PNAS
www.pnas.org

Supplementary Information for

Label-free Hematology Analysis Using Deep-Ultraviolet Microscopy

Ashkan Ojaghi[1], Gabriel Carrazana[1], Christina Caruso[2], Asad Abbas[1], David R. Myers[1,2], Wilbur A. Lam[1,2], Francisco E. Robles[1]*

[1]Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory University, Atlanta, GA.

[2]Aflac Cancer and Blood Disorders Center of Children's Healthcare of Atlanta and Department of Pediatrics, Emory University School of Medicine, Atlanta, GA.

*Corresponding Author: Francisco E. Robles

Email: robles@gatech.edu

This PDF file includes:

www.pnas.org/cgi/doi/10.1073/pnas.2001404117

Supplementary Information:

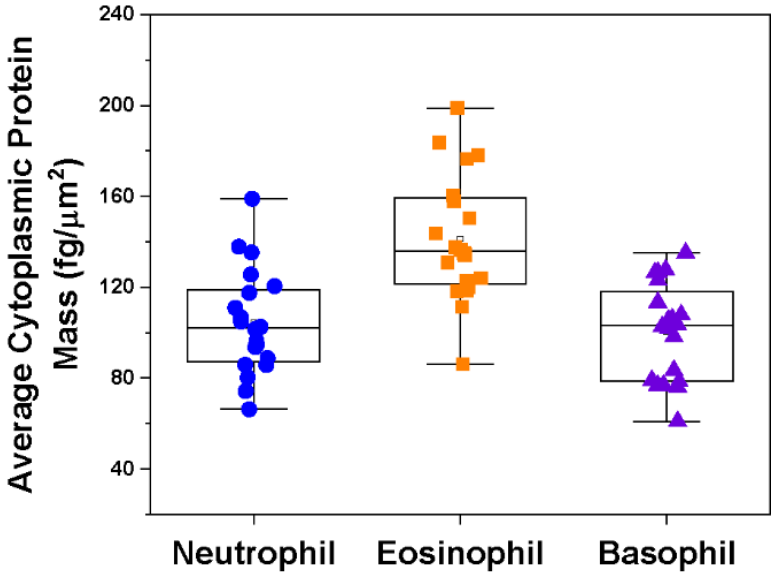

Fig. S1. The average cytoplasmic protein mass for different granulocyte subtypes.

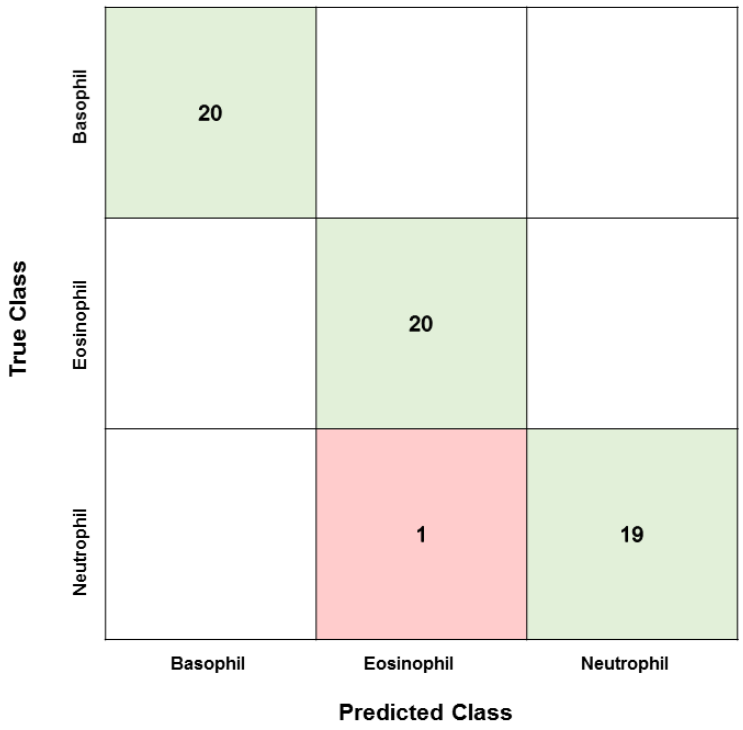

Fig. S2. The confusion matrix for the multi-wavelength SVM model.

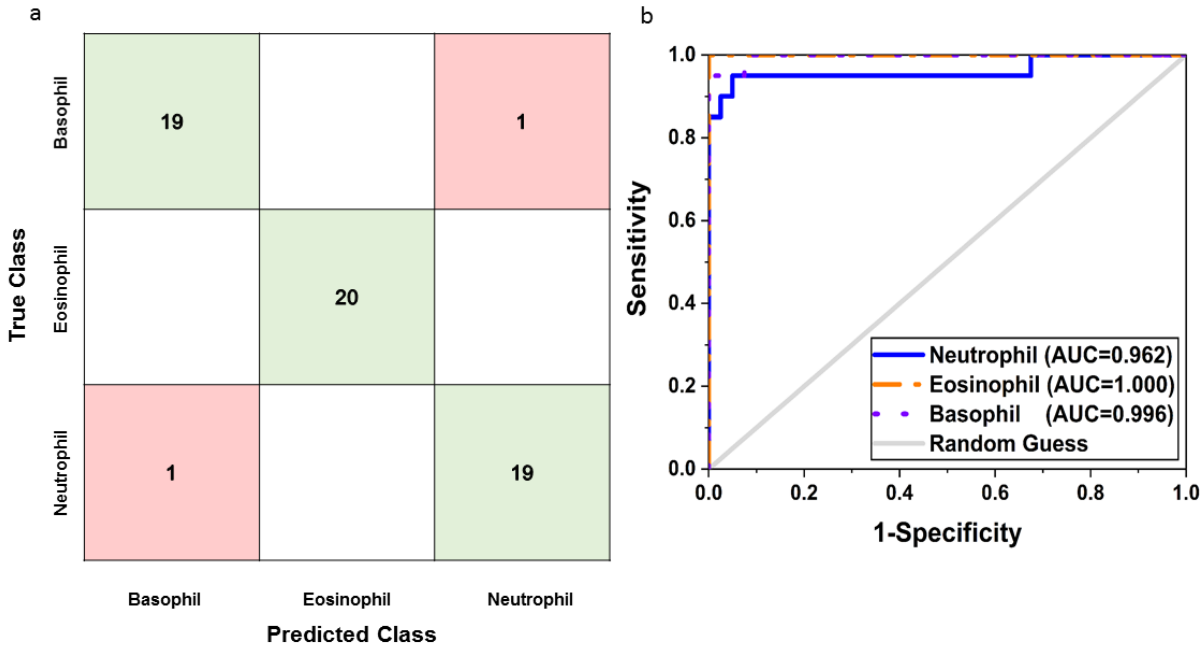

Fig. S3. The confusion matrix (a) and the ROC curves (b) for the single-wavelength SVM model.

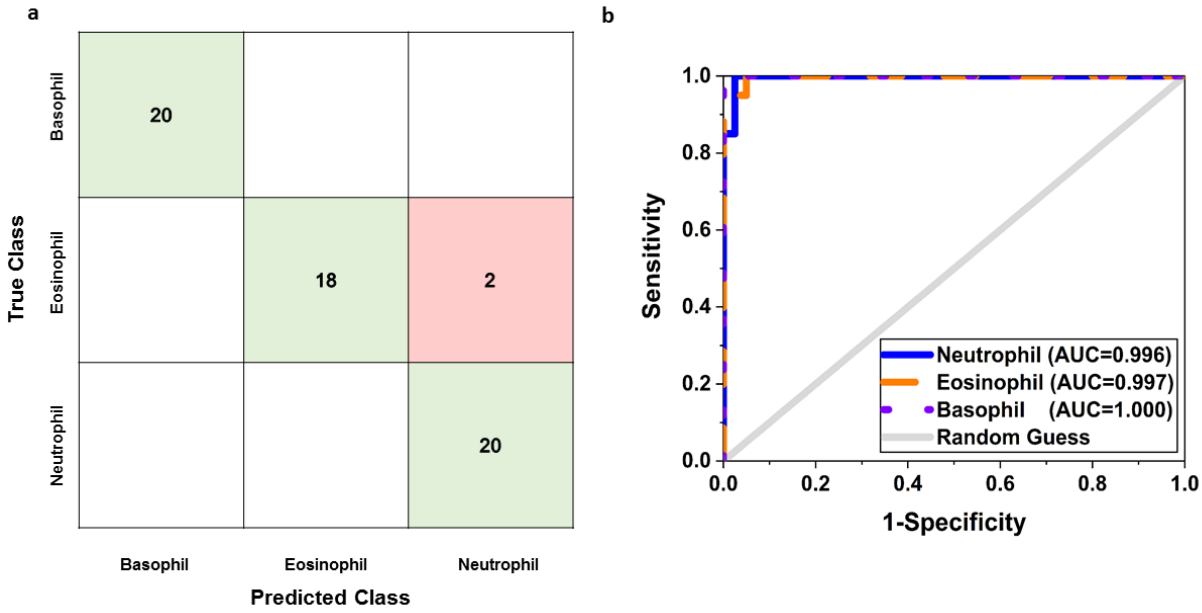

Fig. S4. The confusion matrix (a) and the ROC curves (b) for the SVM model based on the bright-field images of stained granulocytes.

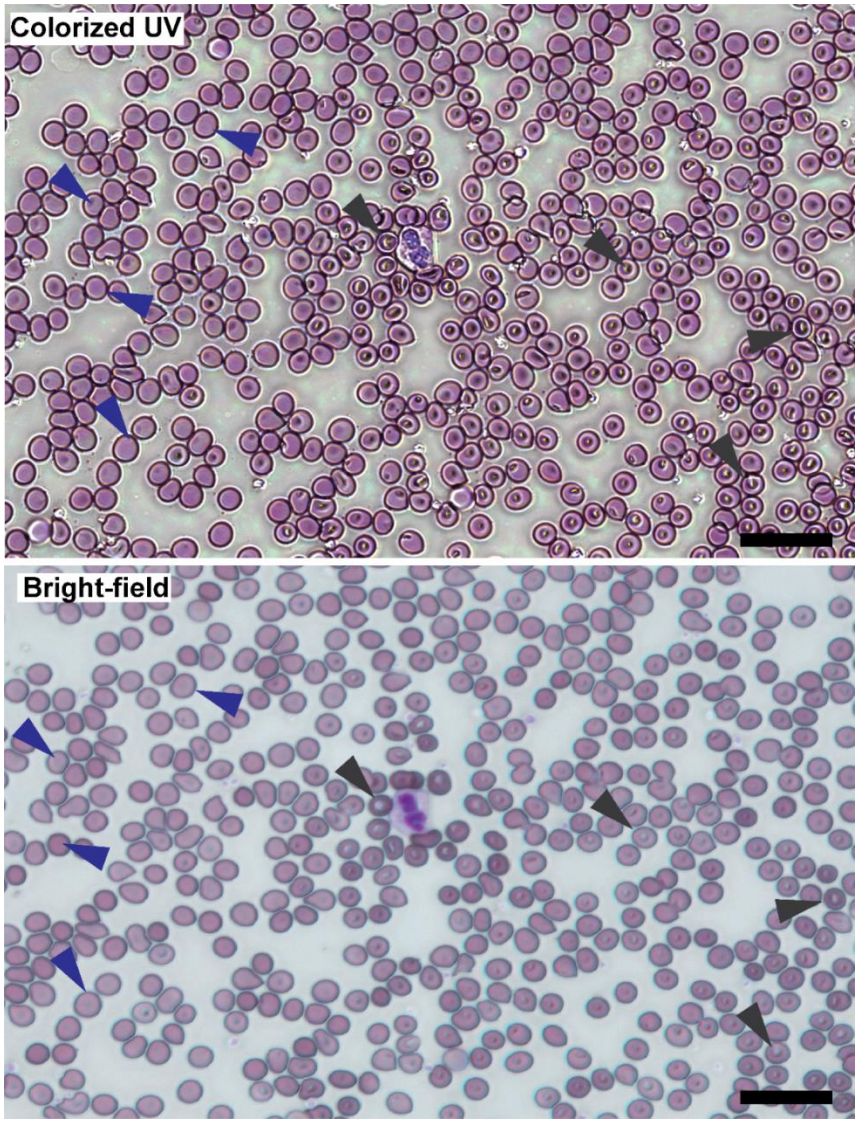

Fig. S5. Example UV colorized and bright-field images of the transition region from the monolayer to the feathered region. Here some RBC lose the typical biconcave RBC shape towards the end of this monolayer region of the smear. This is a common issue with blood smears which can result in misdiagnosis and confusion with other conditions, most typically spherocytosis (1). This issue can be mitigated by ensuring that imaging is performed within the central region of the monolayer area where RBCs maintain their natural shape. The figure shows RBCs with (gray arrowheads) and without (blue arrowheads) the biconcave shape. The scale bars are 30µm.

Table S1. The extracted features for multi- and single-wavelength model training

Feature	Description
Morphological	
Area	Area contained by cell contour
Perimeter	Cell contour perimeter
Eccentricity	The ratio of the distance between the foci of the ellipse and its major axis length. An ellipse whose eccentricity is 0 is a circle, while an ellipse whose eccentricity is 1 is a line segment.
Circularity	Calculated by $4\pi\times$(cell Area)/perimeter^2 . Circularity of a circle is 1.
Major/Minor axis length	Length of the major/minor axis of the ellipse that has the same normalized second central moments as the cell contour
Extent	Ratio of pixels in the cell contour to pixels in the total bounding box.
Solidity	Ratio of contour area to its convex hull area
Second Momentum	Euclidian distance between geometric centroid and each pixel averaged over the whole region.
Statistical	
Mean	Sum of pixel values over the total number of pixels in the region.
Skewness	Skewness is a measure of the asymmetry of the data around the sample mean. If skewness is negative, the data spreads out more to the left of the mean than to the right. If skewness is positive, the data spreads out more to the right. The skewness of the normal distribution (or any perfectly symmetric distribution) is zero.
Kurtosis	Kurtosis is a measure of how outlier-prone a distribution is. The kurtosis of the normal distribution is 3. Distributions that are more outlier-prone than the normal distribution have kurtosis greater than 3; distributions that are less outlier-prone have kurtosis less than 3.
Entropy	Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Defined as sum of the normalized histogram counts times their log.
Standard deviation	Square root of the difference between each pixel value and average pixel values squared over the number of pixels minus one
GLCM	
Contrast	Intensity contrast between a pixel and its neighbor based on GLCM of the image.

Correlation	Measure of how correlated a pixel is to its neighbor over the whole image. Correlation is 1 or -1 for a perfectly positively or negatively correlated image.
Energy	The sum of squared elements in the GLCM of the image. Energy is 1 for a constant image.
Homogeneity	Measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal. Homogeneity is 1 for an image with diagonal GLCM.
Mass	
Protein	Average protein mass (in fg) calculated based on the description in Methods section.
Nucleic acid	Average nucleic acid mass (in fg) calculated based on the description in Methods section.

SI References

1. J. Dacie, *Dacie and Lewis practical haematology* (2006) (September 24, 2019).