J. Entomol. Res. Soc., 25(3): 529-544, 2023
Doi: 10.51963/jers.v25i3.2445

Research Article
Online ISSN:2651-3579

Deep Learning Based Classification for Hoverflies (Diptera: Syrphidae)

Anıl UTKU[1] Zafer AYAZ[2] Derya ÇİFTÇİ[3*] M. Ali AKCAYOL[4]

[1]Faculty of Engineering, Department of Computer Engineering, Munzur University, Tunceli, TÜRKİYE

[2]Department of Management Information Systems, Gazi University, Ankara, TÜRKİYE

[3]Veysel Karani District, University Housing Complex 3E Block No:8, Siirt, TÜRKİYE

[4]Faculty of Engineering, Department of Computer Engineering, Gazi University, Ankara, TÜRKİYE

e-mails: [1]anilutku@munzur.edu.tr, [2]zafer@gazi.edu.tr, [3*]dcanpolat@gmail.com [4]akcayol@gazi.edu.tr

ORCID IDs: [1]0000-0002-7240-8713, [2]0000-0001-7667-7921, [3]0000-0002-7670-4392 [4]0000-0002-6615-1237

*Corresponding author

ABSTRACT

Syrphidae is essential in pollinating many flowering plants and cereals and is a family with high species diversity in the order Diptera. These family species are also used in biodiversity and conservation studies. This study proposes an image-based CNN model for easy, fast, and accurate identification of Syrphidae species. Seven hundred twenty-seven hoverfly images were used to train and test the developed deep-learning model. Four hundred seventy-nine of these images were allocated to the training set and two hundred forty-eight to the test dataset. There are a total of 15 species in the dataset. With the CNN-based deep learning model developed in this study, accuracy 0.96, precision 0.97, recall 0.96, and f-measure 0.96 values were obtained for the dataset. The experimental results showed that the proposed CNN-based deep learning model had a high success rate in distinguishing the Syrphidae species.

Keywords: Convolutional neural network, image classification, automatic species identification

Utku, A., Ayaz, A., Çiftçi, D., & Akcayol, M.A (2023). Deep Learning Based Classification for Hoverflies (Diptera: Sryphidae). *Journal of the Entomological Research Society*, 25(3), 529-544.

Received: April 09 2023 *Accepted: August 21, 2023*

INTRODUCTION

Insects are a class within the arthropods and have an exoskeleton and a characteristic body structure consisting of 3 parts (head, thorax, and abdomen) and three pairs of legs, compound eyes, and a pair of antennae (Martineau et al, 2017; Xia, Chen, Wang, Zhang, & Xie, 2018; Hassan, Rahman, Htike, & Win, 2014). Arthropods show more biodiversity than other groups of living things. Arthropods are important indicators of ecosystem function. It was used to determine the quality of many habitats (agriculture, forest, and meadow) and the fauna richness and diversity of habitats. It is also used to determine the extent of arthropod diversity, habitat fragmentation, and degradation (Martineau et al, 2017; Xia et al, 2018; Hassan et al, 2014).

Traditionally, insect species identification is based on morphological identification. Taxonomists and trained technicians identify using taxa-specific identification keys because species identification requires skills gained through training and experience. Although technicians can identify taxa using identification keys in some cases, some insect taxa also require experts. Furthermore, the need for sufficient experts and technicians in some insect groups delays the insect identification stage. Consequently, alternative and accurate identification methods are required, which at least non-experts can use (Martineau et al, 2017; Xia et al, 2018).

In biodiversity studies, there are other difficulties, in addition to the difficulty of finding an expert insect taxonomist. Some of these are the lack of up-to-date identification keys and catalogues, the scattered family or species-specific sources, the lack of arrangement of synonymous species names, the difficulty of identifying many taxa, and the collection of large numbers of specimens in field studies (Gaston and Neil, 2004; Gaston and May, 1992; Mound and Gaston, 1993).

These difficulties in taxonomy and identification have led to the development of identification methods in the last 30 years (Gaston and Neil, 2004). One of these solutions is the automatic identification process. Image-based insect recognition is widely used, especially in agriculture, ecology, and biodiversity (Hassan et al, 2014; Martineau et al, 2017; Karar, Alsunaydi, Albusaymi, & Alotaibi, 2021).

Fedor, Vaňhara, Havel, Malenovský, & Spellerberg (2009) identified 18 economically important common European species of Thysanoptera with 97% accuracy using the artificial neural network (ANN) model.

Yang, Ma, Wen, Zhan, & Wang (2015) developed a program to identify species with 90-98% accuracy using the wings of Neuroptera. Faria et al (2014) developed an identification method with more than 98% accuracy for Anastrepha fruit pest species, widely distributed in the American tropics and subtropics, using a multimodal fusion approach. The automatic bee identification system (ABIS) made species distinction with Support Vector Machine (SVM) using the front wing photograph of the bee (Arbuckle, Schröder, Steinhage, & Wittmann, 2001). O'Neill (2008) developed a successful invertebrate identification system with (the digital automated identification system) DAISY based on eigen-images recognition.

Zhu et al (2017) presented a study that explores the application of hybrid deep-learning techniques for the automated classification of lepidopteran insect images. The research focuses on leveraging the strengths of various deep learning architectures, combining them in a hybrid approach to achieve improved accuracy in insect classification. By combining Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), the authors create a robust framework capable of capturing spatial and sequential features in insect images. The study demonstrates the efficacy of this hybrid model in accurately classifying lepidopteran insects, offering a promising avenue for automating the identification process in entomology and pest management applications. The research contributes to advancing the field of insect image classification by proposing an innovative approach that effectively addresses the complexities associated with recognizing and categorizing diverse insect species.

Thenmozhi and Reddy (2019) proposed a novel approach for accurately classifying crop pests by utilizing advanced techniques in deep learning. The study employs CNN and transfer learning to enhance the classification accuracy for various crop pest species. By leveraging pre-trained models and fine-tuning them on a specialized dataset of crop pest images, the authors achieve significant improvements in classification performance. Transfer learning enables the network to learn intricate features from the data, resulting in robust and effective pest classification. The research outcomes demonstrate the potential of combining deep learning techniques and transfer learning in addressing complex agricultural challenges such as pest identification and management.

Li et al (2021) presented a comprehensive analysis of research endeavours centered around utilizing deep learning techniques for insect classification and detection in field images, specifically emphasizing applications in intelligent pest management systems. The study surveys a broad range of approaches, methodologies, and advancements in this domain through a systematic review. The authors identify common trends and challenges, highlighting the effectiveness of deep learning models in accurately identifying and categorizing insects from images captured in real-world agricultural environments. The review underscores the potential impact of such methods on improving pest management strategies by enabling timely and targeted interventions. By synthesizing the findings of multiple studies, this paper offers valuable insights into the state-of-the-art techniques and future directions in intelligent pest management using deep learning technologies.

Tiwari et al (2021) present an innovative study focusing on developing and implementing customized deep-learning models for real-time classification of insect pests in soybean crops. The research addresses the challenges the agricultural setting poses, where rapid and accurate insect identification is crucial for effective pest management. The authors create models capable of real-time insect classification by tailoring deep learning architectures to the specific features of soybean crop images and insect pests. This approach demonstrates remarkable accuracy and potentially significantly enhances pest detection and control strategies in soybean farming. The study showcases the importance of adapting deep learning methods to suit the demands of real-world agricultural applications, providing a valuable contribution to intelligent farming and pest management.

Kasinathan, Singaraju, & Uyyala (2020) comprehensively explore the application of contemporary machine-learning methods for insect classification and detection within field crops. The study delves into utilizing state-of-the-art machine learning techniques, such as CNN and SVM, to address the intricate challenge of identifying and detecting insects in complex agricultural environments. By harnessing the power of these advanced techniques, the authors develop effective models capable of accurate insect classification and timely pest detection. The research underscores the potential of modern machine learning approaches to revolutionize pest management strategies by providing timely insights and enabling proactive interventions in field crop protection.

Butterfly species, which are of great ecological importance, are one of the orders that show species richness and are used as indicators in biodiversity studies (Beccaloni and Gaston, 1995; Dennis et al, 2008; Dover, Sparks, Clarke, Gobbett, & Glossop, 2000; Zamora, Verdu, & Galante, 2007; Zupan, Bužan, Grubar, & Jugovic, 2020). Morphological features such as wing shape, structure, and color are used to identify butterfly species. Several automatic machine viewings were designed to make morphological identification accurate, easy, and faster (Kaya & Kaycı, 2014; Wang, Ji, Liang, & Decheng, 2011; Qing et al, 2012; Wen, Guyer, & Li, 2009).

Many methods were used in the literature to classify insects and insect pests (Martineau et al, 2017; Xia et al, 2018; Kasinathan et al, 2020). Classification methods used in the literature are divided into monolithic, combinations, and instance-based groups. Monolithic classification methods are divided into two levels discriminative and generative. The combinations method is divided into two basic levels: boosting and bagging. Instance-based methods use the k-nearest neighbours (k-NN) algorithm.

In discriminative methods, least-squares approximation (Wen et al, 2009; Wen and Zhu, 2010) and fisher linear discriminant (Dietrich and Pooley, 1994; Dietrich, Emigh, & Deitz, 1991; Zayas and Flinn, 1998; Tofiski, 2004; Francoy et al, 2008) were used in the literature to classify different insect species. However, in cases when the data could not be separated linearly, the support vector machine (Qing et al, 2012; Yang et al, 2015; Silva et al, 2015; Wang et al, 2012) was applied using the polynomial kernel or the radial basis function with standard deviation. Neural network architectures were also used as parser classifiers for different insect species by using nonlinear activation functions in the literature (Do, Harp, & Norris, 1999; Al-Saqer and Hassan, 2011; Wang, Lin, Ji, & Liang, 2012; Wen, Wu, Hu, & Pan, 2015; Leow, Chew, Chong, & Dhillon, 2015; Silva, Grassi Sella, Francoy, & Costa, 2015; Xia et al, 2018). Decision tree-based classifiers were also used to classify insect species (Mayo and Watson, 2007; Larios et al, 2008; Silva et al, 2015). However, there are no studies to classify hoverflies using deep learning.

The family Syrphidae of the order Diptera are commonly known as “hoverflies” or “flower flies.” This family is among the insects that visit flowering plants the most. Hoverflies are important pollinators of many plants and crops (Klecka, 2018). The fact that most syrphids are effective natural enemies of different insect groups, especially aphids, allows them to be used in biological control studies (Van Driesche, Hoddle, & Enter, 2008). Adults of all known syrphid species feed almost exclusively on pollen, nectar, or honeydew (Rotheray and Gilbert, 2011) and are usually considered general visitors to the flowers (Klecka, 2018). The fact that they are involved in biodiversity conservation studies, that they are widely distributed, and that various ecological conditions are required for their larvae support the use of this fly family as a bioindicator (Sommaggio, 1999; Sommaggio and Burgio, 2014).

Morphological characters such as abdomen pattern, antenna segment colour, and length, leg colour are used in species identification, but male genitalia should be examined for definitive identification. (Speight, 2018, 2020). In recent years, diagnostic studies with molecular methods have been conducted, especially in problematic groups (Vujić et al, 2015, 2017, 2020; Likov et al, 2020; Kočiš et al, 2018). The identification with both classical methods and molecular methods takes a long time.

In recent years, there have been expanded and detailed identification keys for identifying Syrphidae species distributed in Europe. However, more than these identification keys are needed for the species distributed in Turkey (Speight, 2012, 2018, 2020).

This study aims to develop a habitus images-based deep learning model for faster and more accurate identification of some ecologically important adult Syrphidae (Diptera) species. In the literature, there are artificial intelligence-based studies for classifying and detecting plant and insect species. However, these studies generally focus on identifying some specific species. The datasets used are generally publicly accessible. Unlike the studies in the literature, an original dataset was used in this study. In the dataset used, there are images of 15 syrphid species. The detailed study, which will be carried out in the future, covering the syrphid species distributed in Turkey and neighbouring countries, will be preliminary.

DATASET AND MATERIAL

In this study, different images were used for the dipteran family Syrphidae both in the laboratory and natural environments. To this end, the photographs of some identified syrphid species deposited in Metin Aktaş Zoology Museum of Gazi University (MAZMGU, Ankara, Turkey) were taken for syrphid images stored in the museum environment. Furthermore, the photographs of these syrphid species photographed in their natural environment were taken from the ZMGU database. The number of photographs of the specimens with a sting in the museum and those taken from the database was insufficient for the training set. Therefore, more images of these species were collected from the Internet. The authors collected images using the Google search engine. Care was taken to use the photographs of the species mentioned in Table 1 from different angles. Seven hundred twenty-seven hoverfly images were used to train and test the deep learning model developed in our study. There are a total of 15 syrphid species in the dataset. Figure 1 provides specimen photographs for some species of hoverflies in the dataset.

Figure 1. Some photos of the hoverfly species used to train or test the convolutional neural network (from left to right: *Eristalis tenax, Chrysotoxum vernale, C. bicinctum*).

Four hundred seventy-nine images from the dataset were used for training, and two hundred forty-eight images were used for testing. Table 1 gives the names of each species used in the dataset and the number of images used in the training and test sets. 2/3 of the images in the dataset were allocated to the training set and the remaining 1/3 to the test set.

Table 1. Species names in the dataset and the number of images in the training and test set.

Species Names	Number of images in the training set	Number of images in the test set
Anasimyia contracta Claussen & Torp, 1980	37	18
Anasimyia interpuncta (Harris, 1776)	31	15
Anasimyia lineata (Fabricius, 1787)	29	14
Baccha elongata (Fabricius, 1775)	31	16
Callicera aurata (Rossi, 1790)	33	17
Ceriana conopsoides (Linnaeus, 1758)	22	11
Chrysotoxum arcuatum (Linnaeus, 1758)	33	17
Chrysotoxum bicinctum (Linnaeus, 1758)	31	15
Chrysotoxum vernale Loew, 1841	32	16
Dasysyrphus albostriatus (Fallen, 1817)	32	16
Epistrophe eligans (Harris, 1780)	31	16
Episyrphus balteatus (De Geer, 1776)	37	19
Eristalis interrupta (Poda, 1761)	34	17
Eristalis similis (Fallen, 1817)	33	16
Eristalis tenax (Linnaeus, 1758)	39	20

The total number of samples is 728. 485 samples were used for training and 243 for testing. Samples from the dataset were randomly selected. It was aimed to prevent over-learning by using 10-fold cross-validation. Cross-validation is a technique used to assess the performance of a machine-learning model. In this method, the dataset is divided into 10 equal-sized subsets. The model is then evaluated 10 times, where each time, one of the subsets is used as the test set, and the remaining 9 subsets are used as the training set. This process is repeated for each subset so that each subset can be the test set once. The results from the 10 evaluations are typically averaged to provide an overall performance metric for the model. This approach helps obtain a more robust estimate of the model's performance by reducing the impact of the specific data points in a single train-test split.

DESIGNED DEEP LEARNING MODEL

In this study, a deep-learning model was developed for hoverfly classification. After the developed deep learning model was trained, it was aimed to predict the species of an adult syrphid image entered from outside. To this end, different deep learning architectures were tried, and prediction rates were compared. In this study, a CNN-based deep learning model was developed. The detailed architecture of the developed model is shown in Figure 2.

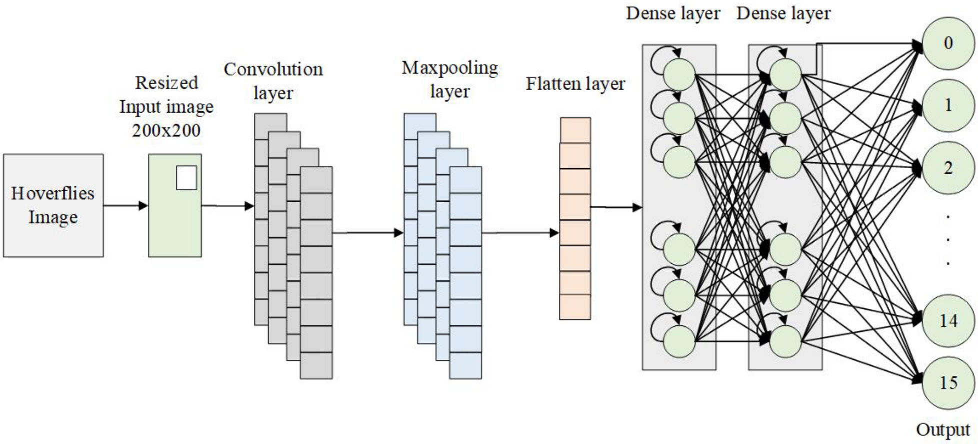

Figure 2. Designed deep learning model.

In the developed model, feature extraction was done with convolution and pooling layers, and classification values were calculated with a fully connected artificial neural network. The probability distribution of the classification categories was determined with the last softmax layer.

X nodes were created in the input layer of the developed model. One input node was used for each pixel value. Since the images have three channels as Red, Green, Blue (RGB), three input layers were used for each image. The developed deep-learning model has 15 nodes in the output layer. An output node was used for each species in the dataset. During the training, the one-hot-encoding method took the node output as 1, representing the species of the entered image, and the others as 0.

In the developed model, values of 0 for padding and 1 for stride were taken in the convolution layers. 3x3 filters were used in the convolution layer. The kernel matrix used as a filter is shown in Figure 3.

-1	-1	1
0	1	-1
0	1	1

Figure 3. The kernel used in the convolution layer.

The same kernel was used for all three channels. In the pooling layer, the window size was taken as 3x3, and the stride value was selected as one. The rectifier linear unit (ReLU) was used as the activation function in the nodes.

EXPERIMENTAL RESULTS

For experimental results, an application was developed in the Google Colabs environment using Python programming language and TensorFlow and Keras deep learning libraries. For the software interface, ReactJS was used on the frontend and a web service created with Python on the backend.

The training of the developed deep learning model was completed after 50 epochs. The categorical cross-entropy (CCE) function, which is widely used in multi-class classification problems, was used to calculate the loss values.

$$CCE = -\frac{1}{N} + \sum_{i=1}^{N}\sum_{c=1}^{C}(p_{ic}\log(y_{ic})) \qquad (1)$$

Here (1), N is the total number of observations, C is the total number of species, p_{ic} is the target (correct result) for the i^{th} observation of species c, and y_{ic} is the predicted probability distribution. The change in the loss value during the training of the model is shown in Figure 4.

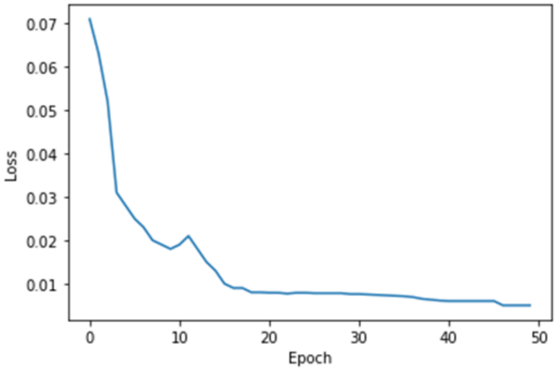

Figure 4. Change of loss values according to epoch.

As seen in Figure 4, the loss value decreases rapidly in the initial epochs, then continues to decrease slowly and reaches the lowest value at the 50th epoch, becoming stable and ending the training. The change in the loss value shows that the developed model was designed successfully. The change in the accuracy value during the training of the model is shown in Figure 5.

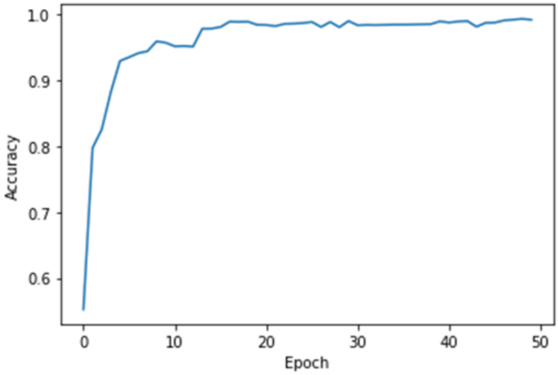

Figure 5. Accuracy value change.

As seen in Figure 5, the classification accuracy rate increases rapidly in the initial epochs, the rate of increase decreases after the 5th epoch, and the increase decreases considerably after the 20th epoch. At the 50th epoch, it reaches its highest value and becomes stable. The change in the accuracy value shows that the developed model is successful.

The true positive (TP), true negative (TN), false positive (FP) and false negative (FN) values for the experimental results obtained for 15 syrphid species are given in Table 2.

Table 2. TP, TN, FP, and FN values for species in the dataset.

Species names	TP	TN	FP	FN
Anasimyia contracta	19	227	2	-
Anasimyia interpuncta	15	232	-	1
Anasimyia lineata	14	233	-	1
Baccha elongata	14	231	1	2
Callicera aurata	17	231	-	-
Ceriana conopsoides	10	236	1	1
Chrysotoxum arcuatum	16	230	1	1
Chrysotoxum bicinctum	15	231	1	1
Chrysotoxum vernale	16	232	-	-
Dasysyrphus albostriatus	16	231	1	-
Epistrophe eligans	16	232	-	-
Episyrphus balteatus	19	229	-	-
Eristalis interrupta	17	231	-	-
Eristalis similis	16	231	-	1
Eristalis tenax	20	228	-	-

The confusion matrix obtained for 15 species is presented in Table 3.

Table 3. Confusion matrix.

		True Class														
		Anasimyia contracta	*Anasimyia interpuncta*	*Anasimyia lineata*	*Baccha elongata*	*Callicera aurata*	*Ceriana conopsoides*	*Chrysotoxum arcuatum*	*Chrysotoxum bicinctum*	*Chrysotoxum vernale*	*Dasysyrphus albostriatus*	*Epistrophe eligans*	*Episyrphus balteatus*	*Eristalis interrupta*	*Eristalis similis*	*Eristalis tenax*
Predicted Class	*Anasimyia contracta*	19	1	-	1	-	-	-	-	-	-	-	-	-	-	-
	Anasimyia interpuncta	-	15	-	-	-	-	-	-	-	-	-	-	-	-	-
	Anasimyia lineata	-	-	14	-	-	-	-	-	-	-	-	-	-	-	-
	Baccha elongata	-	-	-	14	-	1	-	-	-	-	-	-	-	-	-
	Callicera aurata	-	-	-	-	17	-	-	-	-	-	-	-	-	-	-
	Ceriana conopsoides	-	-	-	1	-	10	-	-	-	-	-	-	-	-	-
	Chrysotoxum arcuatum	-	-	1	-	-	-	16	1	-	-	-	-	-	-	-
	Chrysotoxum bicinctum	-	-	-	-	-	-	1	15	-	-	-	-	-	-	-
	Chrysotoxum vernale	-	-	-	-	-	-	-	-	16	-	-	-	-	-	-
	Dasysyrphus albostriatus	-	-	-	-	-	-	-	-	-	16	-	-	-	1	-
	Epistrophe eligans	-	-	-	-	-	-	-	-	-	-	16	-	-	-	-
	Episyrphus balteatus	-	-	-	-	-	-	-	-	-	-	-	19	-	-	-
	Eristalis interrupta	-	-	-	-	-	-	-	-	-	-	-	-	17	-	-
	Eristalis similis	-	-	-	-	-	-	-	-	-	-	-	-	-	16	-
	Eristalis tenax	-	-	-	-	-	-	-	-	-	-	-	-	-	-	20

The total number of TP = 240 and FP = 8. The accuracy value was obtained as 8/248 = 0.9677. The accuracy, recall, precision, and f-measure values for each category are shown in Table 4.

Table 4. Accuracy, precision, recall, and f-measure values for each species.

Class	Accuracy	Precision	Recall	F-Measure
Anasimyia contracta	0.99	0.90	1	0.94
Anasimyia interpuncta	0.99	1	0.93	0.96
Anasimyia lineata	0.99	1	0.93	0.96
Baccha elongata	0.98	0.93	0.87	0.89
Callicera aurata	1	1	1	1
Ceriana conopsoides	0.99	0.90	0.90	0.90
Chrysotoxum arcuatum	0.99	0.94	0.94	0.94
Chrysotoxum bicinctum	0.99	0.93	0.93	0.93
Chrysotoxum vernale	1	1	1	1
Dasysyrphus albostriatus	0.99	0.94	1	0.96
Epistrophe eligans	1	1	1	1
Episyrphus balteatus	1	1	1	1
Eristalis interrupta	1	1	1	1
Eristalis similis	0.99	1	0.94	0.96
Eristalis tenax	1	1	1	1

With the developed model, the accuracy, recall, precision, and f-measure values obtained for each species were very high. The lowest accuracy value was 0.98 in *Baccha elongata*. All others were calculated as 0.99 and 1. The lowest precision value was 0.90 in *Anasimyia contracta* and *Ceriana conopsoides* species. It was found to be 1 in 9 out of 15 species. The recall value was the lowest in *Baccha elongata* with 0.87. It was found to be 1 in 8 out of 15 species. The lowest f-measure value was 0.89 in *Baccha elongata*. It was found to be 1 in 6 out of 15 species. The graph of the obtained accuracy, precision, recall, and f-measure values by species is given in Figure 6.

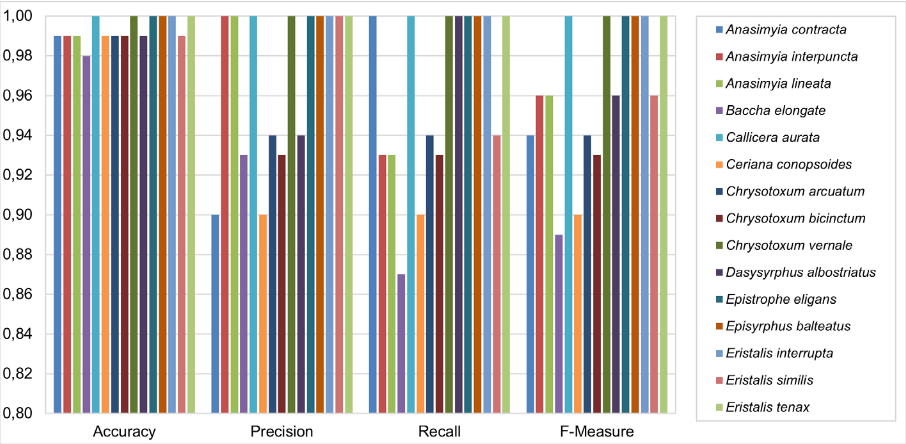

Figure 6. Accuracy, precision, recall, and f-measure values.

The graph of accuracy, precision, recall, and f-measure values for each species is given in Figure 7.

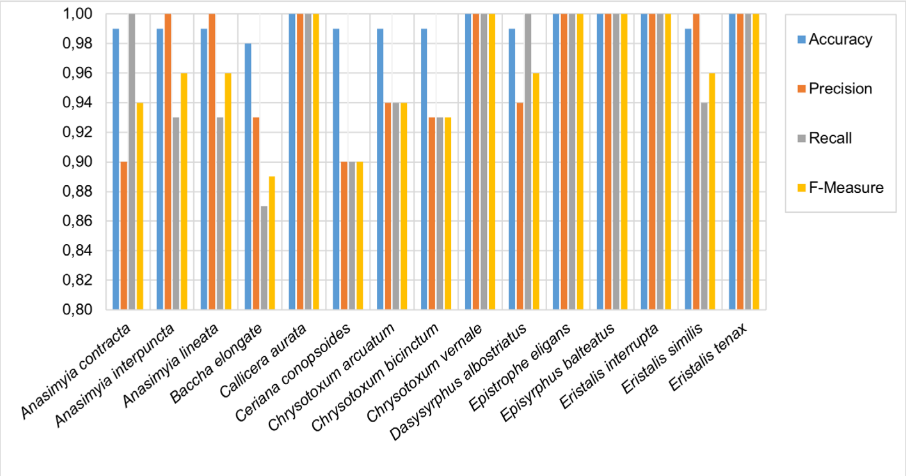

Figure 7. Accuracy, precision, recall, and f-measure values for each class.

The accuracy, precision, recall, and f-measure values obtained for all species are given in Table 5.

Table 5. Average values of accuracy, precision, recall, and f-measure for all species.

	Accuracy	Precision	Recall	F-measure
Average	0.96	0.97	0.96	0.96

As seen in Table 5, the model developed for classification has very high accuracy, precision, recall, and f-measure values for each species. The average values calculated for all classes of the developed deep learning-based classifier were also relatively high. In classification problems, the recall value used in the measurement of the FN value, in other words, the entries that could not be assigned to the correct class targeted in the classification, was relatively high as 0.87 even in the *Baccha elongate* species, in which it was the lowest. The obtained experimental results show that the developed deep-learning model successfully classifies hoverflies.

CONCLUSIONS

In this study, there are 728 samples in the dataset used. 2/3 of these samples were used for training and 1/3 for testing. Samples from the dataset were randomly selected. It was aimed to prevent over-learning by using 10-fold cross-validation. Three input layers, Red, Green, and Blue, are used for each image. 3x3 filters are used in the convolution layer of the developed model. There are 15 nodes in the output layer of the model.

The experimental results showed that the CNN-based deep learning model successfully classified hoverflies. Fifteen species from the family Syrphidae are successfully classified using the developed image-based recognition system. As with other taxonomic groups, the Syrphidae usually has taxonomic characters showing subtle differences in genus and species distinction. Although the characters at the microscopic level are used in the distinction of Syrphidae species (eye bristles, antennal segment lengths, thorax bristles), the macroscopic level is also used in the characters (color of the legs (femura, tibiae, tarsomers), abdomen coloration and pattern) (Speight and Sarthou, 2012). The high success rate in the image-based recognition system developed in our study is that body coloration and patterning in the abdomen tergites are the dominant characters in the distinction at the species level. Although we could not find a consistent error for species in this study, this program can misidentify species that are morphologically very similar and have similar patterns and coloration because the microscopic character is used together with expert opinion to distinguish such species.

This study showed that CNN is a suitable application for distinguishing Syrphidae species. Moreover, image-based recognition systems can be developed to identify more Syrphidae species and species belonging to similar dipteran families.

DISCUSSIONS

This study developed a CNN-based model to classify Syrphidae species, which play an essential role in pollinating flowering plants and grains. Syrphidae is a family with high species diversity in the order Diptera and is used in biodiversity and conservation studies. This is the first study on the classification of Syrphidae species in the literature. There are studies on the classification of insect species in the literature. However, these studies generally focus on classifying certain types using public datasets. An original data set was used in this study.

Experimental results show that the developed model has a very successful classification performance. Increasing the sample size used in the study and the image quality of the used samples will increase the model's success. The quality of the images used in this study is high. Expanding the dataset with new images will contribute to the model's training process. In addition, a more successful classification performance can be achieved with hybrid deep learning models to be developed.

ACKNOWLEDGMENTS

We thank Prof. Dr. Abdullah HASBENLI (Ankara, TURKEY) for allowing the use of syrphid photographs from the database of the MAZMGU and the taking of photographs of syrphid specimens.

REFERENCES

Al-Saqer, S. M., & Hassan G. M. (2011). Artificial neural networks based red palm weevil (*Rynchophorus ferrugineous* Olivier) recognition system. *American Journal of Agricultural and Biological Sciences*, 6, 356-364. doi: https://doi.org/10.3844/ajabssp.2011.356.364.

Arbuckle, T., Schröder, S., Steinhage, V., & Wittmann, D. (2001). Biodiversity informatics in action: identication and monitoring of bee species using ABIS. *15th International Conference on Environmental Informatics*, Zurich, October 10-12.

Beccaloni, G. W., & Gaston, K. J. (1995). Predicting the species richness of Neotropical forest butterflies: Ithominae (Lepidoptera: Nymphalidae) as indicators. *Biological Conservation*, 71, 77–86. doi: https://doi.org/10.1016/0006-3207(94)00023-J.

Dennis, R. L. H., Dapporto, L., Shreeve, T. G., John, E., Coutsis, J. G., Kudrna, O., Saarinen, K., & Ryrholm, N. (2008). Butterflies of European islands: The implications of the geography and ecology of rarity and endemicity for conservation. *Journal of Insect Conservation*, 12, 205–236. doi: https://doi.org/10.1007/s10841-008-9148-3.

Dietrich, C. H., Emigh, T. H., & Deitz, L. L. (1991). Morphometric discrimination among females of sibling species of Aconophorini (Homoptera: Membracidae). *Systematic Entomology*, 16(3), 311-318. doi: https://doi.org/ 10.1111/J.1365-3113.1991.Tb00691.X.

Dietrich, C. H. & Pooley, C. D. (1994). Automated identification of leafhoppers (Homoptera: Cicadellidae: Draeculacephala Ball). *Annals of the Entomological Society of America*, 87(4), 412-423. doi: https://doi.org/10.1093/Aesa/87.4.412.

Do, M. J., Harp, K., & Norris, A. (1999). Test of a pattern recognition system for identification of spiders. *Bulletin of Entomological Research*, 89(3), 217-224. doi: https://doi.org/10.1017/S0007485399000334.

Dover, J., Sparks, T., Clarke, S., Gobbett, K., & Glossop, S. (2000). Linear features and butterflies: the importance of green lanes. *Agriculture, Ecosystems & Environment*, 80, 227–242. doi: https://doi.org/10.1016/S0167-8809(00)00149-3.

Faria, F. A., Perreb, P., Zucchic, R. A., Jorged, L. R., Lewinsohnd, T. M., Rochaa, A., & Torresa, R. da S. (2014). Automatic identification of fruit flies (Diptera: Tephritidae). *Journal of Visual Communication and Image Representation*, 25(7), 1516-1527. doi: https://doi.org/10.1016/j.jvcir.2014.06.014.

Fedor, P., Vaňhara, J., Havel, J., Malenovský, I., & Spellerberg, I. (2009). *Artificial intelligence in pest insect monitoring*. *Systematic Entomology*, 34, 398–400. doi: https://doi.org/10.1111/j.1365-3113.2008.00461.x.

Francoy, T. M., Wittmann, D., Drauschke, M., Müller, S., Steinhage, V., Bezerra-Laure, M. A. F., Jong, D., & Gonçalves, L. S. (2008).Identifiation of Africanized honey bees through wing morphometrics: two fast and efficient procedures. *Apidologie*, 39(5), 488-494. doi: https://doi.org/10.1051/apido:2008028.

Gaston, K. J. & May, R. M. (1992). Taxonomy of taxonomists. *Nature*, 356, 281–281. doi: https://doi.org/10.1038/356281a0.

Gaston, K. J. & Mound, L. A. (1993). Taxonomy, hypothesis testing and the biodiversity crisis. *Proceedings of the Royal Society of London Series B*, 251, 139–142.

Gaston, K. J. & O'Neill, M. A. (2004). Automated species identication: why not?. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 359(1444), 655-667. doi: https://doi.org/10.1098/rstb.2003.1442.

Hassan, S. N. A., Rahman, N.N.S.A., Htike, Z. Z., & Win, S. I. (2014). Vision Based Entomology: A Survey. *International Journal of Computer Science & Engineering Survey*, 5(1), 19-31. doi: https://doi.org/10.5121/ijcses.2014.5103.

Karar, M.E., Alsunaydi, F., Albusaymi, S., & Alotaibi, S. (2021). A new mobile application of agricultural pests recognition using deep learning in cloud computing. *Alexandria Engineering Journal*, 60, 4423–4432. doi: https://doi.org/10.1016/j.aej.2021.03.009

Kasinathan, T., Singaraju, D., & Uyyala, S. R. (2020). Insect classification and detection in field crops using modern machine learning techniques. *Information Processing in Agriculture*, 8(3), 446-457. doi: https://doi.org/10.1016/j.inpa.2020.09.006.

Kaya, Y. & Kaycı, L. (2014). Application of artificial neural network for automatic detection of butterfly species using colour and texture features. *Vis Comput*, 30, 71–79. doi: https://doi.org/10.1007/s00371-013-0782-8.

Klecka, J., Hadrava, J., Biella, P., & Akter, A. (2018). Flower visitation by hoverflies (Diptera: Syrphidae) in a temperate plant-pollinator network. *PeerJ*, 6, e6025. doi: https://doi.org/10.7287/peerj.preprints.26516v2.

Kočiš Tubić, N., Ståhls G., Ačanski J., Đan M., Vidaković D. O., Hayat R., Khaghaninia S., Vujić, A., & Radenković S. (2018). An integrative approach in the assessment of species delimitation and structure of the *Merodon nanus* species group (Diptera: Syrphidae). *Organisms Diversity & Evolution*, 18(4), 479-497.

Larios, N., Deng, H., Zhang, W., Sarpola, M., Yuen, J., Paasch, R., Moldenke, A., Lytle, D. A., Correa, S. R., Mortensen, E. N., Shapiro, L. G., & Dietterich, T. G. (2008). Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects. *Machine Vision and Applications*, 19(2), 105-123. doi: https://doi.org/ 10.1007/s00138-007-0086-y.

Leow, L. K., Chew, L. L., Chong, V. C., & Dhillon, S. K. (2015). Automated identification of copepods using digital image processing and artificial neural network. *BMC Bioinformatics*, 16(18), 1-12. doi: https://doi.org/10.1186/1471-2105-16-S18-S4.

Li, W., Zheng, T., Yang, Z., Li, M., Sun, C., & Yang, X. (2021). Classification and detection of insects from field images using deep learning for smart pest management: A systematic review. *Ecological Informatics*, 66, 101460.

Likov, L., Vujić A., Kočiš Tubić N., Đan M., Veličković N., Rojo R., Pérez-Bañón C., Veselić S., Barkalov A., Hayat R., & Radenković S. (2020). Systematic position and composition of *Merodon nigritarsis* and *M. avidus* groups (Diptera, Syrphidae) with a description of four new hoverflies species. *Contributions to Zoology*, 89, 74-125.

Martineau, M., Conte, D., Raveaux, R., Arnault, I., Munier, D., & Venturini, G. (2017). A survey on image-based insect classification. *Pattern Recognition*, 65, 273–284. doi: https://doi.org/10.1016/j.patcog.2016.12.020.

Mayo, M. & Watson, A. T. (2007). Automatic species identification of live moths. *Knowledge-Based Systems*, 20(2), 195-202. doi: https://doi.org/10.1016/j.knosys.2006.11.012.

O'Neill, M. A. (2008). *DAISY: a practical tool for semi-automated species identication." In Automated Taxon Identication in Systematics: Theory, Approaches and Applications*. N. MacLeod (ed.), London, CRC Press.

Peck, V. (1988). *Family: Syrphidae.* In: A. SOOS and L. Papp. (Eds), Catalogue of Palaearctic Diptera. Syrphidae-Conopidae, Budapest.

Rotheray, G. E. & Gilbert, F. (2011). *The natural history of hoverflies*. Forrest Text.

Qing, Y. A. O., Jun, L. V., Liu, Q.-J., Diao, G.-Q., Yang, B.-J., Chen, H.-M., & Tang, J. (2012). An insect imaging system to automate rice light-trap pest identification. *Journal of Integrative Agriculture*, 11(6), 978-985. doi: https://doi.org/10.1016/S2095-3119(12)60089-6.

Silva, F. L. d., Grassi Sella, M. L., Francoy, T. M., & Costa, A. H. R. (2015). Evaluating classification and feature selection techniques for honeybee subspecies identification using wing images. *Computers and Electronics in Agriculture*, 114, 68-77. doi: https://doi.org/10.1016/j.compag.2015.03.012.

Sommaggio, D. (1999). Syrphidae: can they be used as environmental bioindicators?. *Agriculture, Ecosystems and Environment*, 74, 343-356. doi: https://doi.org/10.1016/S0167-8809(99)00042-0.

Sommaggio, D. & Burgio, G. (2014). The use of Syrphidae as functional bioindicator to compare vineyards with different managements. *Bulletin of Insectology*, 67(1), 147-156.

Speight, M.C.D. (2018). StN key for the identification of the genera of European Syrphidae (Diptera) 2018. *Syrph the Net, the database of European Syrphidae*, 101, 1- 45.

Speight, M.C.D. (2020). StN key for the identification of the genera of European Syrphidae (Diptera) 2020. *Syrph the Net*, 105, 1-46.

Speight, M.C.D. & Sarthou, J. P. (2012). StN keys for the identification of adult European Syrphidae (Diptera) 2012. *Syrph the Net*, 70, 1-130.

Thenmozhi, K. & Reddy, U. S. (2019). Crop pest classification based on deep convolutional neural network and transfer learning. *Computers and Electronics in Agriculture*, 164, 104906.

Tiwari, V., Patel, H., Muttreja, R., Goyal, M., Ojha, M., Gupta, S., & Jain, S. (2021). Real-time soybean crop insect classification using customized deep learning models. *Data Management, Analytics and Innovation: Proceedings of ICDMAI 2021*, 1, 143-156, Springer Singapore.

Tofiski, A. (2004). DrawWing, a program for numerical description of insect wings. *Journal of Insect Science*, 4(1). doi: https://doi.org/10.1093/jis/4.1.17.

Van Driesche R., Hoddle M., & Enter T. (2008). Control of Pests and Weeds by Natural Enemies. An IntroDuction to Biological Control. Blackwell, Malden, MA & Oxford.

Vujić, A., Likov L., Radenković S., Kočiš Tubić N., Đan M., Šebić A., Pérez-Bañón C., Barkalov A., Hayat R., Rojo S., Andrić A., & Ståhls G. (2020). Revision of *Merodon serrulatus* group (Diptera, Syrphidae). *ZooKeys*, 909, 79-158.

Vujić, A., Radenković S., Ačanski J., Grković A., Taylor M., Şenol S. G. & Hayat R. (2015). Revision of the species of the *Merodon nanus* group (Diptera: Syrphidae) including three new species. *Zootaxa*, 4006(3), 439-462.

Vujić, A., Nedeljković Z., Hayat R., Demirözer O., Mengual X., & Kazerani F. (2017). New data on the genus *Chrysotoxum* Meigen (Diptera: Syrphidae) from North-East Turkey, Armenia, Azerbaijan and Iran including descriptions of three new species. *Zoology in the Middle East*, 63(3), 250-268.

Wang, J., Lin, C., Ji, L., & Liang, A. (2012). A new automatic identification system of insect images at the order level. *Knowledge-Based Systems*, 33, 102-110. doi: https://doi.org/10.1016/j.knosys.2012.03.014.

Wang, J., Ji, L., Liang, A., & Decheng, Y. (2011). The identification of butterfly families using content-based image retrieval. *Biosystems Engineering*, 3, 24–32. doi: https://doi.org/10.1016/j.biosystemseng.2011.10.003.

Wen, C., Guyer, D. E., & Li, W. (2009). Local feature-based identification and classification for orchard insects. *Biosystems Engineering*, 104(3), 299-307. doi: https://doi.org/10.1016/j.biosystemseng.2009.07.002.

Wen, C., Wu, D., Hu, H., & Pan, W. (2015). Pose estimation-dependent identification method for field moth images using deep learning architecture. *Biosystems Engineering*, 136, 117-128. doi: https://doi.org/10.1016/j.biosystemseng.2015.06.002.

Wen, C. & Zhu, Q. (2010). Dimension reduction analysis in image-based species classification. *IEEE International Conference on Intelligent Computing and Intelligent Systems (ICIS)*, China, October 29-31.

Xia, D., Chen, P., Wang, B., Zhang, J., & Xie, C. (2018). Insect Detection and Classification Based on an Improved Convolutional Neural Network. *Sensors (Basel)*, 18(12), 4169. doi: https://doi.org/10.3390/s18124169.

Yaakob, S. N. & Jain, L. (2012). An insect classification analysis based on shape features using quality threshold ARTMAP and moment invariant. *Applied Intelligence*, 37(1), 12-30. doi: https://doi.org/10.1007/s10489-011-0310-3.

Yang, H.P., Ma, C.S., Wen, H., Zhan, Q. B., & Wang, X. L. (2015). A tool for developing an automatic insect identification system based on wing outlines. *Scientific Reports*, 5, 1-11. doi: https://doi.org/10.1038/srep12786.

Zamora, J., Verdu, J.R., & Galante, E. (2007). Species richness in Mediterranean agroecosystems: Spatial and temporal analysis for biodiversity conservation. *Biological Conservation*, 134, 113–121. doi: https://doi.org/10.1016/j.biocon.2006.08.011.

Zayas, I. Y. & Flinn, P. W. (1998). Detection of Insects in Bulk Wheat Samples with Machine Vision. *Transactions of the ASAE* 41(3), 883-888. doi: https://doi.org/10.13031/2013.17206.

Zhu, L. Q., Ma, M. Y., Zhang, Z., Zhang, P. Y., Wu, W., Wang, D. D., & Wang, H. Y. (2017). Hybrid deep learning for automated lepidopteran insect image classification. *Oriental Insects*, 51(2), 79-91.

Zupan, S., Bužan, E., Grubar, V. B., & Jugovic, J. 2020. Importance of traditional landscapes in Slovenia for conservation of endangered butterfly. *Open Geosciences*, 12, 610–625. doi: https://doi.org/10.1515/geo-2020-0179.