UDC 004.8
DOI https://doi.org/10.26661/2786-6254-2025-1-13

OPTIMIZING OBJECT DETECTION PERFORMANCE THROUGH IMAGE ENHANCEMENT FOR BEE IMAGES

Zhukov O. O.
Assistant at the Department of Electronics, Information Systems and Software
Zaporizhzhia National University
Universytetska str., 66, Zaporizhzhia, Ukraine
orcid.org/0009-0004-3683-2527
oleksanrei@gmail.com

Horbenko V. I.
Candidate of Physical and Mathematical Sciences,
Associate Professor at the Department of Software Engineering
Zaporizhzhia National University
Universytetska str., 66, Zaporizhzhia, Ukraine
orchid.org/0000-0002-3342-4841
vgorbenko@ukr.net

Key words: *neural networks, clahe, ssd, tensorflow, bee hive, histogram equalization.*

Neural networks and deep learning, being among the technologies experiencing daily growing interest, have found practical application, particularly in bee monitoring. The use of neural networks for identifying bees in images can provide non-invasive monitoring of the condition of individual bees and, consequently, reduce stress on the colony. However, the accuracy of object detection algorithms can be significantly affected by variations in monitoring conditions, such as weather conditions, lighting, or the use of different cameras. Thus, image enhancement can be an important pre-processing step for optimizing object detection algorithms. This study utilizes various image enhancement and color normalization methods to improve the performance of the SSD MobileNet model in detecting the full bodies of bees in images. The data source consists of photos taken with a smartphone, which were subsequently divided into sections to fit within the model's acceptable identification limits. The study examines the effectiveness of three image enhancement methods: Contrast Limited Adaptive Histogram Equalization (CLAHE), a regular histogram equalization algorithm, and Protobuf pipeline functions. Each approach is evaluated on different image sets, including photos of normal beehive sections taken with a smartphone, photos with randomly altered brightness and contrast, and photos with slight blurring. The findings show that image normalization alone does not improve object detection performance on standard bee photos, whereas the CLAHE algorithm proved most effective in other cases, maintaining good results when image brightness and contrast were not optimized. Additionally, it was found that applying blur did not result in an overall increase in the efficiency of identifying bees in photos.

ОПТИМІЗАЦІЯ ПРОДУКТИВНОСТІ ВИЯВЛЕННЯ ОБ'ЄКТІВ ЗА ДОПОМОГОЮ ПОКРАЩЕННЯ ЗОБРАЖЕННЯ ДЛЯ ФОТО БДЖІЛ

Жуков О. О.
асистент кафедри електроніки, інформаційних систем
та програмного забезпечення
Запорізький національний університет
вул. Університетська, 66, Запоріжжя, Україна
orcid.org/0009-0004-3683-2527
oleksanrei@gmail.com

Горбенко В. І.
кандидат фізико-математичних наук,
доцент кафедри програмної інженерії
Запорізький національний університет
вул. Університетська, 66, Запоріжжя, Україна
orchid.org/0000-0002-3342-4841
vgorbenko@ukr.net

Ключові слова: *нейронні мережі, clahe, ssd, tensorflow, бджолиний вулик, вирівнювання гістограми.*

Нейронні мережі та глибоке навчання як технології, інтерес до яких з кожним днем зростає, знайшли практичне застосування, зокрема, у моніторингу бджіл. Використання нейронних мереж для ідентифікації бджіл на зображенні можуть забезпечити неінвазивний моніторинг стану окремих бджіл і, як наслідок, зменшити стрес на колонію. Проте на точність алгоритмів виявлення об'єктів можуть значно вплинути варіації умов моніторингу, як-от погодні умови, освітлення або використання різних камер. Таким чином, покращення зображення може бути важливим етапом попередньої обробки для оптимізації алгоритмів виявлення об'єктів. У цьому дослідженні використовуються різноманітні методи покращення зображення та нормалізації кольору, щоб покращити продуктивність моделі SSD MobileNet у виявленні повних тіл бджіл на зображеннях. Як джерела даних виступають фотографії зроблені на смартфон, які далі були розділені на секції з метою вкладання в допустимі ліміти ідентифікації моделі. У дослідженні розглядається ефективність трьох методів покращення зображення: CLAHE, регулярного алгоритму вирівнювання гістограми та функцій конвеєра Protobuf. Кожен підхід оцінюється на різних наборах зображень, включно з фотографіями звичайних секцій бджолиного вулика, знятих на смартфон, фотографіями з довільно зміненою яскравістю та контрастністю та фотографіями з легким розмиттям. Висновки показують, що нормалізація зображення сама по собі не покращує продуктивність виявлення об'єктів на звичайних фотографіях бджіл, тоді як алгоритм CLAHE виявився найефективнішим в інших випадках, зберігаючи хороші результати у випадках, коли яскравість і контраст зображення не були оптимізовані. Також було виявлено, що застосування розмиття не дало загального приросту в ефективності ідентифікації бджіл на фото.

Introduction. Bee monitoring has emerged as a significant area of interest due to the paramount importance of bees in the ecosystems they inhabit and their integral role in global food production. As quintessential pollinators, bees play a pivotal role in the sustenance of biodiversity and the stability of agricultural systems upon which we depend for a substantial portion of our food production [1; 2]. Alarmingly, the phenomena of Colony Collapse Disorder (CCD) and various health afflictions such as varroa mites have placed bee populations at substantial risk [3; 4].

Given the impracticality of continuous manual hive inspection, which may also introduce stress to the bees, automated, non-intrusive monitoring methods are important. One of the ways to monitor a bee hive is the usage of photo-video feed from the hive and using neural networks to process it and identify bees. Successful detection of bees in images can lead to further identification of the hive's state, providing valuable insights about bee quantity and traffic, varroa mite detection, and information about pollen that bees carry. However, the precision of object detection algorithms can be significantly impacted by variations in monitoring conditions, such as different cameras or lighting conditions [5]. Thus, image enhancement could be an important pre-processing step to optimize object detection algorithms. This paper aims to address this concern by exploring methodologies to enhance bee images, therefore increasing the quality of the end result of the object detection task for different bee images.

Literature Review. The application of neural networks for beehive monitoring has recently gained traction in the domain of bee studies. For instance, this paper [6] showcased how Object Recognition and real-time feed from webcams can be used for tracking bee traffic. Other research showed that the usage of neural networks can detect varroa mite on bees with an accuracy of 70% [7]. On the other hand, this research [8] shows the effectiveness of applying image filters to increase the Image Classification performance on the task of detecting pollen on a bee.

Optimization Adaptive Histogram Equalization (AHE) stands out as a noteworthy image enhancement technique for object detection. Studies like this paper [9] show that using Adaptive Histogram Equalization with an image sharpening algorithm can improve the performance of the Object Detection task in low-light images.

A variant of AHE, Contrast Limited Adaptive Histogram Equalization (CLAHE) has gained popularity due to its controlled approach to histogram equalization. This study [10] successfully utilized CLAHE to improve object detection results on various images, especially ones with shadows or low-contrast areas.

Several other studies managed to successfully use CLAHE in their neural networks pipeline. In a study which set to determine the effects of data augmentation on the CNN-based identification of bee infection, authors managed to successfully use CLAHE to enhance the contrast on foggy bee images, which made them more understandable and led to the better CNN classification results [11]. This other study on honeycomb detection and classification using deep learning has also showed positive results when using CLAHE, as it allowed honeycomb edges to be more distinctive [12].

Materials and Methods. As a data source for the Dataset, photos taken on a Samsung A20 were used. Since model-of-choice by default supports up to 100 detections and some of the photos of the size of 4128 x 3096 pixels contained up to 200 individual bees, it was decided to split images into 4 separate parts for their further use. These parts were annotated manually using the Remo software [13]. During the annotation process, only individual bee bodies were selected, and zones containing only a head or a spine were ignored. After that, 20% of the pictures went to the test set, and the others were used in training and evaluation sets. Figure 1 offers examples of cropped zones from the original dataset images that have been used for training.

(a) (b) (c) (d)

Fig. 1. Cropped annotated bee images that were used for training

Based on the initial evaluation set, three other ones were created: one was created with "corrupted" images, where brightness and contrast were randomly changed in a range of +–20% and +–0.5 accordingly, one with a slight Gaussian blur, and one with the blur applied to the corrupted set.

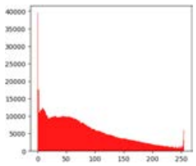
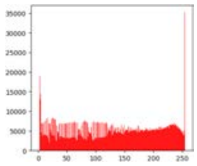
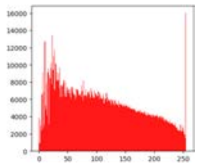
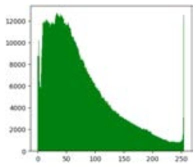
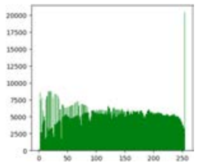
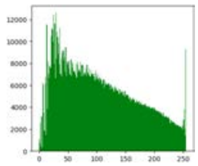
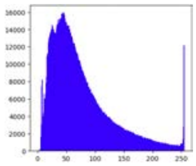
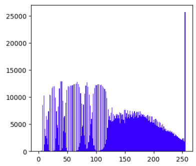
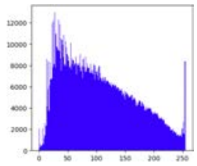
Blurred datasets were taken into account since it is a common way to reduce camera noise on images, which may appear in low-light photos and impact the overall performance of the model. Since the addition of too much blur to the image could make bee detection harder even with a bare eye, the blur was applied using OpenCV's instruments with (3.3) parameters.

For the training, two image enhancement techniques were used: the regular histogram equalization algorithm and CLAHE. Histogram Equalization (HE) is a method to increase the contrast by making the histogram of the image uniform. HE operates on the entire image, thereby globally enhancing contrast which can be advantageous for images with backgrounds and foregrounds that are both dark or both light. Although this method is simple and effective, it can cause color distortion or noise due to changes in the average brightness of the image or excessive contrast increase [14; 15].

Diverging from the global nature of traditional histogram equalization, Contrast Limited Adaptive Histogram Equalization (CLAHE) operates by dividing the image into smaller tiles and then individually equalizing each tile's histogram [16]. By focusing on smaller regions, CLAHE can better address variations in brightness and enhance local contrast, making it particularly suitable for images with varying backgrounds [17]. The performance of CLAHE heavily

Table 1

Applied image enhancement algorithms

	-	HE	CLAHE
			
Levels of red color			
Levels of green color			
Levels of blue color			

depends on its parameters, including the block size and the contrast limit, necessitating fine-tuning for optimal results. In this work, CLAHE was applied using OpenCV's instruments, with (8,8) parameters since based on visual inspection, it gave the optimal results between improved contrast and distortion.

An example of image enhancement application is shown in Table 1. On sample image, both algorithms had spread color levels more evenly, while both CLAHE and Histogram Equalization algorithm made the image brighter, CLAHE did a better job in preserving details.

The SSD MobileNet V2 model was chosen from the TensorFlow 2 Model Zoo repository. Among the other models in the repository, this model stood out for its fast time per detection of 19 milliseconds and its expected input resolution of 320 x 320 pixels, which is good for detecting small objects such as bees in an image. In addition to that, all labeled bounding boxes from the original dataset are in the range of 320 x 320 pixels resolution. Also, usage of SSD model architecture has been proven to yield positive results in terms of identifying bees on image [18].

This model accommodates two architectures – SSD and MobileNet.

The main idea behind SSD architecture is that the network uses only one pass through an image to identify the objects and classify them. Figure 2 shows the structure of the Single-Shot Detector architecture.

The key feature of this architecture is employing multiscale of convolutional frame exits which are connected to a set of feature arrays at the upper part of the neural network. Such a way allows to model potential frame spots efficiently [19].

MobileNet V2 is a deep learning network designed to function on mobile devices and to provide high accuracy of image classification with the least possible model size and number of operations. The main idea of MobileNet V2 lies in designing a light and fast neural network architecture with the help of two key technologies, which are residual connections and inverted residuals. Residual connections are a type

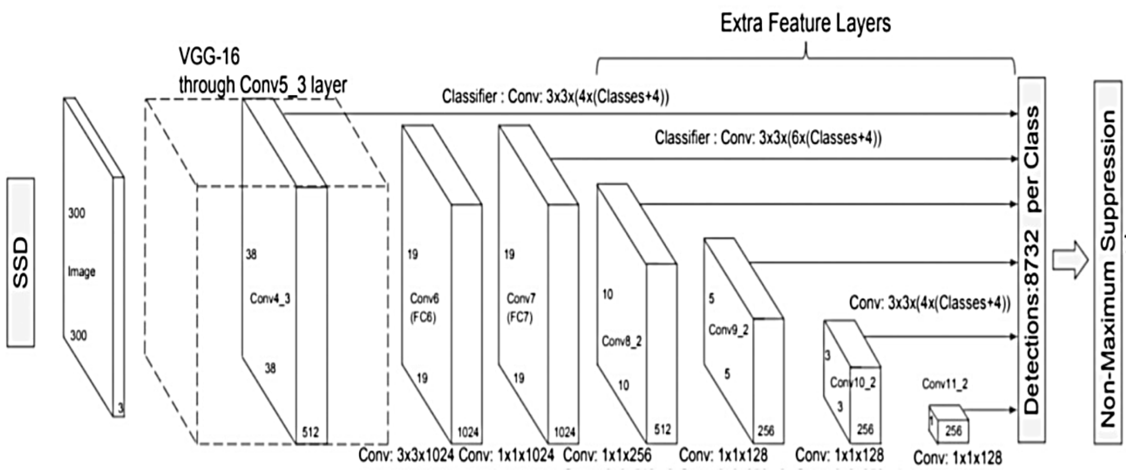


Fig. 2. Single-Shot Detection model architecture

of technology that allows transmitting information immediately from previous layers to the next ones which helps avoid the problem of studying shallow functions and provides a more efficient training of the neural network. Inverted residuals are a technology that allows decreasing the numbers of network parameters by reducing the size of layers and using their increase by applying 1x1 convolutions and linear layers [20].

To evaluate the performance of the applied object detection model, this paper used images that were not included in the original training and test datasets. In object detection approaches, the detection result and the classifier performance are two main indexes to evaluate the performance of the models. For each evaluation, following metrics were measured: Intersection over Union (IOU), recall, precision, and F1 score.

Precision measures what percentage of correctly classified labels is truly positive and uses True Positive (TP) and False Negative (FN) values to calculate it.

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

High precision indicates that the model is reliable in its positive predictions, making few false positive errors. However, precision alone does not account for all relevant aspects of a model's performance, particularly its ability to identify all positive cases in the dataset.

Recall, also known as sensitivity, is a metric that measures what proportion of True Positives out of all objects of the positive class the neural algorithm found

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

A high recall indicates that the model is effective in capturing a large proportion of positive cases, but it does not indicate how many negative cases were incorrectly classified as positive.

Intersection over Union (IoU) metric is used to evaluate the performance of object detection by comparing the ground truth bounding box to the predicted bounding box. The general mean IoU metric was calculated for each evaluation set by calculating it for each individual image, and divided by the evaluation images count.

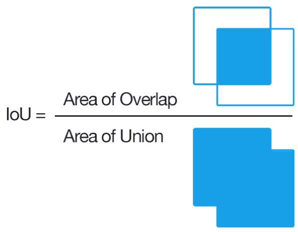


Fig. 3. Graphical representation of Intersection Over Union (IOU) (Source: Adrian Rose- brock/ Creative Commons)

The F1 score represents the harmonic mean of precision and recall, thereby integrating both metrics into a single figure. A higher F1 score illustrates that the model is more robust.

$$F1 = 2 \times \frac{\mathrm{Precision} \times Recall}{\mathrm{Precision} + Recall}; \quad (3)$$

The training and evaluation process unfolds a sequence of structured steps. Initially, input images are pre-processed, which includes applying image enhancement techniques as required to prepare the input data. Similarly, evaluation data undergo a pre-processing phase that may also encompass image enhancement to ensure consistency with the training data conditions. Following these preparatory steps, the dataset is transformed into two TFRecord files, which serve as the foundation for the training and validation processes. After it, the model is trained using Tensorflow Object Detection API instruments with the prepared training and testing input data.

Each model was trained for 22000 epochs with 112 batch size parameter. In addition to the HE-based and CLAHE-based models, a model which utilized Protobuf's pipeline features was added. For this pipeline, following data augmentation methods were used:

- Hue adjustment, which randomly alters hue by a value of up to 0.02.
- Saturation adjustment, which randomly changes saturation by a value between 0.8 and 1.25.
- Contrast adjustment, which randomly scales contrast by a value between 0.8 and 1.25.
- Brightness adjustment, which randomly changes image brightness by a value between 0.8 and 1.2.

These were applied to a random images during the model training to the input data, and no image enhancement algorithm was applied to them beforehand.

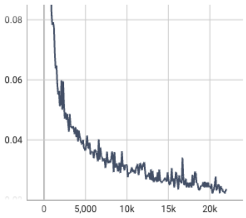
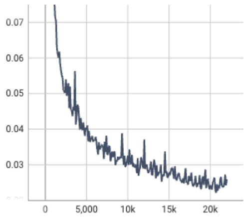
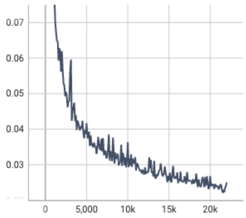
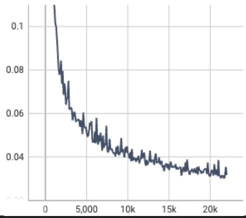
When training was finished, each model was evaluated on the according set of images and the mean IoU metric was calculated for them. Additionally, a manual assessment is conducted to verify the True Positives and False Negatives and to gather other relevant metrics. Only detections with a confidence score over 50% were taken into account during the evaluation process.

All models were trained using a Google Colab Pro high-RAM instance with a GPU accelerator, following the recommended training procedures. Table 2 shows the loss value and the progression of it during training of the models.

For model training, the cosine decay learning rate strategy was used, providing an adaptive adjustment of the learning rate over time. This approach follows a cosine curve to decrease the learning rate from an initial high value to near zero, facilitating both rapid exploration of the solution space and fine-tuning of model parameters. Since cosine decay learning rate depends on the initial learning rate value and the

Table 2

Loss values and their progression during training

Dataset	Final Loss Value	Loss value graph
Original	0.023	
Norm	0.025	
CLAHE	0.025	
PROTOC	0.031	

number of steps, which was identical for all models, the learning rate value progression is exactly the same for all 4 trained models, as shown on the Figure 4.

Results and Discussion. Based on the results from Table 3, adding Image Enhancement techniques didn't increase the end recall and precision results. On the other hand, even though dataset with CLAHE algorithm applied didn't show the best result, its precision and recall are similar to the original dataset.

Upon further investigation, IoU metric didn't show meaningful results in this particular task because there were a lot of detected zones that overlapped with each other, and, additionally, it was affected by the false positive results which were also heavily overlapping with the expected bounding boxes in zones with lots of bees present. Adding blur to the images also didn't show an increase in results.

Based on the data presented in Table 3, in the case of the corrupt dataset scenario, CLAHE showed the best results in successfully detecting bees.

Table 5 shows the object detection result in one of the evaluation images with CLAHE algorithm applied.

Table 3

Regular dataset performance results

	Regular				**Regular + Blur**			
	Rec.	**Prec.**	**IoU**	**F1**	**Rec**	**Prec**	**IoU**	**F1**
Original	81.8%	90%	0.562	0.857	74.4%	90.0%	0.581	0.815
Norm	71%	85%	0.600	0.774	61.2%	89.0%	0.583	0.725
CLAHE	79.9%	87.0%	0.577	0.833	76.1%	90.0%	0.606	0.825
PROTOC	62.3%	86.0%	0.603	0.723	56.6%	87.0%	0.616	0.686

Table 4

"Corrupt" dataset performance results

	Corrupt				**Corrupt + Blur**			
	Rec.	**Prec.**	**IoU**	**F1**	**Rec**	**Prec**	**IoU**	**F1**
Original	70.6%	92.0%	0.537	0.799	63.6%	91.0%	0.591	0.749
Norm	60.2%	90.0%	0.535	0.721	54.0%	89.0%	0.522	0.672
CLAHE	80.5%	89.0%	0.557	0.845	70.8%	89.0%	0.590	0.789
PROTOC	62.9%	87.0%	0.576	0.730	55.7%	84.0%	0.621	0.670

 ISSN 2786-6254

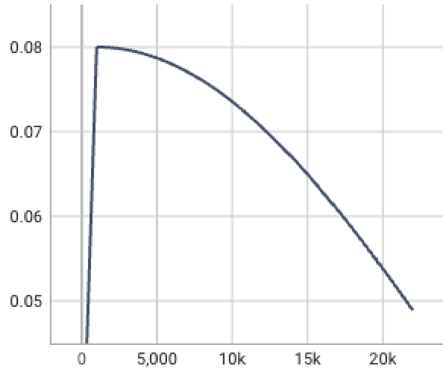


Fig. 4. Learning rate progression over training steps graph

Conclusions. The main goal of the current study was to determine whether using image enhancement algorithms can lead to better object detection results. The results of this investigation show that using HE and CLAHE algorithms have not increased the precision of the object detection model in general. Another major finding is that usage of CLAHE algorithm may lead to more stable results in various lighting conditions and small defects in image coloring. However, using HE, blur, and data augmentation features of Protobuf pipeline did not show significant results in either test.

Further research might explore the possibilities of using not only the image normalization techniques, but also applying color filters. An issue that was not explored in this study is the appearance of camera noise in low light environment, so reviewing the models' performance and ways to counter the noise can be a subject for further investigation.

Table 5

Object detection results

Dataset	Resulting Image
Original	
Original + Blur	
Corrupt	
Corrupt + Blur	

REFERENCES

1. Dicks, L.V., Breeze, T.D., Ngo, H.T., Senapathi, D., An, J., Aizen, M.A., Basu, P., Buchori, D., Galetto, L., Garibaldi, L.A., Gemmill-Herren, B., Howlett, B.G., Imperatriz-Fonseca, V.L., Johnson, S.D., Kovács-Hostyánszki, A., Kwon, Y.J., Lattorff, H.M.G., Lungharwo, T., Seymour, C.L., . . . Potts, S.G. (2021). A global-scale expert assessment of drivers and risks associated with pollinator decline. *Nature Ecology & Evolution*, 5 (10), 1453–1461. https://doi.org/10.1038/s41559-021-01534-9.
2. Potts, S.G., Biesmeijer, J.C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W.E. (2010). Global pollinator declines: trends, impacts and drivers. *Trends in Ecology & Evolution*, 25 (6), 345–353. https://doi.org/10.1016/j.tree.2010.01.007.
3. Neov, B., Shumkova, R., Palova, N., & Hristov, P. (2021). The health crisis in managed honey bees (Apis mellifera). Which factors are involved in this phenomenon? *Biologia*, 76 (8), 2173–2180. https://doi.org/10.1007/s11756-021-00684-2.
4. Rosenkranz, P., Aumeier, P., & Ziegelmann, B. (2010). Biology and control of Varroa destructor. *Journal of Invertebrate Pathology*, 103, S96–S119. https://doi.org/10.1016/j.jip.2009.07.016.
5. Ito, K., Higashi, H., Hietanen, A., Fält, P., Hine, K., Hauta-Kasari, M., & Nakauchi, S. (2022). The optimization of the light-source spectrum utilizing neural networks for detecting oral lesions. *Journal of Imaging*, 9 (1), 7. https://doi.org/10.3390/jimaging9010007.
6. Ryu, J.-S., Jung, J.-W., Jeong, C.-H., Choi, B.-J., Lee, M.-L., & Kwon, H.W. (2021). Honeybee in-out monitoring system by object recognition and tracking from real-time webcams. Han'gug Yangbong Haghoeji [Apicultural Society of Korea], 36 (4), 273–280. https://doi.org/10.17519/apiculture.2021.11.36.4.273.

7. Voudiotis, G., Moraiti, A., & Kontogiannis, S. (2022). Deep learning beehive monitoring system for early detection of the Varroa mite. *Signals*, 3 (3), 506–523. https://doi.org/10.3390/signals3030030.
8. Braga, D., Madureira, A., Scotti, F., Piuri, V., & Abraham, A. (2021). An intelligent monitoring system for assessing bee hive health. IEEE Access: Practical Innovations, Open Solutions, 9, 89009–89019. https://doi.org/10.1109/access.2021.3089538.
9. Guo, H., Lu, T., & Wu, Y. (2021). Dynamic low-light image enhancement for object detection via end-to-end training. 2020 25th International Conference on Pattern Recognition (ICPR).
10. Wang, T.-S., Kim, G. T., Kim, M., & Jang, J. (2023). Contrast enhancement-based preprocessing process to improve deep learning object task performance and results. *Applied Sciences* (Basel, Switzerland), 13 (19), 10760. https://doi.org/10.3390/app131910760.
11. Kaur, M., Ardekani, I. T., Sharifzadeh, H., & Varastehpour, S. (2022). A CNN-Based Identification of Honeybees' Infection using Augmentation. 2022 International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME). https://doi.org/10.1109/iceccme55909.2022.9988275.
12. Alves, T.S., Pinto, M.A., Ventura, P., Neves, C.J., Biron, D.G., Cândido, A., De Paula Filho, P.L., & Rodrigues, P. (2020). Automatic detection and classification of honey bee comb cells using deep learning. Computers and Electronics in Agriculture, 170, 105244. https://doi.org/10.1016/j.compag.2020.105244.
13. Remo Remo.ai. Available at: https://github.com/rediscovery-io/remo-python (Accessed: January 20, 2024).
14. Hum, Y.C., Lai, K.W., & Mohamad Salim, M.I. (2014). Multiobjectives bihistogram equalization for image contrast enhancement. *Complexity*, 20 (2),22–36. https://doi.org/10.1002/cplx.21499.
15. Chen Soong Der and Sidhu, M.S. (2009). Re-evaluation of Automatic Global Histogram Equalization-based Contrast. eJCSIT – *Electronic Journal of Computer Science & Information Technology*, 1, 13–17.
16. Zuiderveld, K. (1994). Contrast limited adaptive histogram equalization. In P.S. Heckbert (Ed.), Graphics Gems (pp. 474–485). Elsevier.
17. Reza, A.M. (2004). Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement. *The Journal of VLSI Signal Processing Systems for Signal Image and Video Technology*, 38 (1), 35–44. https://doi.org/10.1023/b:vlsi.0000028532.53893.82
18. Cai, J., Makita, Y., Zheng, Y., Takahashi, S., Hao, W., & Nakatoh, Y. (2022). Single shot multibox detector for honeybee detection. Computers & Electrical Engineering, 104, 108465. https://doi.org/10.1016/j.compeleceng.2022.108465.
19. Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A.C. (2015). SSD: Single Shot MultiBox Detector. https://doi.org/10.48550/ARXIV.1512.02325.
20. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). MobileNetV2: Inverted residuals and linear bottlenecks. https://doi.org/10.48550/ARXIV.1801.0438.