DOI 10.18699/SBB-2022-655

# Classification of fruit flies by gender in images based on the YOLOv4-tiny neural network using mobile devices

Genaev M.A.[1, 2], Komyshev E.G.[1, 2], Afonnikov D.A.[1, 2]*

[1] *Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

[2] *Kurchatov Genomic Center of the Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia*

* *ada@bionet.nsc.ru*

**Key words:** YOLOv4-tiny, fruit flies detection, gender classification, mobile devices, FlyCounter, Android

*Motivation and Aim*: The fruit fly *Drosophila melanogaster* is a classic research object in genetics and systems biology. In genetic analysis of flies, a routine task is to determine the offspring size and gender ratio in their populations. Currently, these estimates are made manually, which is a very time-consuming process. The counting and gender determination of flies can be automated by using image analysis with deep learning neural networks on mobile devices.

*Results*: We proposed an algorithm based on the YOLOv4-tiny network to identify drosophila flies and determine their gender based on the protocol of taking pictures of insects on a white sheet of paper with a cell phone camera. 406 images with labels (total number of flies 35025) were collected for training the network. The increase of the training sample was performed due to generation of synthetic images (superposition of the contours with the artificial background, as well as mosaic transposition of image fragments). which allowed increasing the recall parameter from 0.981 to 0.991, precision from 0.628 to 0.726, and F1 from 0.766 to 0.838. The high value of the recall parameter indicates a high accuracy of fly identification in the image. Gender recognition is less accurate. We compared the accuracy of the gender recognition by neural network algorithm with the results obtained by expert geneticists on the same set of images. On the set of images obtained from mobile devices, the recognition accuracy of machine learning and expert geneticists was close (F1 = 0.726 vs 0.716), while on a sample of images under high-quality illumination and digital camera conditions the machine learning algorithms were noticeably worse (F1 = 0.824 vs 0.900). The results also show that the gender determination for females (F1 = 0.812) is worse than that of males (F1 = 0.873). Among the factors most strongly influencing the accuracy of fly gender recognition, the factor of fly position on the paper was the most important: flies with the lateral position are recognized the worst, but their proportion is the highest. In addition, increased light intensity and higher quality of the device cameras have a positive effect on the recognition accuracy.

*Conclusion*: We implement our method as the Android app FlyCounter for mobile devices, which performs all the image processing steps using device processors only. The time that the YOLOv4-tiny algorithm takes to process one image is less than 4 seconds.

*Acknowledgements*: The data analysis performed using computational resources of the "Bioinformatics" Joint Computational Center supported by the budget project No. FWNR-2022-0020.