Classification accuracy and compatibility across devices of a new Rapid-E+ flow cytometer

Branko Sikoparija[1], Predrag Matavulj[2], Isidora Simovic[1], Predrag Radisic[1], Sanja Brdar[1], Vladan Minic[1], Danijela Tesendic[3], Evgeny Kadantsev[4], Julia Palamarchuk[4] and Mikhail Sofiev[4]

[1]BioSense Institute Research Institute for Information Technologies in Biosystems, University of Novi Sad, Novi Sad, 2100, Serbia
[2]Institute for Data Science, University of Applied Sciences North Western Switzerland, Windish, 5210, Switzerland
[3]Department of mathematics and informatics, Faculty of Sciences, University of Novi Sad, Novi Sad, 21000, Serbia
[4]Finnish Meteorological Institute, Helsinki, Erik Palmenin Aukio 1, FI-00560, Finland

Correspondence to: Mikhail Sofiev (Mikhail.Sofiev@fmi.fi)

Anonymous Referee #2 (Citation: https://doi.org/10.5194/egusphere-2024-187-RC2)

A new device from Plair SA company Rapid-E+ is investigated in current study. A two-step classification was applied. At the first step of classification pollen are separated from non-pollen particles. At the second step pollen are classified into 27 pollen classes. It as established, that as with previous device rapid-E remains a large discrepancy between the signals measured by different devices. Therefore individual models need to be trained for every device. In overall the paper is well prepared. Some minors points must be corrected before final publication.

Reply: The authors would like to thank Referee #2 for reviewing the manuscript and positive opinions. We are grateful for helpful comments, which we have used to improve our manuscript. Below we answer the questions and indicate the changes we have made to the revised manuscript.

The paragraph about the used model (135-150) should be extended. ResNet-18 has 4 2-layer blocks. What does mean 4-block-layer or 3-block-layer? In context of ResNet style models, a block is a container of layers. It means that a block is a larger unit than a layer. It seems that not all neural networks have 18 layers, because their architectures are different. That to present the architectures to readers, a good point would be to prepare a architecture table as Table 3 in the paper (https://arxiv.org/pdf/1803.06131). It would also be useful to show the size of the inputs arrays received by each mode sub-network.

Reply: The paragraph has been extended as requested, and now reads:

"The ResNet architecture with shortcut connections was chosen for its proven superior performance in classifying pollen using Rapid-E measurements (Matavulj et al., 2023; Daunys et al., 2022). Given the variability of input data, we adapted the ResNet model inspired by the 18-layer version. Specifically, we implemented a 4-block layer for the fluorescence spectrum and lifetime, a 3-block layer for the 447 nm laser scattering images, and a 1-block layer for the 637 nm laser scattering image. Details of these configurations are provided in Table B1. These architectures were selected because they demonstrated the best performance for the respective data types in the previous device version (Matavulj et al., 2023). The block-layers contained three convolutional layers, where we captured a residual following the initial convolution. Subsequently, at the closure of each block layer, we established a residual connection to the layer's output. Following the completion of all block layers, an additional convolutional layer was integrated. This was followed by a global average pooling, which averaged over the spatial dimensions of the images. The network initially learned from each type of input separately. After this initial training, we transferred the learned features from these individual inputs (specifically, the parts of the network responsible for feature extraction, known as convolutional blocks) to a new network. This new network processed all different inputs together by

equalizing the features from each input using a fully connected (FC) layer, which were then merged. Finally, the network was trained only to classify this combined data using another FC layer with a SoftMax function. During this phase, the weights of the feature extractors (the convolutional blocks) were kept unchanged. This means that while the network was learning to classify the merged data, the initial parts that extract features from each input type did not undergo any further changes."

Table B1: Feature extractors for each data type. The convolutional layers are represented as N x M, F, where N X M represents the filter size for the 2D convolution, while F represents the number of feature maps.

Input type:	Scattered light images	Fluorescence spectrum	Fluorescence lifetime	Infrared image
Input dimension:	120x14	5x14	3x22	4x4
conv1	7 x 7, 70	1 x 7, 70	1 x 7, 70	3 x 3, 70
block1	3 x 3, 70 3 x 3, 70 3 x 3, 70	1 x 3, 70 1 x 3, 70 3 x 3, 70	1 x 3, 70 1 x 3, 70 3 x 3, 70	3 x 3, 70 3 x 3, 70 3 x 3, 70
block2	5 x 5, 140 5 x 5, 140 3 x 3, 140	1 x 7, 140 1 x 5, 140 3 x 3, 140	1 x 5, 140 1 x 5, 140 3 x 3, 140	
block3	7 x 1, 200 5 x 5, 200 3 x 3, 200	1 x 5, 200 1 x 5, 200 3 x 3, 200	1 x 3, 200 1 x 5, 200 3 x 3, 200	
block4		1 x 3, 300 1 x 5, 300 3 x 3, 300	1 x 3, 300 1 x 5, 300 3 x 3, 300	
final_conv	3 x 3, 200	3 x 3, 300	3 x 3, 300	4 x 4, 70

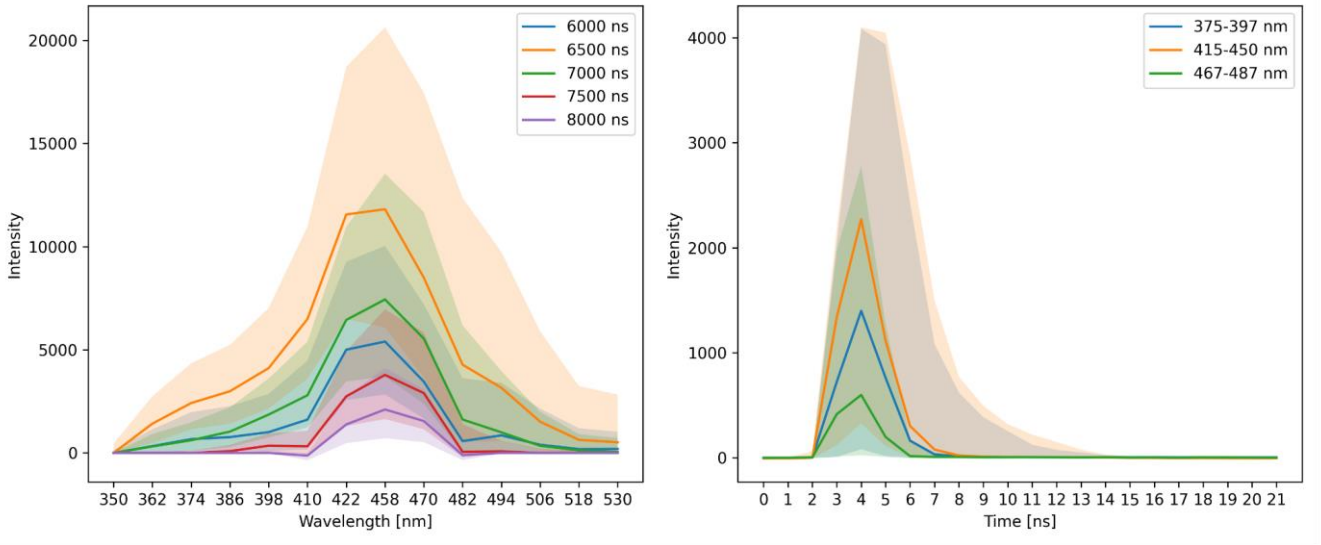
The scattering images of Rapid-E were of variable length. What is case in Rapid-E+? If they are of variable size, how the issue was solved?

Reply: The scattering image in Rapid-E+ has a fixed length of 120 acquisitions across 14 scattering angles. We have now noted that in chapter 2.1 "The 447 nm laser scattering is measured now in two polarization planes at a narrower angle window and fixed duration limited to 120 acquisitions."

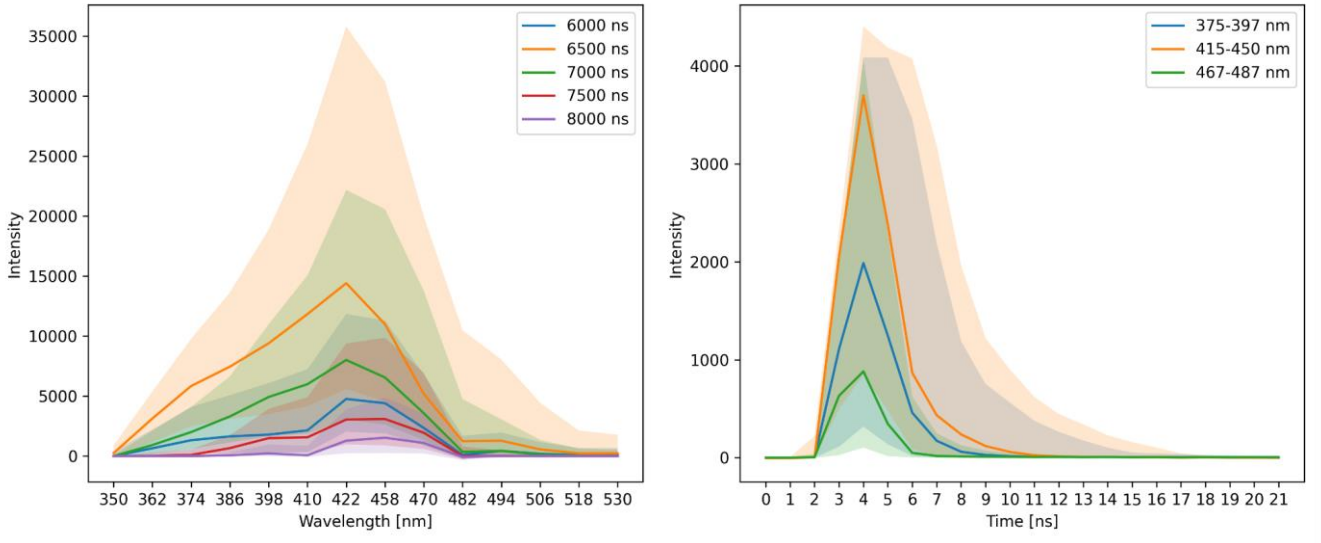
It would seem that in the graphs shown in Figure B2 of Appendix B, the intensity should be positive. However, a large part of the shadow, which is bounded by the curvatures calculated adding and subtracting standard deviation to/from the mean, is in the negative range. The standard deviation is appropriate to characterize the dispersion when the values follow a normal distribution. In this case, the distribution does not appear to be normal and, moreover, asymmetric. In this case, it is preferable to represent in the center by solid line a median curve and to delimit the shaded area by curves corresponding to quantiles symmetrical with respect to the median.

Reply: Figure B2 of Appendix B has been changed accordingly, where a solid line now represents a median and the shaded area represents the interquartile range (25th - 75th percentiles).

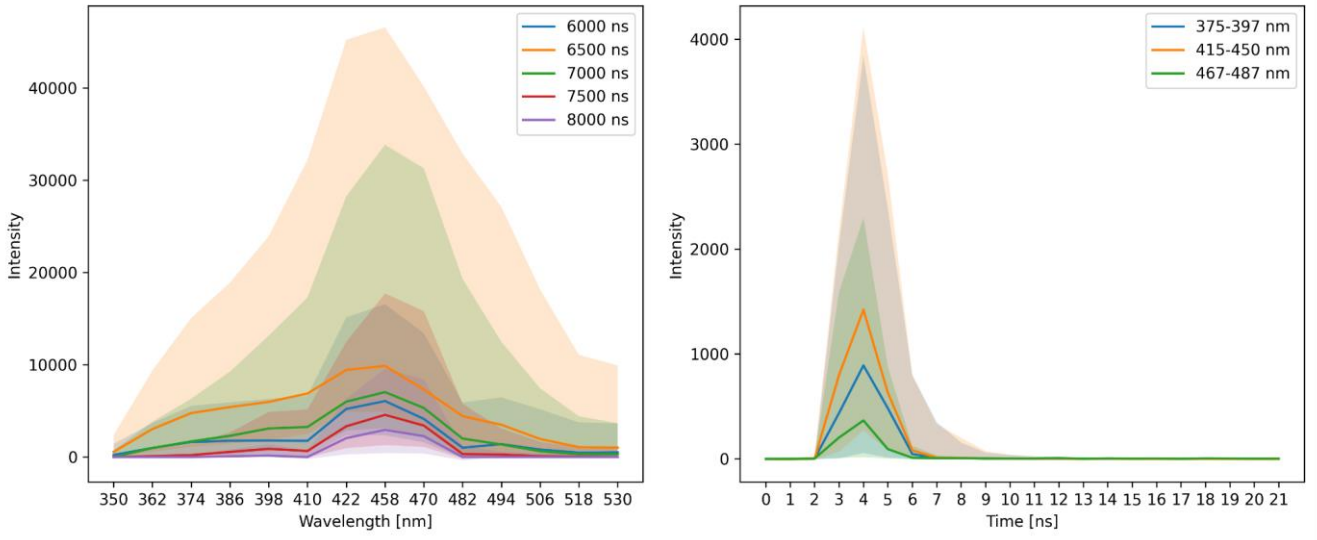
(A)


6000 ns
6500 ns
7000 ns
7500 ns
8000 ns
Intensity
Wavelength [nm]
375-397 nm
415-450 nm
467-487 nm
Intensity
Time [ns]

(B)


6000 ns
6500 ns
7000 ns
7500 ns
8000 ns
Intensity
Wavelength [nm]
375-397 nm
415-450 nm
467-487 nm
Intensity
Time [ns]

(C)



(D)

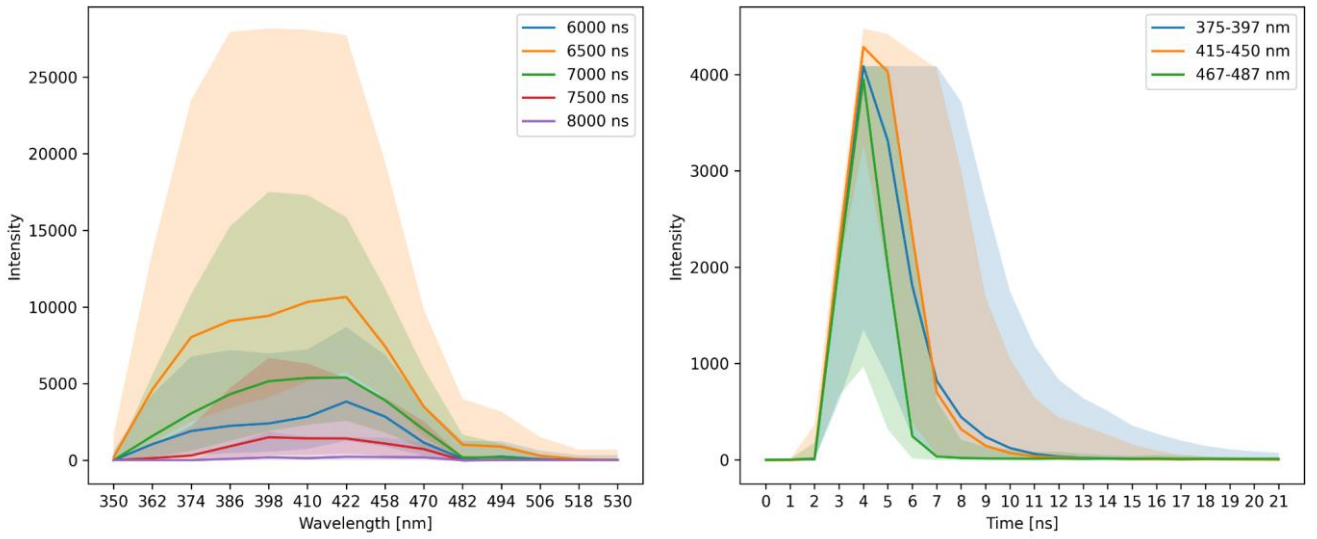


Figure B2: Median (with the interquartile range 25th - 75th percentiles depicted by area around lines) fluorescence spectrum (left side) and lifetime (right side) measurements after preprocessing for: (A) *Betula pendula*, (B) *Fraxinus pennsylvanica*, (C) *Juglans regia* and (D) *Platanus orientalis* reference pollen measured in "pollen mode" on Novi Sad Rapid-E+ device. (y-axis is "unitless")

References

Matavulj, P., Panić, M., Šikoparija, B., Tešendić, D., Radovanović, M., and Brdar, S.: Advanced CNN Architectures for Pollen Classification: Design and Comprehensive Evaluation, Applied Artificial Intelligence, 35, 1, e2157593, https://doi.org/10.1080/08839514.2022.2157593, 2023.

Daunys, G., Šukienė, L., Vaitkevičius, L., Valiulis, G., Sofiev, M., and Šaulienė, I.: Comparison of computer vision models in application to pollen classification using light scattering. Aerobiologia, https://doi.org/10.1007/s10453-022-09769-0, 2022.