İSTANBUL UNIVERSITY
CERRAHPAŞA

ORIGINAL ARTICLE
Electrica 2025; 25: 1-12

Evaluation of Classification Performance of Honey Bee Species With Deep Learning Models

Mücahit Çalışan[1], Nevzat Olgun[2]

[1]Department of Computer Engineering, Bingöl University Faculty of Engineering and Architecture, Bingöl, Türkiye
[2]Department of Software Engineering, Afyon Kocatepe University Faculty of Engineering, Afyon, Türkiye

Cite this article as: M. Çalışan and N. Olgun, "Evaluation of classification performance of honey bee species with deep learning models," *Electrica*, 25, 0157, 2025. doi: 10.5152/electr.2025.24157.

WHAT IS ALREADY KNOWN ON THIS TOPIC?

- *Deep learning, particularly Convolutional Neural Networks (CNN), has proven effective in image-based classification tasks, including biological species identification.*
- *Honey bees play a critical role in maintaining ecological balance and agricultural productivity, and their classification is vital for biodiversity monitoring.*

WHAT THIS STUDY ADDS ON THIS TOPIC?

- *This study introduces and evaluates two deep learning models (Model1 based on classical CNN and Model2 integrating CNN-ResNet) specifically developed for honey bee species classification.*
- *The models achieved high accuracy rates (94.91% for Model1 and 93.63% for Model2), demonstrating that such approaches can be used for fast, accurate, and low-cost classification in biodiversity conservation and smart agriculture applications.*

Corresponding author:
Mücahit Çalışan, mcalisan@bingol.edu.tr

Received: October 22, 2024
Revision requested: December 6, 2024
Last revision received: February 24, 2025
Accepted: March 1, 2025
Publication Date: June 30, 2025

DOI: 10.5152/electr.2025.24157

ABSTRACT

Bees, which have a critical role in the sustainability of our ecosystem, are of vital importance in maintaining ecological balance and biodiversity. More than one-third of the world's food production is dependent on pollinators, and one of the most important of these is the honey bee. As the decline of bee populations threatens ecological balance and biodiversity, the importance of computer vision techniques for the automatic detection of these species is increasing. In this study, two different deep learning models based on classical convolutional neural network (CNN) and Residual Neural Network (ResNet) architectures are proposed and their effectiveness in classifying honey bee species is examined. Model1 is based on the CNN structure and Model2 uses CNN-ResNet integration. In the experiments on the BeeImage dataset containing images of six different honey bee species, it was determined that Model1 achieved 94.91%, and Model2 achieved 93.63% with the K-fold Cross-Validation technique. The results show that deep learning models can classify bee species in a fast and low-cost way. It is seen that these technologies have significant potential in biological research. The high success rates obtained in the study show that the developed deep learning models can contribute to the protection of biodiversity.

***Index Terms*—** Classification, convolutional neural networks., deep learning, honey bee

I. INTRODUCTION

Honey bees, which play a critical role in the continuity of the ecosystem by providing important contributions to agricultural production, are the most important pollinators in nature. Different reasons such as global warming, the effect of Varroa-like parasites, and modern agricultural practices jeopardize the populations of this species. This fact necessitates the development of innovative methods for monitoring, detection, and protection of honey bees.

Accurately identifying honey bee species serves a critical role in both ecosystem health and agricultural productivity. Honey bees are of vital importance for agriculture and ecosystem balance due to their significant role in pollination processes. Depending on the climate and vegetation of each region, there are bee species with different characteristics that can survive. On different continents, honey bees are divided into physiologically, genetically, and behaviorally diverse species. While some species can survive in wide regions, others are distributed in more limited regions [1]. This situation makes it important to select the appropriate bee species for the region in terms of productivity. Choosing the wrong species can damage the ecosystem by increasing bee mortality. The conservation of the bee population is of critical importance both for maintaining the natural habitat and for a sustainable life. Therefore, accurate identification of bee species and selection of appropriate species is of crucial importance. However, the identification and classification of different honey bee species is often a challenging and time-consuming process. Traditional methods often require expertise and longer observations, reducing efficiency for studies requiring large amounts of data and high accuracy.

In recent years, the development of deep learning algorithms has contributed to significant advances in image processing and object recognition. In particular, methods such as convolutional neural networks (CNN) can provide high accuracy rates on large datasets. These models can also be used in the classification of honey bee species. One of the main criteria for determining honey bee species is morphological features such as wing length and color. Observing and measuring these characteristics and comparing them with other honey bee species are processes carried out by experts. These processes have a high potential for failure. In order to identify honey bee species accurately, computer vision techniques can be utilized in this field. There are many computer vision techniques that provide fast results by imitating the processes performed by the human eye. Deep learning techniques, a sub-discipline of machine learning, are among the most popular computer vision techniques [2].

This study aims to evaluate the classification performance of the proposed deep learning models for honeybee species. The focus of the study is on developing two different deep learning models, applying them to honeybee images, and analyzing their performance. In this context, the study will both provide information on how deep learning techniques can be applied to biological datasets and reveal the potential and limitations of these technologies in the classification of honey bee species.

The significant contributions of this study can be summarized as follows:

- Two different deep learning models based on classical CNN (Model1) and CNN-ResNet (Model2) architectures have been proposed for monitoring honey bees, which are critical for ecosystem sustainability and agricultural production.
- The proposed Model1 and Model2 achieved successful results in classifying honey bee species quickly, accurately, and at low cost, with 94.91% and 93.63% accuracy rates, respectively.
- The results obtained show that these technologies can be used effectively in biodiversity conservation and smart agriculture applications.

The present study is organized as follows: Section 2 presents the related work in the literature. Section 3 describes the honey bee image dataset and introduces the proposed deep learning models. Section 4 presents the experimental studies and results. In the last section, the results are discussed and recommendations are given.

II. RELATED WORKS

Classification of honey bee species is of crucial importance for the conservation of biodiversity and increasing agricultural productivity. Computer scientists have conducted intensive studies on this species, which is extremely important for the ecosystem. In most of these studies, deep learning techniques are effective in the classification of honey bee species. CNN, one of the deep learning techniques, has significant potential in such classification tasks.

De Nart et al. proposed a convolutional neural network-based model for discriminating between honey bee subspecies using image recognition techniques applied to honey bee wings [3]. Another study examining the applicability of deep learning methods in the classification of honey bee species showed that a CNN-based model can classify honey bee images collected from different geographical regions with high accuracy rates [4]. In their study, deep learning models trained by taking morphological variations into account provided faster and more accurate results compared to traditional methods. The technique proposed by Karthiga et al. aims to classify various honey bee species and identify diseases that predispose honey bees [5]. Biscaro et al. developed a data augmentation and transfer learning-supported method based on the wing morphology of 47 bee species using CNN [6]. Using pre-trained deep learning models, high accuracy rates were quickly achieved regardless of the size of the dataset. The study highlights the adaptability of deep learning models and how they can be optimized on different datasets. Yoo et al. proposed a transformer-based network architecture called BeeNet for classifying the health status of bees [7]. This architecture consists of two main parts. In the first part, deep features were extracted using the ResNet50 network architecture, and in the second part, these deep features were classified using the Transformer model. A more comprehensive literature survey on honey bee images is given in Table I. Table I gives an overview of the studies detailing the variables analyzed, data types, methods, advantages, and disadvantages. The Metrics column gives accuracy and other performance values derived from relevant studies. For studies that do not use such metrics, the corresponding fields are left blank to preserve data integrity.

When the studies given Table I are analyzed, it is seen that they focus on honey bee monitoring, disease detection, and bee classification. Most of these studies in the literature use image-based datasets. It is evident that different machine learning and deep learning techniques such as CNN, Recurrent neural network (RNN), support vector machines (SVM), k-nearest neighbors (KNN), and Artificial neural network (ANN) are used in these studies. Some of the studies include specific topics such as real-time monitoring, disease detection, and bee counting, while others include more general topics such as hive health and classifying different bee species.

However, approaches in the literature seem to have some limitations. While most of these approaches produce solutions for monitoring bee movements or hive health, the detection of subspecies has been largely ignored. The proposed study makes an important contribution to beekeeping and biodiversity research by focusing directly on the classification of honey bee subspecies rather than just general hive monitoring or disease detection.

III. MATERIAL AND METHODS

The material and methods section of the study consists of 2 parts. In the first part, the dataset is introduced, and in the second part, the deep learning architectures used in the study are introduced. The general block diagram of the study is given in Fig. 1.

A. Dataset

The dataset used in the study consists of a total of 5172 bee images collected from different regions of the United States in 2018 [19]. These images are taken from bee videos recorded by Stanford University and shared under a public license. Each image contains only one bee and the background has been edited to ensure a clear view of the bee. Fig. 2. shows sample honey bee images from the dataset. In addition to the bee images, the dataset also contains the date and time of the images, the location of the images, and the species information of the bees.

TABLE I. DEEP LEARNING–BASED STUDIES ON HONEY BEE

References	Variables Analysed	Type of Data	Methods	Advantages	Disadvantages	Metrics
Chen et al., 2012 [8]	Beehive monitoring (bee counting)	Image	Principal component analysis, support vector machines (SVM)	Efficient in detecting bee movement patterns and activity changes.	Limited to static images; lacks real-time processing.	Accuracy: 86% and 98%
Ngo et al., 2019 [9]	Monitoring bee activities	Image, video	Kalman filter, Hungarian algorithm	Allows real-time tracking of multiple bees with improved accuracy.	Computationally expensive and requires high processing power.	Accuracy: 93.9%
Sevli, 2019 [10]	Bee prediction	Image	CNN	Effective for predicting bee species based on morphological features.	Limited dataset may affect generalization of the model.	Accuracy: 91.1% Recall: 78%–100% Precision: 65%–100% F1-score:71%–97%
Sun and Gaydecki, 2021 [11]	Monitoring bee flight behavior	Image, video	Gaussian mixture model	Enables 3D flight trajectory reconstruction for behavioral studies.	High complexity due to 3D reconstruction process.	–
Braga et al., 2021 [12]	Bee health	Image	CNN	Provides an intelligent system for hive health monitoring.	Depends on high-quality images for reliable results.	Bee health classification accuracy: 95% Bee detection accuracy: 82%
Dghim et al., 2021 [13]	Nosema disease detection	Image	Artificial neural networks, CNN, SVM	Capable of detecting Nosema disease with high accuracy.	Requires extensive training data for high performance.	Accuracy: 96.25%
Kulyukin, 2021 [14]	Beehive monitoring	Image	K-nearest neighbors (KNN), SVM, random forest, CNN	Combines multiple classifiers for robust hive monitoring.	High computational cost due to multiple classifiers.	Accuracy: 99.08%
Wachowicz et al., 2022 [15]	Varroa detection	Image	KNN, CNN	Improves accuracy in detecting Varroa mite infestation.	Sensitive to image noise and variations.	Accuracy: 80% Precision: 89% Sensitivity: 68% Specificity: 92%
Andrijević et al., 2022 [16]	Bee counting	Image	Recurrent neural networks, long short-term memory	Accurately predicts bee population trends over time.	Requires large datasets for effective training.	Root mean square error (RMSE): 378
Rigakis et al., 2023 [17]	Beehive monitoring	Sensor signals	Kalman filter, Hungarian algorithm	Uses low-cost multisensory devices for energy-efficient monitoring.	Limited scalability to large beekeeping operations.	–
Üzen et al., 2024 [18]	Bee disease classification	Image	VGG16, ResNet101	Ensemble learning improves classification performance for bee diseases.	May require extensive computational resources for inference.	Accuracy: 99.31% Recall: 98.92% Precision: 98.83% F1-score: 98.87%

The images in the dataset have different resolutions and are not of standard quality. It has been determined that the average values for image sizes are between 50 and 100 pixels, while the extreme values can range from a few pixels to approximately 500 pixels. This causes inconsistencies in the image processing process. In order to provide efficient learning for the proposed deep learning models, the resolutions should be standardized. For this purpose, before starting the learning process, all images are rescaled to 64 × 64 pixels to provide a balanced and optimized data input to the deep learning models. The scaling process was performed both to reduce the computational cost and to improve the performance of the model by preserving the distinctive features of the images.

The original images were colored and processed as 3-channel images without any modification during the data processing. The dataset contains a total of six different bee species, including five different bee subspecies and one species mixed with the local race. The "Italian honey bee" species has the highest number of data with 3008. This is followed by the Russian honey bee (527) and the Carniolan honey bee (501). In addition, there are 428 undefined

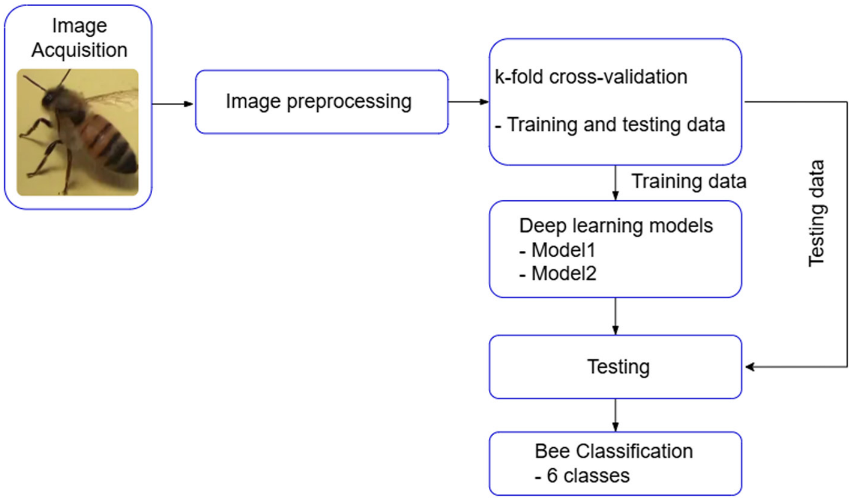

Fig. 1. Block diagram of the proposed bee classification.

species in the dataset that were not used in the study. To eliminate the imbalance in the number of images between the classes in the dataset, the number of samples in the "Italian honey bee" class was subsampled and mapped to the "Russian honey bee" species. The distribution of the number of images corresponding to each class in the dataset and the distribution of images in the balanced dataset after the preprocessing step are shown in Fig. 3.

The dataset used in the study consists of images obtained under different environmental conditions (light, shadow, camera angle). This situation necessitated the application of pre-processing and screening processes on the images. Low-resolution images, the absence of a bee in the image, or the presence of a very small part of the bee in the image led to the elimination process in the images. Fig. 4. shows examples of images extracted from the dataset. Since a large number of meaningful and labeled data are needed for the successful application of deep learning techniques, the elimination phase has become a necessity [20].

The aim of the study is to accurately detect the honey bee subspecies using the images in the existing dataset. For this purpose, the data is divided into training and test sets to test how accurately all images

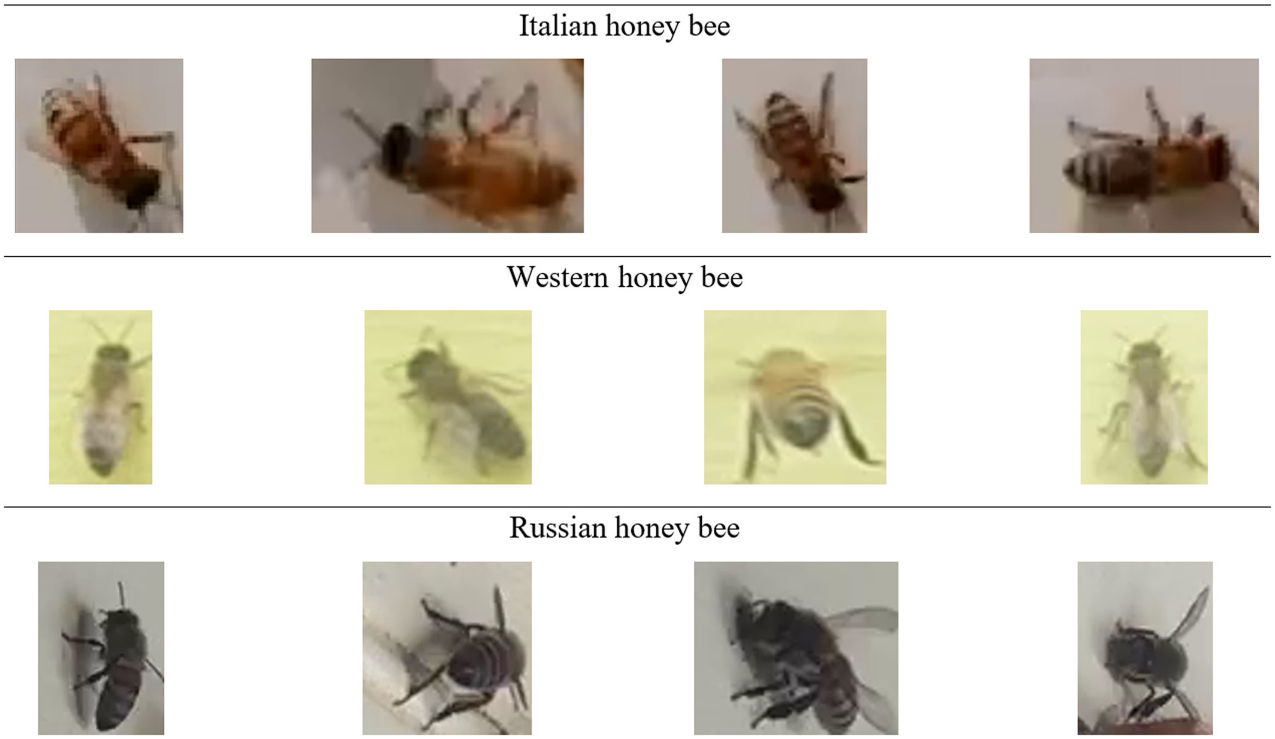

Fig. 2. Example bee subspecies in the dataset.

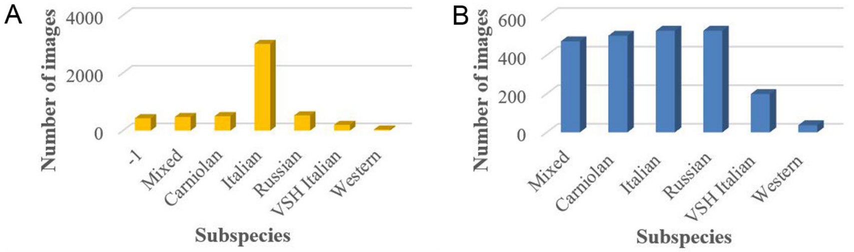

Fig. 3. A) Honey bee subspecies numbers B) Balanced dataset.

in the dataset are classified. The training set will be used to train the model, while the test set will be used to test the accuracy of the model against new, fresh data not used in training. This distribution of the dataset is first divided into two sets: 80% training set and 20% test set.

B. Used Deep Learning Architectures

In this study, two different deep learning models are proposed to classify honey bee images. Deep learning is a branch of machine learning based on deep neural networks, including representation learning [20]. Deep neural networks consist of artificial neural networks with many hidden layers. In deep learning, especially CNNs are utilized for feature extraction from images. CNNs are used to create feature maps from input images by applying filters to the input images. Although there are many different CNN architectures, the general CNN structure consists of adding several convolutional and pooling layers, one after the input layer where the image is fed.

The first model (Model1) used in the study is based on a layered CNN structure. Input images ranging between 0–255 are converted into intensity pixel values represented as a matrix through RGB (red, green, blue) channels. In the first stage, the pixels forming the image are scanned from left to right and from top to bottom with the help of filters defined in the CNN layer. The filter is defined in the first step to define the convolution. The output of the first layer is given as input to successive layers, which act as a feedback mechanism. The rectified linear unit activation function, which is faster than activation functions such as Hyperbolic Tangent and Sigmoid, eliminates negative values from the convolution layer and transfers positive values to the next layer. In the deep learning architecture proposed with Model1, the spatial dimension of the feature map is reduced by using the maximum pooling layer after each convolution layer. The feature maps obtained after the CNN layers are reduced to one dimension and presented as input to the fully connected layer. Finally, all layers are connected to the output layer which produces an n-dimensional vector. The architecture of Model1 used for the classification of honey bee subspecies is shown in Fig. 5. There is no universal optimal model for all datasets. For different problems, models with different structures are designed to achieve the best performance. For more information about the architectures of CNN-based deep learning models, please refer to the related literature [21],[22].

In the second model (Model2) used in the study, a layered CNN structure and Residual blocks are used. ResNet is an architecture that significantly improves image recognition performance [23]. ResNet is a deep learning architecture that uses skip connections as shown in Fig. 6. This architecture, which uses residual blocks, directly transfers the output of the previous layer to the next layer. With residual blocks, inputs are transferred faster over the remaining connections between layers. Here, the input x is multiplied by the weights of the layer and then passes through the activation function $f(x)$, and the output is taken as $H(x)$.

In Model2, unlike existing deep learning methods, a new CNN model is proposed that takes advantage of the ResNet model. The CNN-ResNet architecture proposed with Model2 is shown in Fig. 7.

C. Cross-Validation Technique

The K-fold cross-validation technique is used to evaluate both deep learning models used in the study. The purpose of the K-fold technique is to divide the data into groups and treat each group as a test case to evaluate the model [24]. The remaining data groups are used as training data. The K-fold process as an evaluation model can be expressed as in Fig. 8.

In this study, the stratified K-fold technique was used to avoid an unbalanced distribution of the dataset and to increase the generalization ability of the model. In the classical K-fold technique, some classes are likely to be missing or imbalanced, especially in

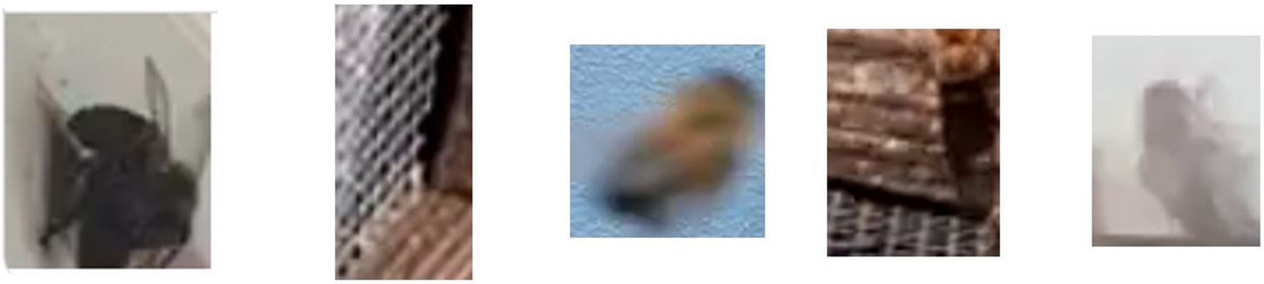

Fig. 4. Some examples of images eliminated from the dataset.

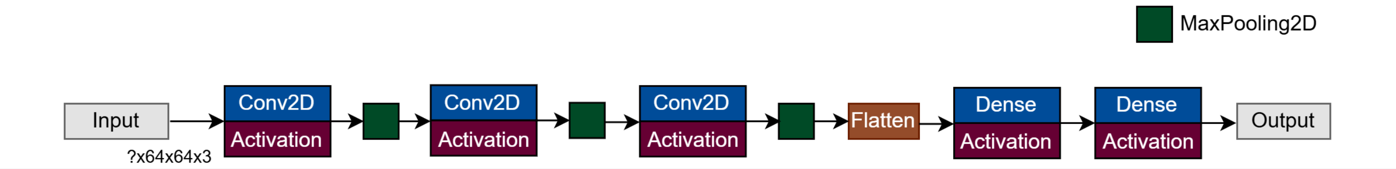

Fig. 5. Model1 architecture used in the study.

unbalanced datasets. With the stratified K-fold technique, a balanced distribution of each subtype in the training and test datasets is ensured.

D. Validation Metrics

Sensitivity, recall, F1-score, and accuracy metrics are utilized in this study to comprehensively evaluate the performance of the proposed models [25]. These metrics are widely used to evaluate the performance of classification models. The metrics used are shown in (1), (2), (3), and (4) [26].

$$\text{Precision} = \frac{TP}{TP+FP} \quad (1)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (2)$$

$$\text{f1}-\text{score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (3)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

With these metrics, the overall classification performance was evaluated by analyzing the sensitivity of the models to false positive and false negative predictions.

IV. EXPERIMENTAL RESULTS

In this study, basic CNN and CNN-ResNet based deep learning models named Model1, Model2 were used for honey bee subspecies classification. The experiments were performed on a computer system equipped with an Intel i7-5500U 2.4 GHz CPU, Nvidia GeForce 840M GPU, and 12 GB RAM. Since the dataset used in the study contains images with different resolutions and quality, the stratified K-fold technique was used instead of randomly separating training and test data. With this technique, the generalization ability of the models can be increased by ensuring that the distribution between different honey bee subspecies is consistent between training and test sets. For this purpose, the dataset is divided into 5 equal parts, and each time 80% of the dataset is used as training data and 20% as test data.

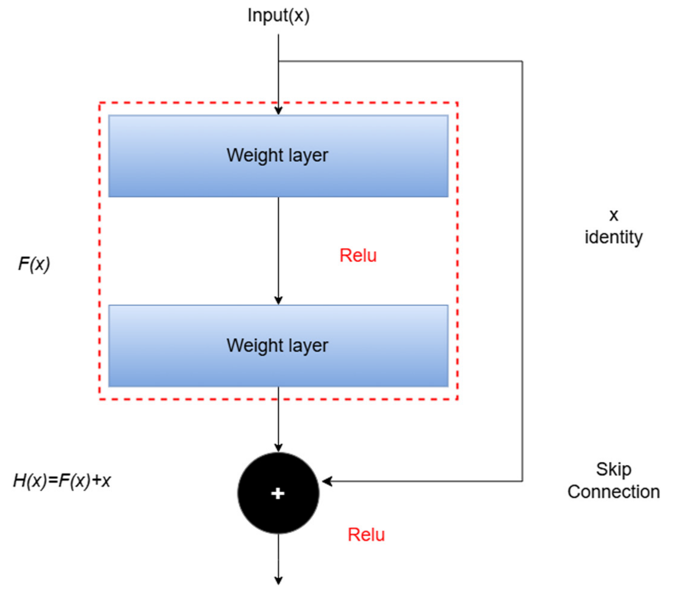

Fig. 6. Core building block of deep residual networks.

The hyperparameters and value ranges used in Model1 and Model2 were chosen to be the same and performance comparisons were made. Table II shows the hyperparameters and test ranges used in both models.

Rectified linear unit is used as an activation function in the models used. The Softmax function used in the last layer ensures that the input image of the system is included in only one class from six different subtypes.

Model1 has 1 784 646 trainable parameters while Model2 has 752 710 trainable parameters. This difference in the number of parameters shows that Model2 can provide less memory usage and faster inference time. However, due to its lower capacity, it may lose performance in certain classification tasks. Model1, which has more parameters, has an advantage in learning more complex features.

In the study, training and test times were also analyzed in order to determine the inference times as well as the performance of the models. For this purpose, training and test times for Model1 and Model2 for honey bee subspecies classification were measured under the same hardware conditions. Training time was calculated for 1810 samples. In order to evaluate the suitability of the models for real-time applications, the test time was calculated for a single image sample. It was observed that Model1 completed the training process in approximately 2873 seconds and Model2 completed the training process in 620 seconds. The approximate test time of Model1 was 5.8 ms, while the approximate test time of Model2 was 1.3 ms. It is seen that Model2 is more advantageous in classification applications that require speed.

The results obtained in the study were analyzed in order to compare the classification performance of Model1 and Model2 in detail and to analyze their generalization capabilities. In the cross-validation process, the validation results in the second fold (Fold2) were the highest in both models, and the results obtained for each class are given in Table III.

Due to the unbalanced distribution of the number of images of honey bee subspecies in the dataset, other metrics should be taken into account in addition to the Accuracy metric. When Table III is analyzed, the F1-score metric reveals the overall performance and balanced classification ability of the models. As F1-score is the

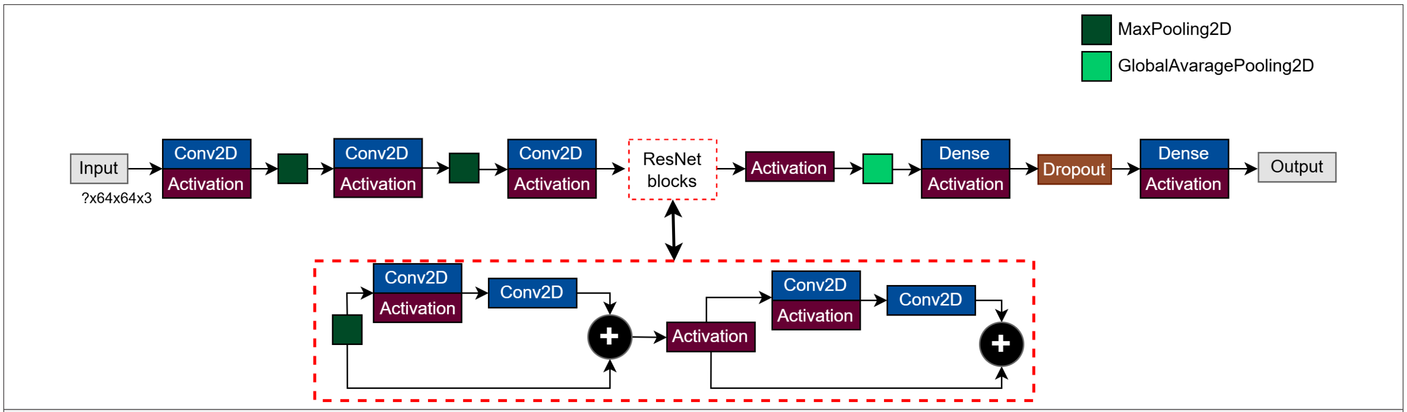

Fig. 7. Model2 architecture used in the study.

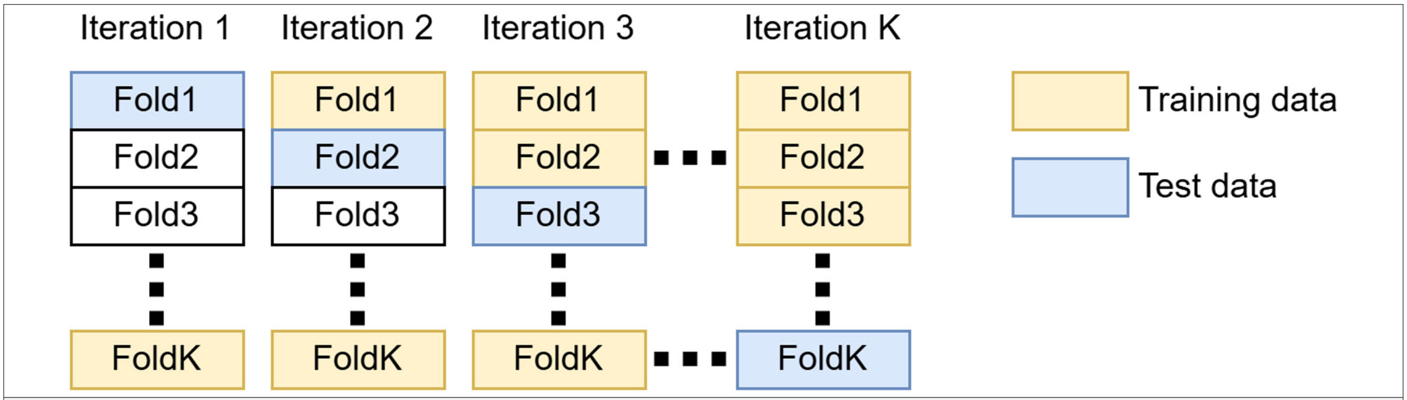

Fig. 8. K-fold diagrams.

harmonic mean of precision and recall values, it is a balanced metric that reflects both the rate of correct positive predictions and the effect of false negatives. The high F1-score values obtained despite the imbalanced dataset indicate that the generalization ability of the models is strong in all classes. As seen in Table III, it is seen that the generalization capacity of both models is high in the range of 93%–100% in Mixed, Carniolan, Italian, and Russian bee subspecies with a high number of data. Compared to the bee subspecies with a high number of data, VSH Italian and Western subspecies contain less data. However, when the metrics of these subspecies are analyzed, it is seen that the models can successfully discriminate these subspecies as well. Fig. 9. shows the average F1-score values of both models.

The accuracy values and averages of both models used in the study in the cross-validation steps are shown in Fig. 10. When the figure is analyzed, it is seen that the accuracy values are similar to the F1-score values.

TABLE II. HYPERPARAMETER VALUES TESTED FOR THE MODELS USED

Hyperparameters	Tested Range	Best Value
Number of epocs	30–100	80
Learning rate	0.0001–0.1	0.001
Batch size	16-64	32

In order to evaluate whether the difference between the accuracy rates of both deep learning models proposed in the study is statistically significant, a paired *t*-test was applied [27]. The dependent *t*-test results were calculated using the accuracy rates obtained for both models in the cross-validation steps shown in Fig. 10. As a result of these statistical procedures, the *t*-statistic was 3.51 and the p-value was 0.0246. Since $p < 0.05$, it was concluded that the difference in accuracy ratio between Model1 and Model2 was significant. These results show that Model1 is more performant than Model2.

Confusion matrices were also used to evaluate the overall performance of both models. When examining the confusion matrices in Fold 2 in Fig. 11, where the highest success is achieved, it is seen that Italian and Mixed bee species are predicted interchangeably in both model types. This situation can be explained by the presence of mixed bee species with local races in the dataset. When the results are analyzed in Model1, it is seen that 97 correct and 1 incorrect prediction were made in the Mixed class. In addition, it is seen that Carniolan, Russian, and Western species are completely correctly classified. Compared to the other classes of Model1, Italian, VSH Italian, and Mixed subspecies were predicted interchangeably. Similarities arising from genetic and environmental factors can create common points in the morphological structures of these species. Especially, the similarities between the wing structures, color patterns, and body sizes of bee subspecies may make it difficult to visually distinguish these species from each other. When the

TABLE III. HIGHEST CLASSIFICATION RESULTS OBTAINED IN CROSS-VALIDATION

		Mixed	Carniolan	Italian	Russian	Varroa sensitive hygiene (VSH) Italian	Western
Model1	Accuracy	99	100	89	100	95	100
	Precision (%)	92	100	96	100	98	100
	Recall (%)	99	100	89	100	95	100
	F1-score (%)	95	100	93	100	96	100
Model2	Accuracy	88	100	95	100	95	100
	Precision (%)	95	96	85	100	100	100
	Recall (%)	88	100	95	95	95	100
	F1-score(%)	91	98	90	98	98	100

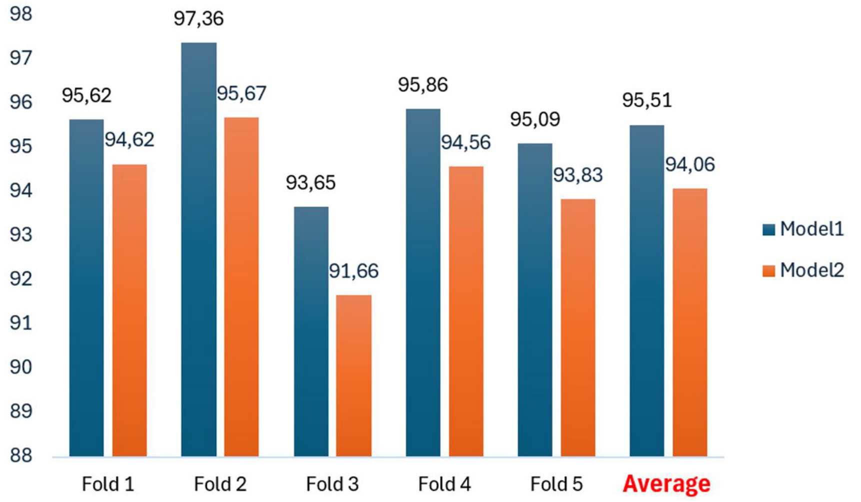

Fig. 9. F1-score values of the models.

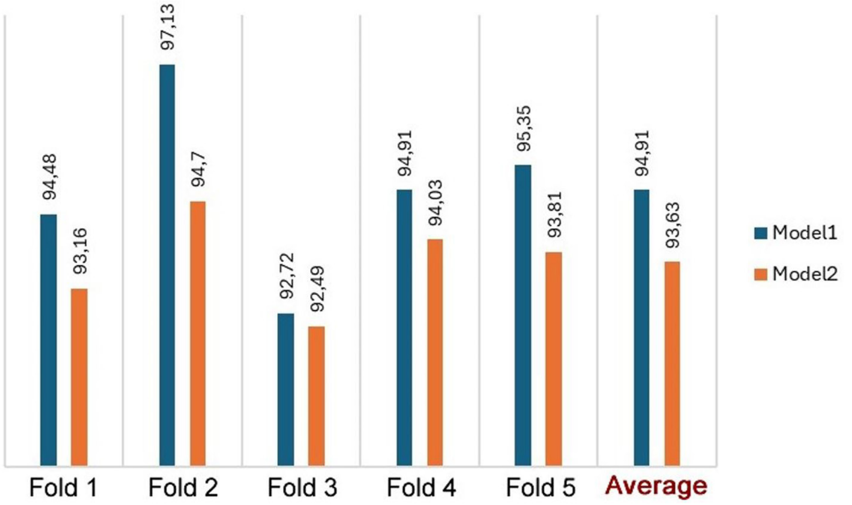

Fig. 10. Accuracy values and averages of the models.

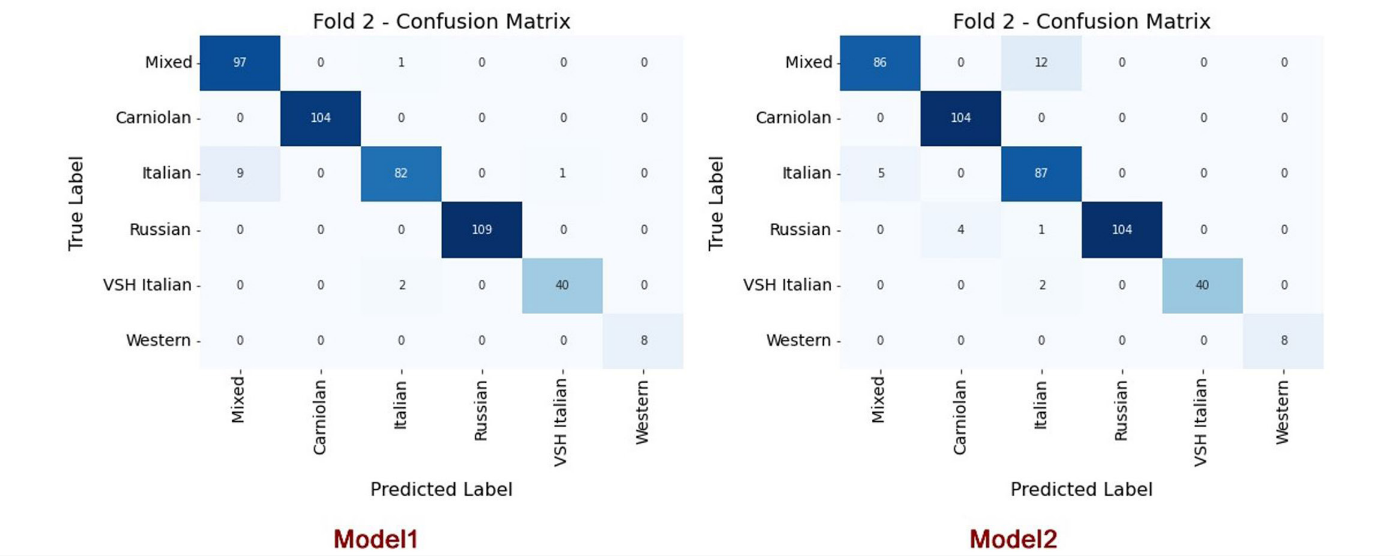

Fig. 11. Confusion matrices of the models in Fold 2, where the highest performance is achieved.

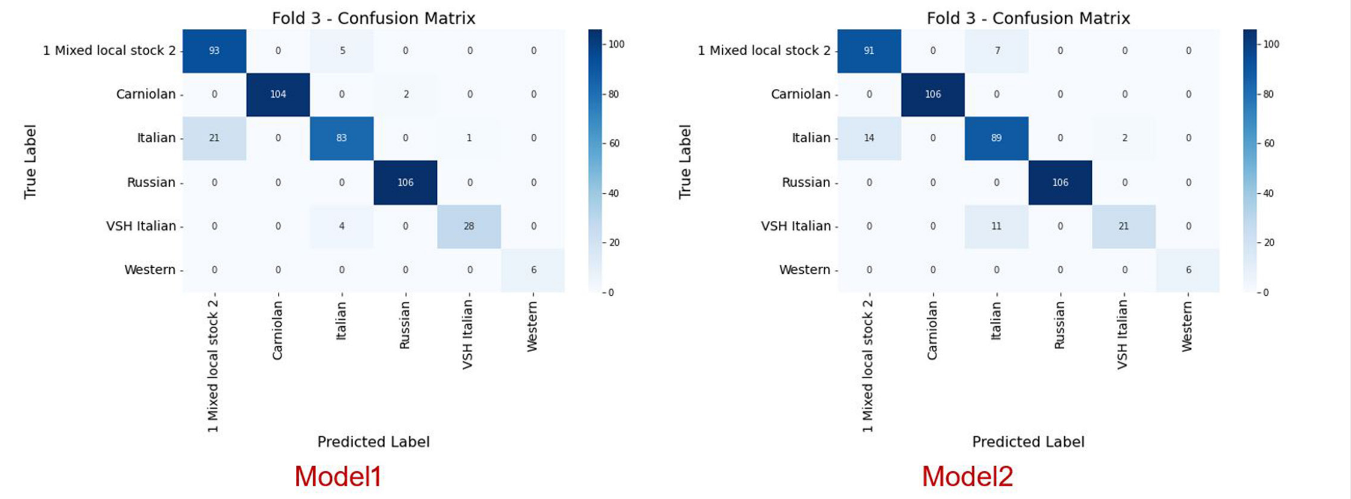

Fig. 12. Confusion matrices of the models in Fold 3 with the lowest performance.

Fig. 13. A) Training image samples, B) Test image samples.

results of Model2 are examined, it is seen that the accuracy rates for Carniolan and Western species are similar to Model1. In the Italian bee subspecies, although the number of false negatives is lower than Model1, the classification performance of Model1 is higher than Model2.

When the Confusion matrices in Fold 3, where the lowest success is achieved in Fig. 12. it is seen that Italian, VSH Italian, and Mixed bee subspecies are predicted instead of each other, similar to Fold 2, where the highest success is achieved. It is seen that the number of false negatives in the Italian bee subspecies increased to 21. This shows that the generalization ability of Model1 in this class decreases significantly in Fold 3. As a result, it can be said that the generalization ability of Model1 is superior to Model2 in all cases.

When the Confusion matrices of the models are examined, it is seen that the generalization ability of the models is low, especially in some classes. For a better understanding of this situation, sample training and test images used in the study are given in Fig. 13

The accuracy and F1-score rates of our study are quite high compared to other studies focusing on the same topic. This shows that the proposed models provide an effective classification method. For example, in the study by Sevli [10], CNN was used to classify honey bee subspecies, and an accuracy rate of 91.1% was obtained. On the other hand, in the proposed study, Model1 and Model2 showed a higher performance with 94.91% and 93.63% accuracy rates, respectively. When examining Table III, it is observed that the class-wise accuracy values in Model1 range between 88% and 100%. These results show that the generalization capability of both models is good. In our study, balanced partitioning of the dataset with the stratified K-fold technique and hyperparameter optimization were important in improving the generalization ability of both models. Furthermore, aside from Sevli [10], the other studies presented in Table I focus on different data types (e.g., bee tracking, segmentation, disease detection) and employ various analytical approaches, making direct comparisons infeasible. However, the high accuracy rates achieved by the proposed models in our study clearly demonstrate that they provide a reliable and effective approach for honey bee species classification.

V. CONCLUSION

In this study, which aims to classify various honey bee species from images using deep learning techniques, two different models are proposed. The first model (Model1) presents a structure based on the classical CNN architecture, while the second model (Model2) is based on the integration of the CNN-ResNet architecture. In the validation studies performed with the K-fold technique, it was observed that Model1 achieved an average of 94.91% and Model2 achieved an average of 93.63%. The obtained results show that it is possible to classify honey bee species in a fast and low-cost way with image recognition techniques.

Considering the high performances of these models, it is thought that they can make important contributions in terms of monitoring and protecting bee populations and realizing efficient beekeeping. However, the low number of layers of the proposed models shows that they can be used in real-time smart agriculture and beekeeping applications. In the future, it is planned to develop new models based on data fusion by expanding the dataset.

Data Availability Statement: The data that support the findings of this study are available on request from the corresponding author.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept – M.Ç., N.O.; Design – M.Ç., N.O.; Supervision – M.Ç.; Data Collection and/or Processing – N.O.; Analysis and/or Interpretation – M.Ç., N.O.; Literature Review – M.Ç., N.O.; Writing – M.Ç.; Critical Review – N.O.

Declaration of Interests: The authors have no conflicts of interest to declare.

Funding: The authors declare that this study received no financial support.

REFERENCES

1. F. Ruttner, *Biogeography and Taxonomy of Honeybees (eBook ed.)*. Springer Science & Business Media, 2013. (https://doi.org/10.1007/978-3-642-72649-1)
2. G. Kambale, and N. Bilgi, “A survey paper on crop disease identification and classification using pattern recognition and digital image processing techniques,” *IOSR* , vol. 4, pp. 14–17, 2017.
3. D. De Nart, C. Costa, G. Di Prisco, and E. Carpana, “Image recognition using convolutional neural networks for classification of honey bee subspecies,” *Apidologie*, vol. 53, no. 1, 2022. [CrossRef]
4. A. Ratnayake, H. M. Yasin, A. Ghani Naim, and P. E. Abas, “Classification of subspecies of honey bees using convolutional neural network,” in 2023 6th International Conference on Applied Computational Intelligence in Information Systems (ACIIS), Bandar Seri Begawan, Brunei Darussalam, 2023, pp. 1–6. [CrossRef]
5. M. Karthiga, S. Sountharrajan, S. S. Nandhini, E. Suganya, and S. Sankarananth, “A Deep Learning Approach to classify the Honeybee Species and health identification,” in 2021 Seventh International Conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 2021, pp. 1–7. (https://doi.org/10.1109/ICBSII51839.2021.9445173)
6. H. Bíscaro, D. S. Assunção, L. A. Digiampietri, and T. M. Francoy, *A Fully Automatic Classification of Bee Species Using CNN with Data Augmentation and Transfer Learning Techniques (Preprint)*, SSRN, 2023. ()https://doi.org/10.2139/ssrn.4658136
7. J. Yoo, R. Siddiqua, X. Liu, K. A. Ahmed, and M. Z. Hossain, “BeeNet: An end-to-end deep network for bee surveillance,” *Procedia Comput. Sci.*, vol. 222, pp. 415–424, 2023. [CrossRef]
8. C. Chen, E. C. Yang, J. A. Jiang, and T. T. Lin, “An imaging system for monitoring the in-and-out activity of honey bees,” *Comput. Electron. Agric.*, vol. 89, pp. 100–109, 2012. [CrossRef]
9. T. N. Ngo, K. C. Wu, E. C. Yang, and T. T. Lin, “A real-time imaging system for multiple honey bee tracking and activity monitoring,” *Comput. Electron. Agric.*, vol. 163, p. 104841, 2019. [CrossRef]
10. O. Sevli, “Evrişimsel Sinir Ağları ile Bal Arısı Irklarının Tahminlenmesi,” in Proceedings on 2nd International Conference on Technology and Science, Burdur, Türkiye, Nov. 2019.
11. C. Sun, and P. Gaydecki, “A visual tracking system for honey bee (hymenoptera: Apidae) 3D flight trajectory reconstruction and analysis,” *J. Insect Sci.*, vol. 21, no. 2, p. 17, 2021. [CrossRef]
12. D. Braga, A. Madureira, F. Scotti, V. Piuri, and A. Abraham, “An intelligent monitoring system for assessing bee hive health,” *IEEE Access*, vol. 9, pp. 89009–89019, 2021. [CrossRef]
13. S. Dghim, C. M. Travieso-González, and R. Burget, “Analysis of the Nosema cells identification for microscopic images,” *Sensors (Basel)*, vol. 21, no. 9, p. 3068, 2021. [CrossRef]
14. V. Kulyukin, “Audio, image, video, and weather datasets for continuous electronic beehive monitoring,“ *Appl. Sci.*, vol. 11, no. 10, p. 4632, 2021. [CrossRef]
15. A. Wachowicz, J. Pytlik, B. Małysiak-Mrozek, K. Tokarz, and D. Mrozek, “Edge computing in IoT-enabled honeybee monitoring for the detection of Varroa destructor,” *Int. J. Appl. Math. Comput. Sci.*, vol. 32, no. 3, 2022. [CrossRef]
16. N. Andrijević, V. Urošević, B. Arsić, D. Herceg, and B. Savić, “IoT monitoring and prediction modeling of honeybee activity with alarm,” *Electronics*, vol. 11, no. 5, p. 783, 2022. [CrossRef]

17. I. Rigakis, I. Potamitis, N. A. Tatlas, G. Psirofonia, E. Tzagaraki, and E. Alissandrakis, "A low-cost, low-power, multisensory device and multivariable time series prediction for beehive health monitoring," *Sensors (Basel)*, vol. 23, no. 3, p. 1407, 2023. [CrossRef]
18. H. Üzen, M. Altın, and İ. Balıkçı Çiçek, "Bal Arı Hastalıklarının Sınıflandırılması için ConvMixer, VGG16 ve ResNet101 Tabanlı Topluluk Öğrenme Yaklaşımı," *Mühendislik Bilimleri Derg.*, vol. 36, no. 1, pp. 133–145, 2024. [CrossRef]
19. J. Yang, "*The BeeImage dataset: Annotated honey bee images | Kaggle*. Accessed: Mar. 2024. [Online]. Available: https://www.kaggle.com/jenny18/honey-bee-annotated-images.
20. J. J. Lv, X. H. Shao, J. S. Huang, X. D. Zhou, and X. Zhou, "Data augmentation for face recognition," *Neurocomputing*, vol. 230, pp. 184–196, 2017. [CrossRef]
21. K. O'Shea, "An introduction to convolutional neural networks," *arXiv Preprint ArXiv*:1511.08458, 2015.
22. Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, 2015. [CrossRef]
23. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2016, pp. 770–778. [CrossRef]
24. N. Olgun, and I. Türkoğlu, "Classification of live/lifeless assets with laser beams in different humidity environments," in 2020 8th International Symposium on Digital Forensics and Security (ISDFS), pp. 1–6. [CrossRef]
25. M. Çalışan, V. Gündüzalp, and N. Olgun, "Evaluation of U-net and ResNet architectures for biomedical image segmentation," *International Journal of 3D Printing Technologies and Digital Industry*, vol. 7, no. 3, pp. 561–570, 2023. [CrossRef]
26. M. Irfan et al., "Energy theft identification using Adaboost ensembler in the smart grids," *Comput. Mater. Continua*, vol. 72, no. 1, pp. 2141–2158, 2022. [CrossRef]
27. A. Yala, T. Schuster, R. Miles, R. Barzilay, and C. Lehman, "A deep learning model to triage screening mammograms: A simulation study," *Radiology*, vol. 293, no. 1, pp. 38–46, 2019. [CrossRef]

Mücahit ÇALIŞAN was born in Bingöl, Turkey in 1985. He received the B.S. degree in computer science from Fırat University, Elazığ, Turkey, in 2007. He received the Ph.D. degree in computer engineering from İnönü University in 2021. His current interests include machine learning, pattern recognition, and computer vision.

Nevzat OLGUN received his master's degree in Electronics and Computer Education from Fırat University in 2014. He received his PhD degree in Software Engineering from Fırat University in 2022. His research interests are brain-computer interaction, deep learning, machine learning, and signal processing.