/10.1016/j.compag.2016.01.011

27. Venkatesan, R., Li, B.: Convolutional Neural Networks in Visual Computing: A
Concise Guide, 1st edn. Taylor & Francis Group, Phoenix (2018)