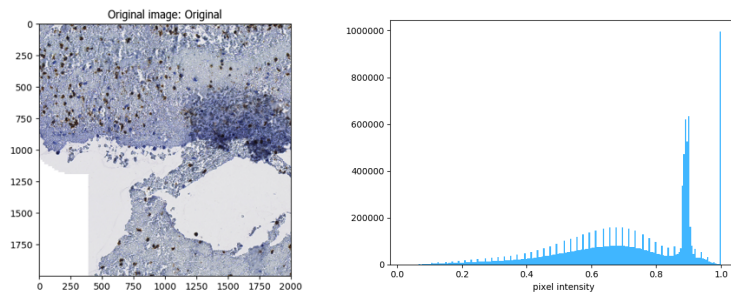
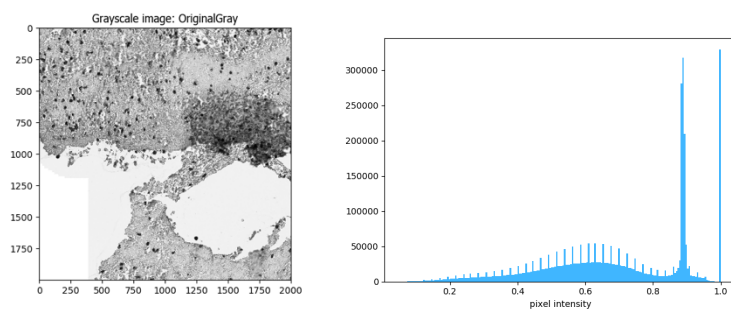


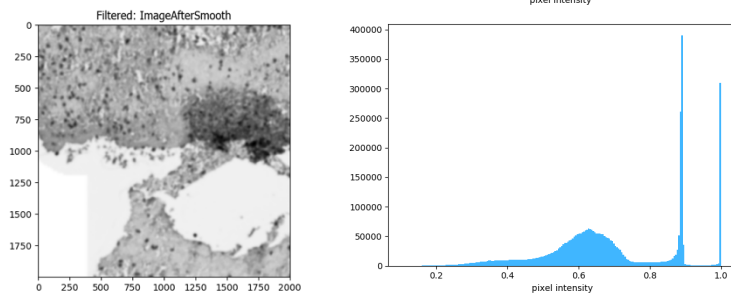
A.



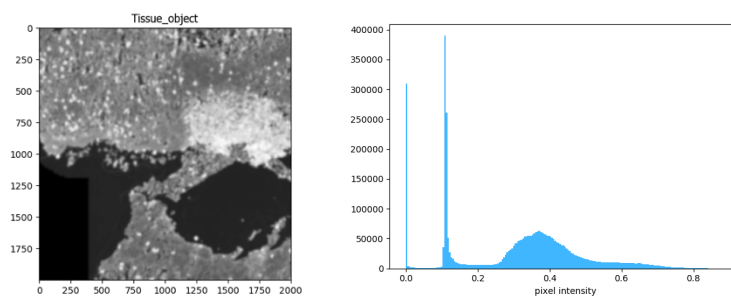
B.



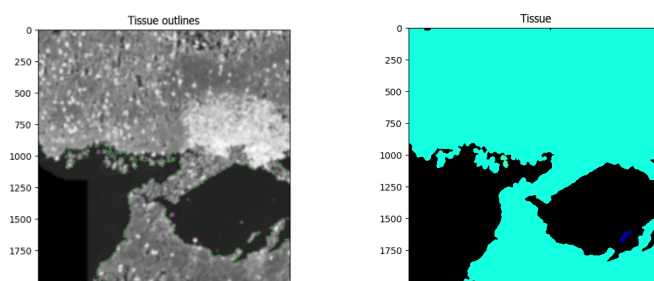
C.



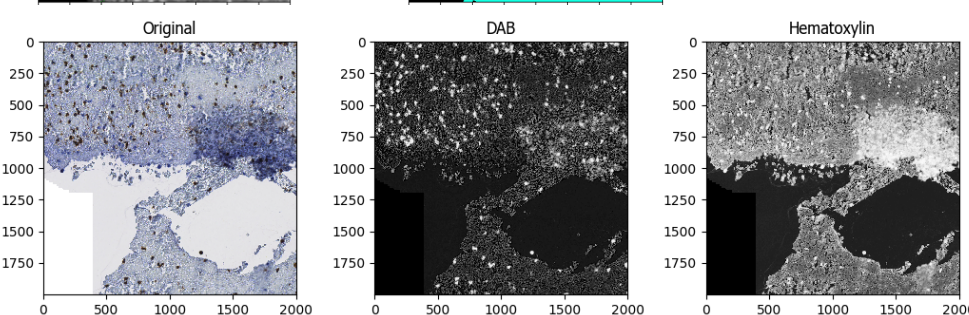
D.



E.



F.



### The CD66b CellProfiler pipeline workflow. (Part 1)

**[A].** The original image (left) is masked using PathProfiler Tissue Segmentation Unet and used as input by CellProfiler 4.1.3; the graph (right) shows the tonal distribution in the digital whole-slide image on a RGB scale.

**[B].** The input image is converted to a gray scaled image (left); the graph (right) shows the tonal distribution in the gray scaled image.

**[C].** A Gaussian filter is applied to smoothen the image and reduce image artefacts (artifact size 20 pixels) and noise (left); the graph (right) shows the tonal distribution after smoothening.

**[D].** The gray scaled image is inverted, *i.e.* non-tissue will become black (left); the graph (right) shows the tonal distribution after inverting.

**[E].** The tissue area is identified, as demarcated by the green line in the left image; the total tissue area size is calculated in pixels (right image) and tabulated (table).

**[F].** The colors, *i.e.* stains, are unmixed using the original image (left): DAB (middle), and Hematoxylin (HE, right).

**[G].** The red color channel is extracted from the input image (left); the graph (right) shows the tonal distribution of the extracted red channel image.

**[H].** The red channel image is inverted, *i.e.* non-tissue will become black (left); the graph (right) shows the tonal distribution after inverting.

**[I].** The DAB image (left) and red channel image (middle) are multiplied to enhance the saliency of CD66b positively stained cells (right). the graph (right) shows the tonal distribution of the enhanced CD66b stained image.

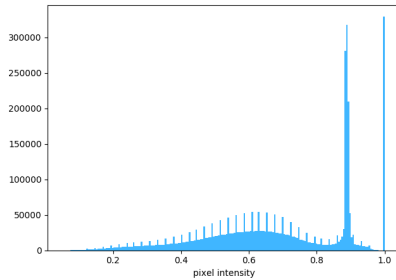
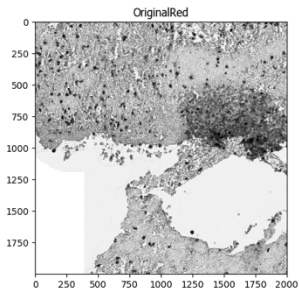
**[J].** The CD66b-positive objects, *i.e.* cells, are identified, white areas in the left image; the CD66b-positive objects are demarcated by a green line in the middle image, areas that are excluded due to size (minimal size 13 pixels) are demarcated in magenta; the right image shows all the identified CD66b-positive objects in random colors; the total number of identified objects is calculated and tabulated (table).

**[K].** The HE image (left) is masked by overlaying the identified CD66b-positive objects (right).

**[L].** The CD66b nuclei objects are identified, white areas in the left image; the positive nuclei objects are demarcated by a green line in the middle image, areas that are excluded due to size (minimal size 2 pixels) are demarcated in magenta; the right image shows all the identified positive nuclei objects in random colors; the total number of identified objects is calculated and tabulated (table).

# of accepted objects	3
10th pctlle diameter	68.7 pixels
Median diameter	74.4 pixels
90th pctlle diameter	1874.5 pixels
Area covered by objects	69.2 %
Thresholding filter size	1.0
Threshold	0.19

G.



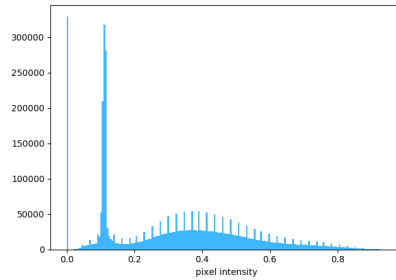
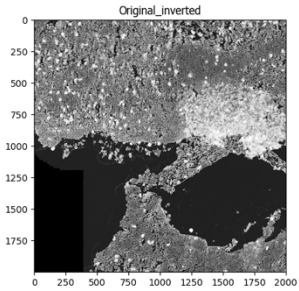
**The CD66b CellProfiler pipeline workflow. (Part 2)**

**[M].** The CD66b-positive identified objects are filtered using the identified nuclei objects. If a CD66b-positive object does not include a nuclei object, it is discarded.

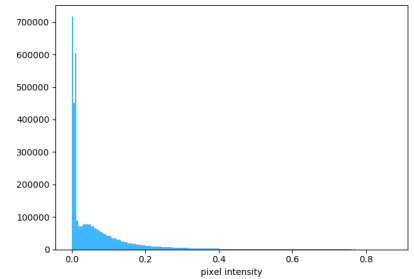
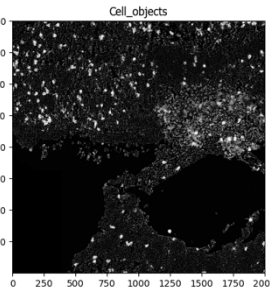
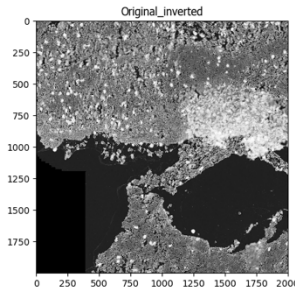
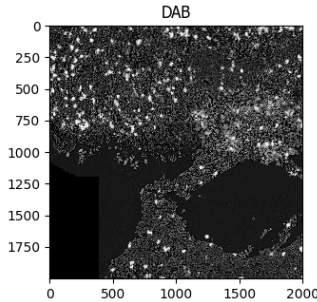
**[N].** Finally, the data for each tile are saved in a comma-separated table, including meta-data such as tile positions, image location, object counts (there could be multiple patches of stained areas or tissue). The original image (top-left) is used to overlay the CD66b-positive objects (using random colors, top-right). The tissue area (red), CD66b-positive objects (green), and CD66b nuclei objects (blue) are all demarcated in the bottom-left image. The table (bottom-right) shows the areas occupied by each object class.

Sample used: AE1531.UMC.CD66b.ndpi. [Tile= X22000, Y14000]

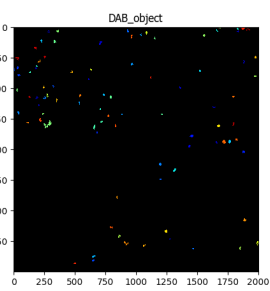
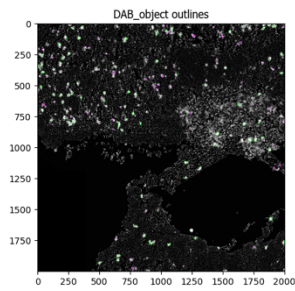
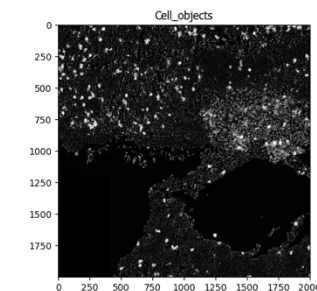
H.



I.

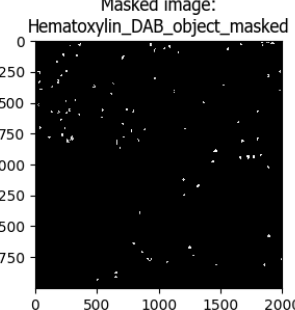
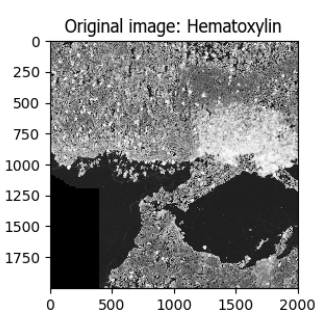


J.

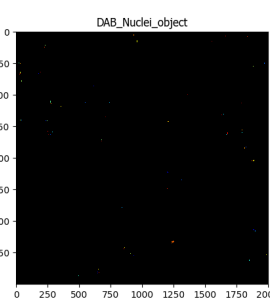
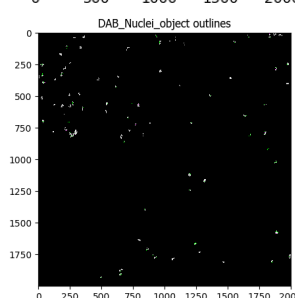
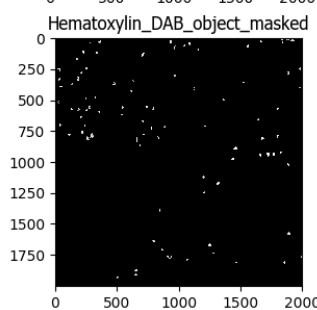


# of accepted objects	91
10th pctlile diameter	14.1 pixels
Median diameter	17.6 pixels
90th pctlile diameter	23.2 pixels
Area covered by objects	0.6 %
Thresholding filter size	1.0
Threshold	0.59

K.



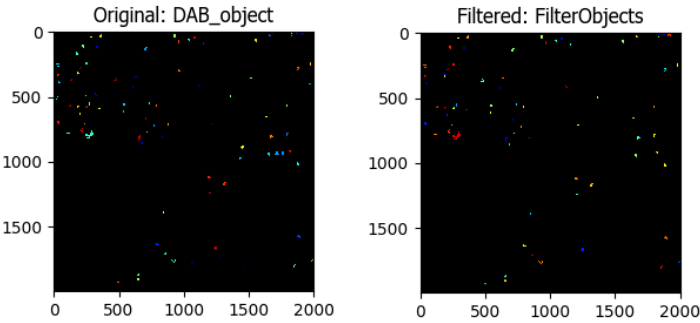
L.



# of accepted objects	163
10th pctlile diameter	2.3 pixels
Median diameter	3.6 pixels
90th pctlile diameter	6.4 pixels
Area covered by objects	0.1 %
Thresholding filter size	1.0
Threshold	0.805

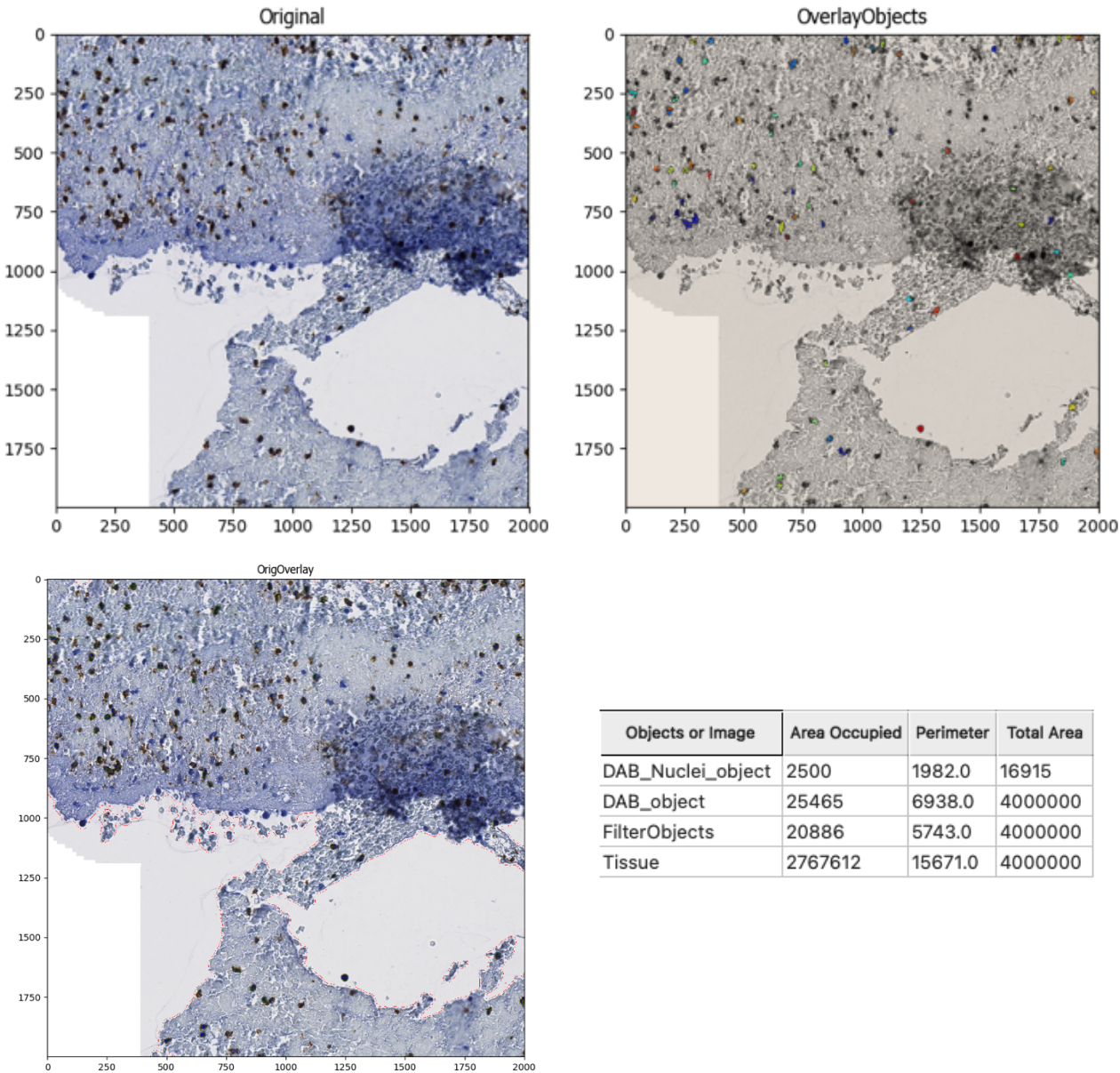


M.



Number of objects pre-filtering	91
Number of objects post-filtering	74

N.



Objects or Image	Area Occupied	Perimeter	Total Area
DAB_Nuclei_object	2500	1982.0	16915
DAB_object	25465	6938.0	4000000
FilterObjects	20886	5743.0	4000000
Tissue	2767612	15671.0	4000000