scientific reports

OPEN

Assessment of deep convolutional neural network models for species identification of forensically-important fly maggots based on images of posterior spiracles

Darlin Apasrawirote[1], Pharinya Boonchai[2], Paisarn Muneesawang[2], Wannacha Nakhonkam[3] & Nophawan Bunchu[3]✉

Forensic entomology is the branch of forensic science that is related to using arthropod specimens found in legal issues. Fly maggots are one of crucial pieces of evidence that can be used for estimating post-mortem intervals worldwide. However, the species-level identification of fly maggots is difficult, time consuming, and requires specialized taxonomic training. In this work, a novel method for the identification of different forensically-important fly species is proposed using convolutional neural networks (CNNs). The data used for the experiment were obtained from a digital camera connected to a compound microscope. We compared the performance of four widely used models that vary in complexity of architecture to evaluate tradeoffs in accuracy and speed for species classification including ResNet-101, Densenet161, Vgg19_bn, and AlexNet. In the validation step, all of the studied models provided 100% accuracy for identifying maggots of 4 species including *Chrysomya megacephala* (Diptera: Calliphoridae), *Chrysomya* (*Achoetandrus*) *rufifacies* (Diptera: Calliphoridae), *Lucilia cuprina* (Diptera: Calliphoridae), and *Musca domestica* (Diptera: Muscidae) based on images of posterior spiracles. However, AlexNet showed the fastest speed to process the identification model and presented a good balance between performance and speed. Therefore, the AlexNet model was selected for the testing step. The results of the confusion matrix of AlexNet showed that misclassification was found between *C. megacephala* and *C.* (*Achoetandrus*) *rufifacies* as well as between *C. megacephala* and *L. cuprina*. No misclassification was found for *M. domestica*. In addition, we created a web-application platform called thefly.ai to help users identify species of fly maggots in their own images using our classification model. The results from this study can be applied to identify further species by using other types of images. This model can also be used in the development of identification features in mobile applications. This study is a crucial step for integrating information from biology and AI-technology to develop a novel platform for use in forensic investigation.

The fly larvae of *Chrysomya megacephala* (Diptera: Calliphoridae), *Achoetandrus rufifacies* (Diptera: Calliphoridae), *Lucilia cuprina* (Diptera: Calliphoridae), and *Musca domestica* (Diptera: Muscidae) have long been recognized as crucial clues in medicolegal forensic entomology, especially in Thailand[1,2]. The estimation of the minimum postmortem interval (PMIm) from fly maggots plays a primary role in estimating the postmortem interval[3]. Moreover, human myiasis, the infestation of fly maggots in a living person, may be found in neglected

[1]Department of Business Administration, Faculty of Business Economics and Communications, Naresuan University, Muang, Phitsanulok 65000, Thailand. [2]Department of Electrical and Computer Engineering, Faculty of Engineering, Naresuan University, Muang, Phitsanulok 65000, Thailand. [3]Department of Microbiology and Parasitology, Faculty of Medical Science, Naresuan University, Muang, Phitsanulok 65000, Thailand. ✉email: nophawanb@nu.ac.th

people[1]. Therefore, fly larvae can be used to estimate the minimum abandonment duration[1]. The accuracy of PMIm determination and abandonment duration mainly relies on species identification[3].

Identification of fly maggots is more difficult because the general morphology of each species is quite similar[4–6]. It is vermiform, circular in cross section, tapering to a point, and having no legs[4–6]. Morphology of the posterior spiracle is one of the important characteristics that have been used for identification of the larval stage[4–6]. Posterior spiracles are the organs involved in the gas exchanges, which are in the last segment and are joined to the dorsal longitudinal trunks of the internal tracheal system. The spiracles are placed in two chitinized plates surrounded by the peritreme. Typically, the number of slits can be used to classify the stage of larvae[6,7]. For example, the third-stage larva has three slits in each side of the posterior spiracle. The pain points of species identification based on morphological characteristics are the requirement of taxonomic experts and time consuming. In addition, the number of taxonomists and classification experts has drastically decreased so far. Although molecular techniques and proteomic analysis have been currently used for identification of fly species, specific equipment and reagents are required in the laboratory[8–10].

To overcome the pain point, alternative automatic identification methods with expert-level classification accuracy are highly required in this field. Automatic species identification is important not only for flies in forensics, but also for other insects in general since it contributes to various purposes such as environment monitoring, pest diagnostics, and vector epidemiology[11]. Many previous studies indicated that automatic classification of insect species achieved high classification accuracy[12–18]. Currently, many classification models based on convolutional neural networks (CNNs) have been proposed in the field of computer vision. CNN is the type of multi-layered network learning algorithm, which usually contains multiple convolution layers, batch normalization layers, and fully connected layers[19]. Many previous reports demonstrated that the CNNs models have achieved species identification of insects with high accuracy and precision[12–18]. Based on our information, there have been no reports on the identification modeling of forensically-important fly maggots based on CNNs. Therefore, the novel method was proposed and the identification model was developed in 2 parts, including a custom object detection model for detection of position by using LabelImg, and YOLO programs, consecutively, and the multiple deep learning models to identify posterior spiracle images of four fly species including *C. megacephala* (Diptera: Calliphoridae), *C.* (*Achoetandrus*) *rufifacies* (Diptera: Calliphoridae), *L. cuprina* (Diptera: Calliphoridae), and *M. domestica* (Diptera: Muscidae). In this study, the assessment of the identification was compared with state-of-the-art CNNs including ResNet-101, Densenet161, Vgg19_bn, and AlexNet. The best performance model was selected and used to develop a web-application for automatic species identification.

Materials and methods

Preparation of posterior spiracle. Slides of the posterior spiracle were prepared following the method of Bunchu et al.[20]. Briefly, the morphology of the posterior spiracle was investigated by using the hydroxide clearing method. Maggots of each species (*C. megacephala*, *C.* (*A.*) *rufifacies*, *L. cuprina*, and *M. domestica*) were laboratory strains, which were obtained from the Medical Entomology Laboratory, Department of Microbiology and Parasitology, Faculty of Medical Science, Naresuan University. Adults of all species were identified by using the taxonomic key of Kurahashi and Bunchu[21]. The posterior spiracles of each species were photographed under the light microscope connected to the digital camera. All the images were used in the training process. For demonstration purposes, example images of each species were inverted using the rgb colors from rgb (255, 255, 255) to rgb (0, 0, 0) by using the image adjustment function of Adobe photoshop 2021 to increase the clarity of the images.

Image data set. In this study, the images of the posterior spiracles of four forensically-important fly species as mentioned above were analyzed. All original images have been categorized as "verified" after they were identified and confirmed by an expert. In total, 17,144 original images were used in this study and divided into training (12,000 images), validation (3428), and testing (1716) groups with a number ratio of 7:2:1. For each species, the number of images for training, validation, and testing were 3000, 857, and 429 respectively.

In the first step, we developed a custom object detection model for detecting the position of posterior spiracle in the input image by using LabelImg program and YOLO object detection algorithm. The process of the custom image detection was shown in Fig. 1. Briefly, this process divided the input image into subregions to predict multiple bounding boxes with their class probabilities for each region, training to create the most accurate object detection model for bounding boxes. Five hundred images were used for training the object detection model. The program used for labeling data for YOLO training was the LabelImg program. For the first step, the input images were added to the model to generate images in multiple scales. Then, the LabelImg program was used to label the position of the posterior spiracle in the image to detect the object of interest for further training (Fig. 1A). After that, the labeled images were used in training of the convolutional neural network (CNN) using YOLO object detection algorithm. The learning process was repeated to adjust the parameter several times to receive the highest accuracy. After passing through several convolutional layers, a fully connected layer was obtained to adjust the model output for detecting the position of the object in the image within bounding boxes (Fig. 1B). Therefore, this training model can be used to label the input image with the maximum confidence to select the final detection. The images were cropped tightly to the position of posterior spiracle (final detection). The purpose of this process was to ensure that each image was labeled precisely. This custom object detection data set was encoded to the JSON format with 224 × 224 pixels (detection result) (Fig. 1C). The cropped images with resolutions lower than 224 × 224 pixels were discarded from this study. The image data set was encoded to the JPEG format for the model training in the next step. This custom object detection model provided confidence and accuracy ≥ 80%.
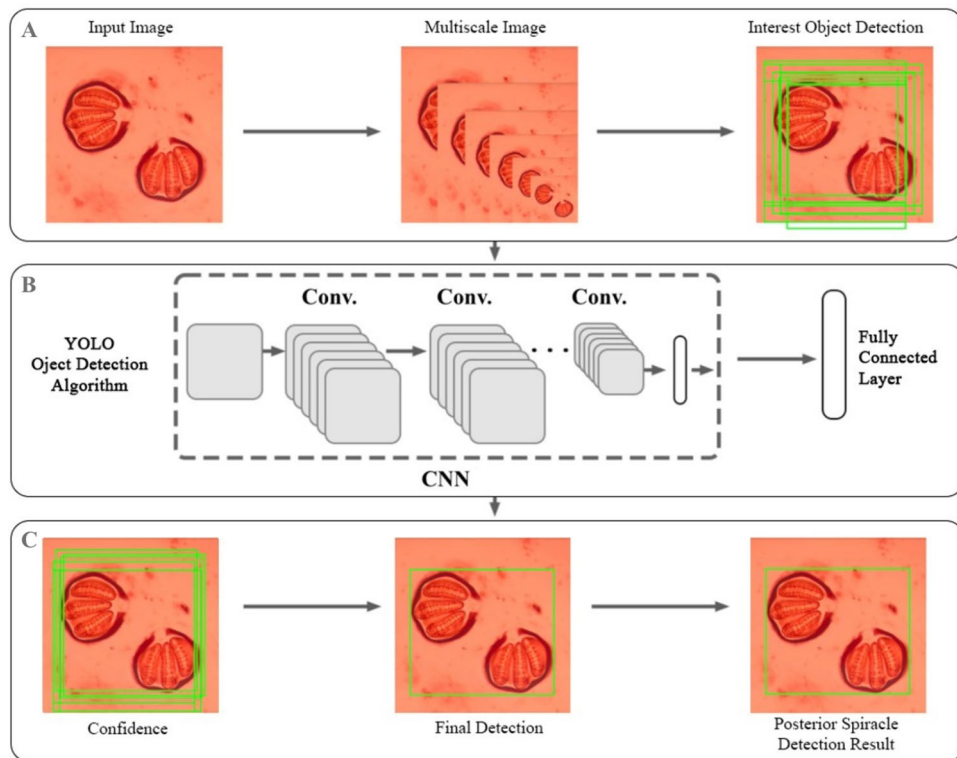
Figure 1. The processes of the custom object detection model used in this study was developed by using LabelImg program and YOLO object detection algorithm consisting 3 steps; (**A**) preparation process, (**B**) training process, and (**C**) detection process.

Model training. In this study, we compared the performance of four widely used models that vary in complexity of architecture to evaluate tradeoffs in accuracy and speed for species classification including ResNet-101, Densenet161, Vgg19_bn, and AlexNet. In the first step, we used the model from the previous step as mentioned above to speed up the training process. After randomly shuffling images in each species, the images within the species were divided by 7:2:1 ratio into training, validation, and testing groups. For each model, cropped images were resampled to the standard size for the input model. After using these models initially, we used 224×224 pixel images for Resnet-101, Densenet161, Vgg19_bn, and AlexNet. We used an image augmentation strategy following the previous study by Spiesman et al.[16]. Stochastic gradient descent (SGD) optimizer with an initial learning rate of 0.01 for all models was used. We used batch normalization, and models were trained for the 50 epochs using Nvidia K80, T4, P4 and P100 GPUs. At the species level, two metric model performances were calculated, including precision (1), recall (2), accuracy (3), and F1 score (4). Macro precision or mean of species-level precision scores and macro-recall or mean of species-level recall scores for each model were determined. The formulas for these indicators are as follows:

$$Precision = \frac{Tp}{Tp + Fp}, \tag{1}$$

$$Recall = \frac{Tp}{Tp + Fn}, \tag{2}$$

$$Accuracy = \frac{Tp + Tn}{Tp + Fp + Tn + Fn}, \tag{3}$$

$$F1score = \frac{2 \times Precision \times Recall}{Precision + Recall}, \tag{4}$$

where *Tp, Tn, Fp,* and *Fn* represent the number of true positive, true negatives, false positive and false negatives, respectively. For the support value, it is the number of actual occurrences of the class in the specified dataset.

In addition, a confusion matrix table was constructed to realize a class-wise comparison to examine the accuracy of the identifications (generalized accuracy) and ensure that the model could detect and classify objects in a reliable manner. The mean accuracy determined from the table was analyzed and assessed for all the models. The mean time needed to predict the image in the test data set was also quantified as the model speed. All speed
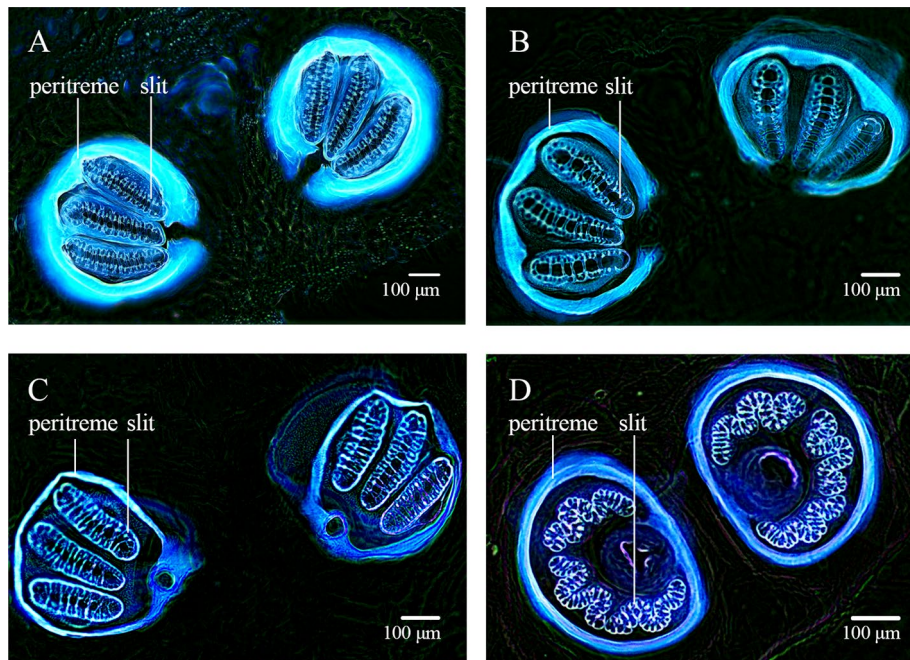
Figure 2. Morphology of posterior spiracles of four different fly species after inverting the image colors; (**A**) *Chrysomya* (*Achoetandrus*) *ruffifacies*, (**B**) *Chrysomya megacephala*, (**C**) *Lucilia cuprina*, (**D**) *Musca domestica*.

Model	#Params (million)	Model speed (s)	Model size (mb)	Macro precision	Macro recall	f1-score	support
Resnet101	44.6	246.0	170.96	1.00	1.00	1.00	3388
Densenet161	14.2	326.4	55.33	1.00	1.00	1.00	3388
Vgg19_bn	20.6	210.0	78.67	1.00	1.00	1.00	3388
AlexNet	**2.7**	**10.8**	**10.58**	1.00	1.00	1.00	3388

Table 1. Comparison of model size, speed, and performances of each studied model (The text in bold indicates the best value in each category).

tests run on the same system using Nvidia K80s, T4s, P4s and P100s GPUs. The best performance model from this study was selected and used to develop web applications for automatic species identification.

To confirm the results of this study, we tested this model with other images from outsources (internet and personal contact) and visualized the results using the PyTorch CNN visualizations[22].

Results and discussion

Of which at the third instar, the external morphology of larvae is quite similar; thus, the morphological identification used to differentiate between its genera or species, generally includes cephalophalyngeal skeleton, anterior spiracle, and posterior spiracles. The morphology of the posterior spiracle is one of the important characteristics for identification. A typical morphology of the posterior spiracle of third stage larvae was shown in Fig. 2. Based on studying under light microscopy, the posterior spiracle of *M. domestica* was clearly distinguished from the others. On the other hand, the morphology of the posterior spiracle of *C. megacephala* and *A. rufifacies* was quite similar. For *C. megacephala* and *C. rufifacies*, the peritreme, a structure encircling the three spiracular openings (slits), was incomplete and slits were straight as shown Fig. 2A,B, respectively. The complete peritreme encircling three slits was found in *L. cuprina* and *M. domestica* as in Fig. 2C,D, respectively. However, only the slits of *M. domestica* were sinuous like the M-letter (Fig. 2D). Their morphological characteristics found in this study were like the descriptions in the previous reports[23–25].

For model training, four of the CNN models used for species-level identification of fly maggots provided 100% accuracy rates and 0% loss. Number of parameter (#Params), model speed, model size, macro precision, macro recall, f1-score, and support value were also presented in Table 1. The result demonstrated that the AlexNet model provided the best performance in all indicators when compared among four models. The AlexNet model used the least number of parameters while the Resnet101 model used the most. For model speed, the AlexNet model provided the fastest speed, while the Densenet161 model provided the slowest speed. For the model size, the AlexNet model was the smallest, while the Resnet101 model was the largest which corresponded to the number of parameters used. Macro precision, macro recall, f1-score and support value of all models were the same.
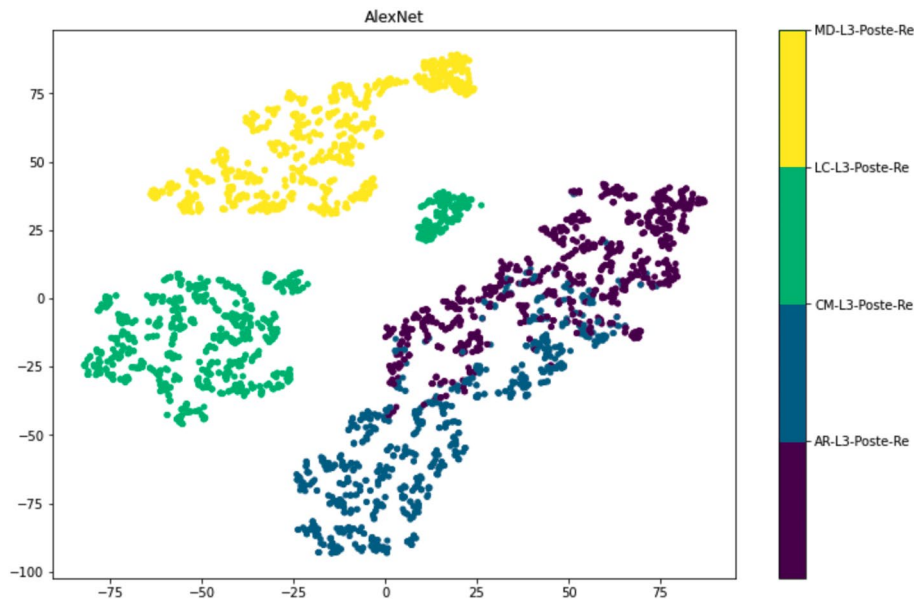
Figure 3. tSNE visualization of the AlexNet model by dimensionality reduction of the penultimate features (The test data are shown in colors for different classes; yellow refers to data set of *Musca domestica*, green refers to data set of *Lucilia cuprina*, greenish blue refers to data set of *Chrysomya megacephala*, and dark purple refers to data set of *Chrysomya* (*Achoetandrus*) *ruffifacies*).

As the training results presented in the supplementary data (Fig. S1), all models provided 100% accuracy and 0% loss in the early stage of training (< 10 epochs). This could be due to the training and testing processes with cropping specific portions of the fly images using our custom object detection model. Moreover, all images were from the laboratory strains, and their variation of morphological characteristics might be less than their wild counterparts. As a result, the training time was short, and the accuracy of the model was high. Of the four models tested, AlexNet demonstrated a good balance between performance and speed. This model can proceed the system the fastest and its model size was the smallest. The speed and accuracy of AlexNet made it useful for web-based and mobile applications that relied on both speed and reliable predictions. Speed was a factor in user satisfaction and would be important for future development such as video-based applications[16]. Therefore, we focused on the AlexNet results for the remainder of this article.

By using tSNE visualization, the AlexNet model can separate species explicitly into distinct groupings based on characteristics extracted from the model as shown in Fig. 3. All four species were separated with overlapping data of *C. megacephala* and *C. ruffifacies*. It could be due to the similarity of morphological characteristics of both species[24]. This result indicated that the performance of the AlexNet model was equal to human practice.

The five hidden convolutional layers of AlexNet on four example images of posterior spiracles were visualized as shown in Fig. 4. In each layer, the posterior spiracle of each species was randomly selected for visualization. The color images were generated to clearly show the patterns. The multiple hidden layers could extract a rich hierarchy of patterns from the input image, including low-level, middle-level, and high-level patterns that can be observed. The lower layers extracted the detailed local patterns from images, such as the textures, margins, and others. The complexity and variation of the visualized patterns increased when it comes to the middle layers. The higher layers extracted specific pattern features.

The classification results (validation and test) for each image are displayed in the confusion matrices (Fig. 5) which show how predicted species (columns) correspond to the actual species (rows). The values along the diagonal indicate the number of correct predictions, whereas off-diagonal values indicate misclassifications. Interestingly, no misclassification was found after testing the model by using the test images (Fig. 5A). Therefore, the results indicated that the predictions of the AlexNet model match the taxonomic expert classification. The confusion matrix showed misclassification between *C. megacephala* and *C. ruffifacies* (Fig. 5B), corresponding to the results of tSNE visualization. When the model was tested with the outsource images, the accuracy of the classification for *C. megacephala*, *C. ruffifacies, L. cuprina*, and *M. domestica* was 94.94, 98.02, 98.35, 100%, respectively (Fig. 5B). The results from using the Heatmap program showed that the prediction accuracy of this model was still high (98.70–100%), depending on image conditions (Fig. 6).

The framework of the AlexNet model was demonstrated in Fig. 7. This model consists of 5 convolutional layers (Conv1, Conv2, Conv3, Conv4, and Conv5) and 3 fully connected (FC6, FC7, and FC8) layers[19]. The convolutional layers extract features, and then the fully connected layers combine the features to model the outputs. The example of characteristics of the images presented in all 5 convolutional layers were shown in Fig. 4.

Previously, CNNs have been used successfully to identify different cells or species[11–18,26,27]. This study also confirmed the efficiency of CNNs in identifying fly species.

Finally, we created a web application called "The Fly" by using our classification model for identifying species of fly maggots. The web application is available at https://thefly.ai. Users can identify species of fly maggots
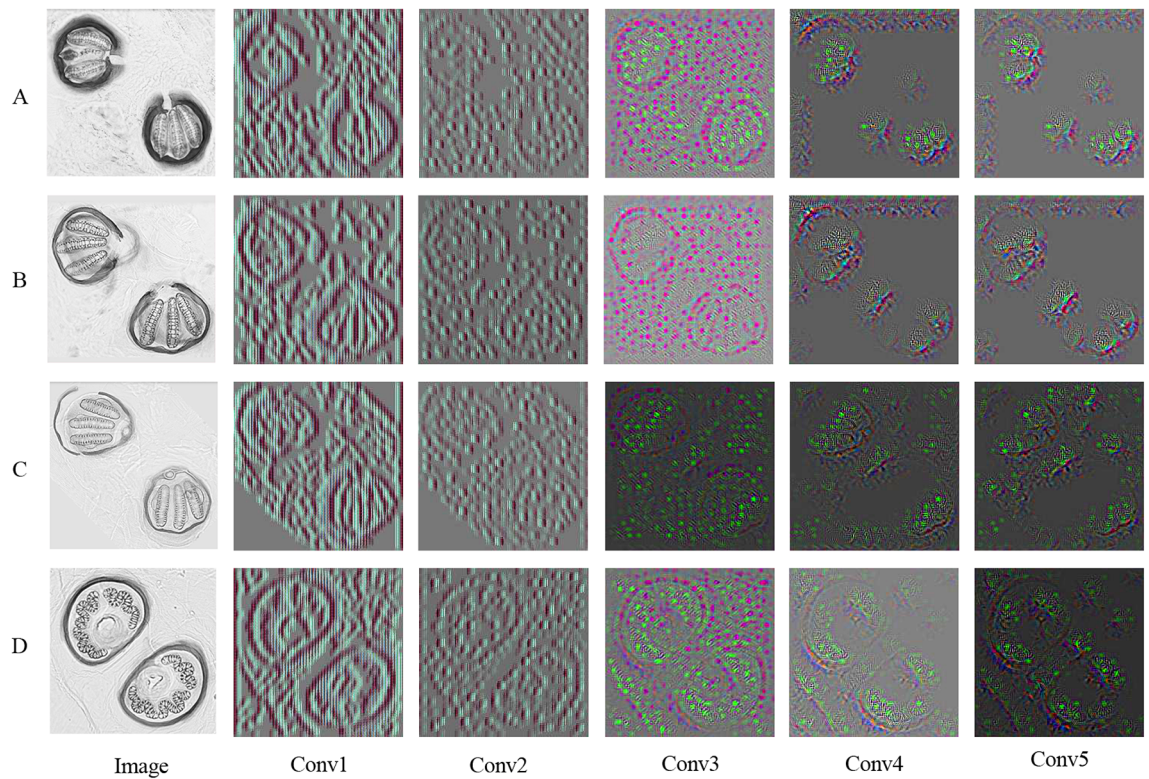
Figure 4. Visualization of hidden convolutional layers in AlexNet for four example images (To clearly show the patterns, we generated the color images; (**A**) *Chrysomya* (*Achoetandrus*) *ruffifacies*, (**B**) *Chrysomya megacephala*, (**C**) *Lucilia cuprina*, (**D**) *Musca domestica*).
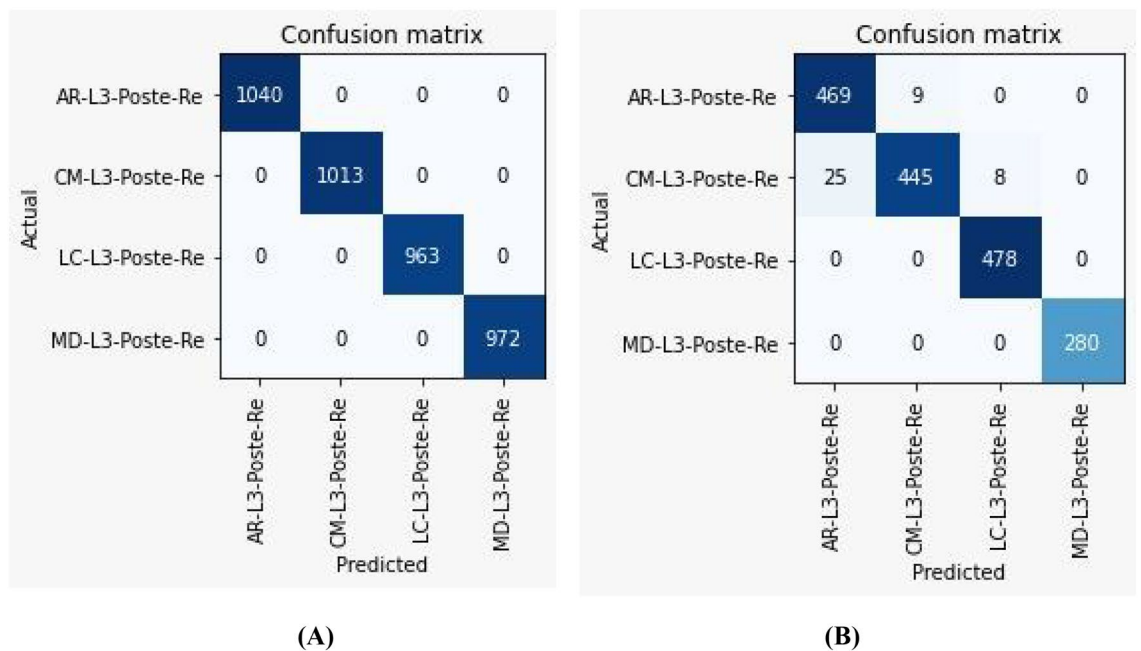

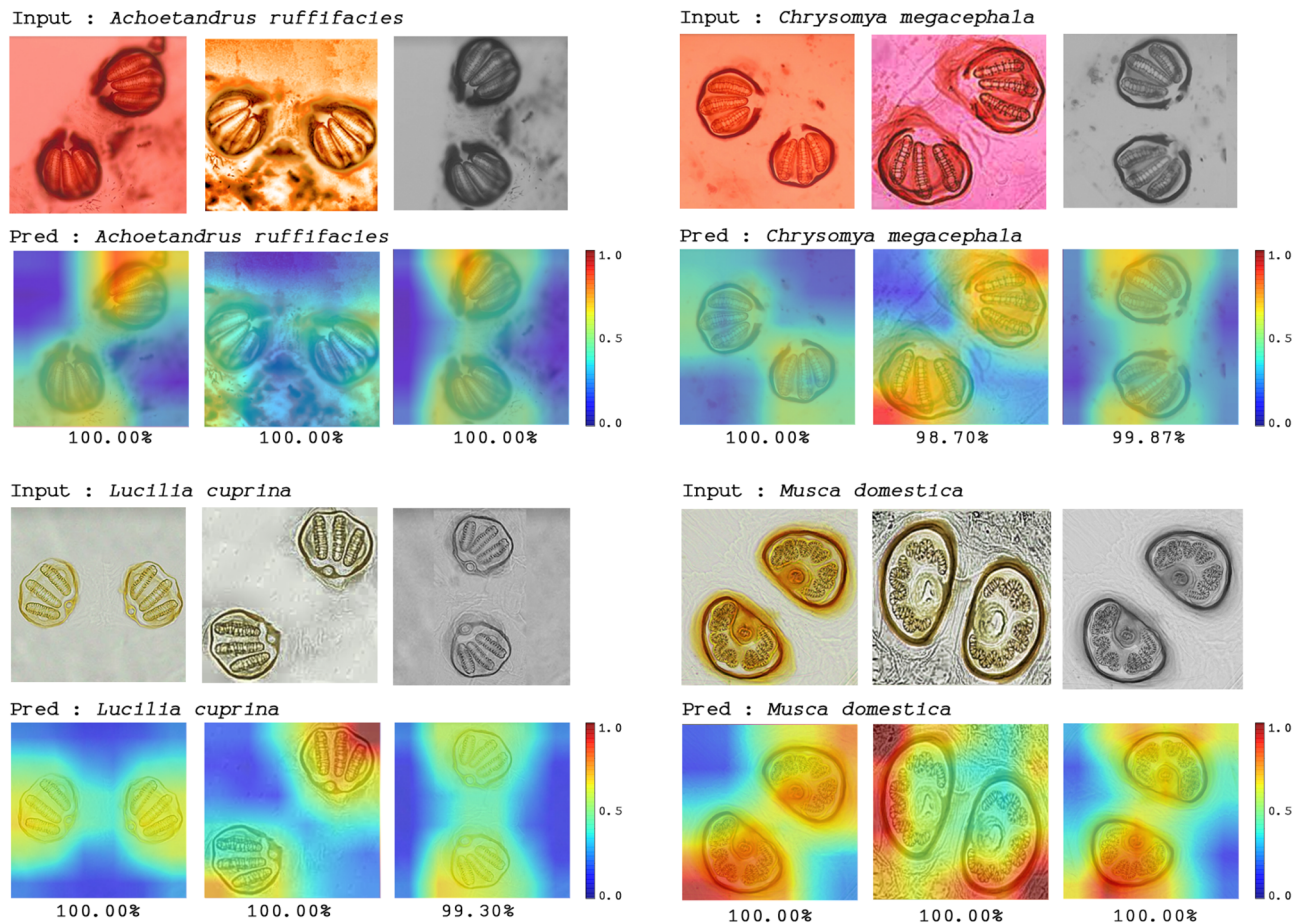**Figure 5.** The confusion matrix achieved by the AlexNet model: (**A**) Validation (**B**) Test dataset.

6

Figure 6. Heatmap of attention maps of AlexNet on example images showing prediction accuracy (98.70–100%) of this model for classification of each fly species in different image conditions.

by uploading their images of posterior spiracles and the result with associated probability percentage will then be shown. This web application can be accessed and used on both desktop and mobile browsers. In terms of performance limitations, this web application was designed to identify only four species of fly maggots using images of posterior spiracles. This web application is the beginning step of the development of automatic species-identification for fly species in Order Diptera. More images of these four species and other species must be studied in the future. In addition, the results from this study will be applied to develop a feature as a microservice for the identification of fly maggots in a mobile application called iParasites which is currently available on AppStore and GooglePlay. We, nonetheless, wish to project that taxonomic experts are still important and critical for the development of this automatic identification by AI-based imaging system as mentioned in a previous report[11].

Conclusion

Currently, CNNS have been successfully used in a wide range of scientific discipline including Entomology[11–19,26,27]. This study demonstrated that all state-of-the-art CNNs used in this study including ResNet-101, Densenet161, Vgg19_bn, and AlexNet provided high efficiency in the computer vision aspect. AlexNet showed the fastest speed to process the identification model and could reliably identify species of fly maggots based on images of posterior spiracles. However, the results from this study confirmed that the AI-based imaging system will be useful in developing automatic species identification for flies and other insects. In the next step, more images of other fly species should be collected as training data sets. Moreover, the results from this study can be used in the development of identification features in mobile applications and in further species identification by using other types of images.
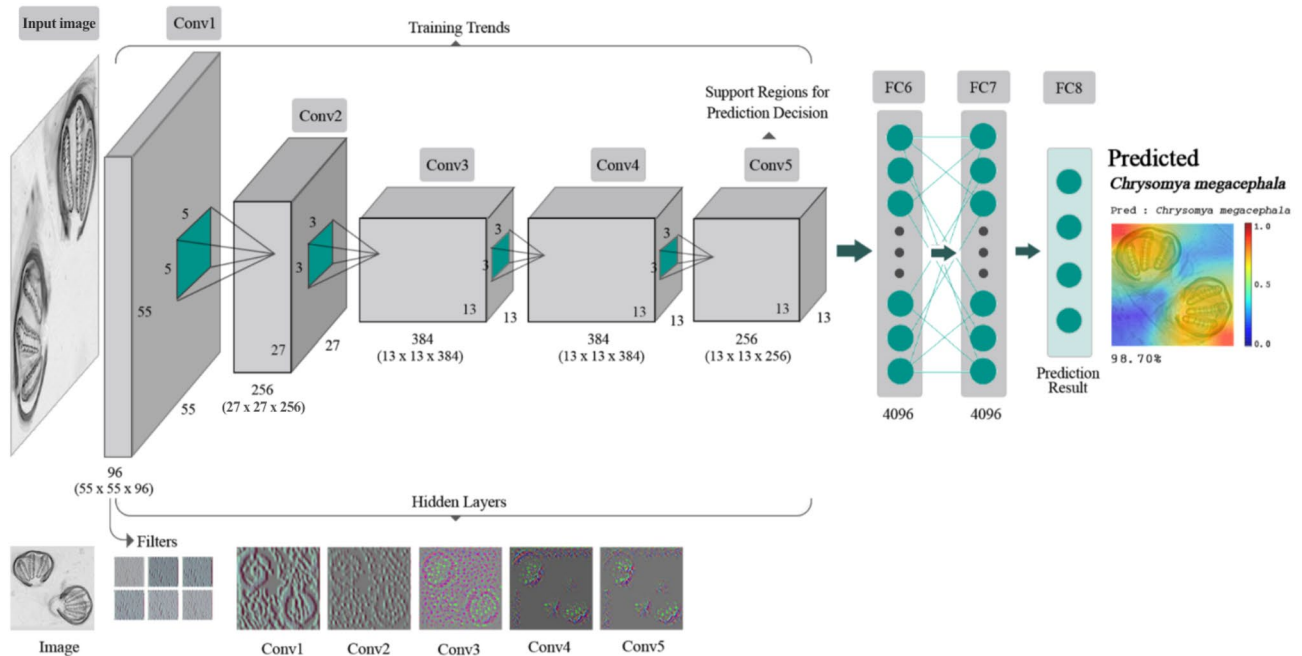
Figure 7. Framework of the proposed interpretation architecture for the deep learning model, AlexNet, in this study. The AlexNet model contains eight learned layers with weights (five convolutional layers and three fully connected layers), namely Conv1 as a convolutional layer that accepts an input image tensor ($224 \times 224 \times 3$) and performs Convolution to obtain the position and strength of the input image properties causing tensor output ($55 \times 55 \times 96$), Conv2 as a convolutional layer that generates tensor output ($27 \times 27 \times 256$), Conv3 as a convolutional layer that creates tensor output ($13 \times 13 \times 384$), Conv4 as a convolutional layer that creates tensor output ($13 \times 13 \times 384$), Conv5 as a convolutional layer that creates tensor output ($13 \times 13 \times 256$), FC6 as a fully connected layer (4096) which flattens tensor output from Conv5 into a single vector by weight sharing, resulting in tensor output (4096×1), FC7 as a fully connected layer which performs the same actions as FC6 and generates tensor output (4096×1), and FC8 as a fully connected layer to generates tensor output (1000×1) which is the prediction result and which contains 1000 neurons corresponding to the 1000 natural image classes.

Data availability

The dataset and source codes for this study are publicly available through the second author's GitHub repository: https://github.com/pharinyab/Fly-PosteriorSpiracles-AlexNet.

References

1. Bunchu, N. Blow fly (Diptera: Calliphoridae) in Thailand: Distribution, morphological identification and medical importance appraisals. *Int. J. Parasitol. Res.* **4**, 57–64 (2012).
2. Sukontason, K. *et al.* Forensic entomology cases in Thailand: A review of cases from 2000 to 2006. *Parasitol. Res.* **101**, 1417–1423 (2007).
3. Harvey, M. L., Gasz, N. E. & Voss, S. C. Entomology-based methods for estimation of postmortem interval. *Res. Rep. Forensic Med. Sci.* **6**, 1–9 (2016).
4. Greenberg, B. & Szyska, M. L. Immature stages and biology of 15 species of Peruvian calliphoridae (Diptera). *Ann. Entomol. Soc. Am.* **77**, 488–517 (1984).
5. Greenberg, B. & Tantawi, T. I. Different developmental strategies in two boreal blow flies (Diptera: Calliphoridae). *J. Med. Entomol.* **30**, 481–484 (1993).
6. Liu, D. & Greenberg, B. Immature stages of some flies of forensic importance. *Ann. Entomol. Soc. Am.* **82**, 80–93 (1989).
7. Flores, D. *et al.* Fluorescence imaging of posterior spiracles from second and third instars of forensically important *Chrysomya rufifacies* (Diptera: Calliphoridae). *J. Forensic Sci.* **61**, 1578–1587 (2016).
8. Gemmellaro, M. D., Hamilton, G. C. & Ware, J. L. Review of molecular identification techniques for forensically important Diptera. *J. Med. Entomol.* **27**, 887–902 (2019).
9. Park, J. H. *et al.* Molecular identification of forensically important calliphoridae and sarcophagidae species using ITS2 nucleotide sequences. *Forensic Sci. Int.* **284**, 1–4 (2018).
10. Phuwanatsarunya, P., Hongsrichan, N., Chaiwong, T., Panya, M. & Bunchu, N. Species identification of forensically important fly larvae by integrating morphological characteristics and protein profiles. *Songklanakarin J. Sci. Technol.* **42**, 1215–1220 (2020).
11. Høye, T. T. *et al.* Deep learning and computer vision will transform entomology. *Proc. Natl. Acad. Sci. U.S.A.* **2**, 118 (2021).
12. Hernández-Serna, A. & Jiménez-Segura, L. F. Automatic identification of species with neural networks. *PeerJ* **2**, e563. https://doi.org/10.7717/peerj.563 (2014).
13. Motta, D. *et al.* Application of convolutional neural networks for classification of adult mosquitoes in the field. *PLoS ONE* **14**, e0210829 (2019).

14. Ong, S. Q. *et al.* Implementation of a deep learning model for automated classification of *Aedes aegypti* (Linnaeus) and *Aedes albopictus* (Skuse) in real time. *Sci. Rep.* **11**, 9908 (2021).
15. Park, J. *et al.* Classification and morphological analysis of vector mosquitoes using deep convolutional neural networks. *Sci. Rep.* **10**, 1012 (2020).
16. Spiesman, B. J. *et al.* Assessing the potential for deep learning and computer vision to identify bumble bee species from images. *Sci. Rep.* **11**, 7580 (2021).
17. Ye, S., Lu, S., Bai, X. & Gu, J. ResNet-locust-BN network-based automatic identification of east asian migratory locust species and instars from RGB images. *Insects* **11**, 458. https://doi.org/10.3390/insects11080458 (2020).
18. Visalli, F., Bonacci, T. & Borghese, N. A. Insects image classification through deep convolutional neural networks. *Smart Innov. Syst. Technol.* **184**, 217–228 (2021).
19. Boukaye, B. T., Bernard, K. F. & Fana, T. Deep convolution neural network for image recognition. *Ecol. Inform.* **48**, 257–268 (2018).
20. Bunchu, N. *et al.* Morphology and developmental rate of the blow fly, *Hemipyrellia ligurriens* (Diptera: Calliphoridae): Forensic entomology applications. *J. Parasitol. Res.* **2012**, 371243. https://doi.org/10.1155/2012/371243 (2012).
21. Kurahashi, H. & Bunchu, N. The Blow flies recorded from Thailand, with the description of a new species of *Isomyia* WALKER (Diptera, Calliphoridae). *Jpn. J. Syst. Ent.* **17**, 237–278 (2011).
22. Ozbulak, U. *Convolutional Neural Network Visualizations* (2019). https://github.com/utkuozbulak/pytorch-cnn-visualizations. (Accessed 24 August 2021).
23. Sukontason, K. L. *et al.* Larval morphology of *Chrysomya megacephala* (Fabricius) (Diptera: Calliphoridae) using scanning electron microscopy. *J. Vector Ecol.* **28**, 47–52 (2003).
24. Sukontason, K., Piangjai, S., Siriwattanarungsee, S. & Sukontason, K. L. Morphology and developmental rate of blowflies *Chrysomya megacephala* and *Chrysomya rufifacies* in Thailand: Application in forensic entomology. *Parasitol. Res.* **102**, 1207–1216 (2008).
25. Sukontason, K., Sukontason, K. L., Ngern-klun, R., Sripakdee, D. & Piangjai, S. Differentiation of the third instar of forensically important fly species in Thailand. *Ann. Entomol. Soc. Am.* **97**, 1069–1075 (2004).
26. Acevedo, A. *et al.* Recognition of peripheral blood cell images using convolutional neural networks. *Comput. Methods Progr. Biomed.* **180**, 105020 (2019).
27. Tan, C. *et al.* Identification of different species of *Zanthoxyli Pericarpium* based on convolution neural network. *PLoS ONE* **15**, e0230287 (2020).

Acknowledgements

Author contributions

D.A. conceived the study, contributed model training, and provided acquisition of funding. P.B. performed the modeling, designed all visualization, and created a web application (https://thefly.ai). P.M. provided the critical feedback, editing, and approval of the final draft. W.N. prepared slides of posterior spiracles, and took photographs under a stereomicroscope, N.B. initiated and designed the study, identified the specimens, prepared the data sets, performed training of models, wrote the first draft of the manuscript. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-08823-8.

Correspondence and requests for materials should be addressed to N.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.