Large-scale extraction of interpretable features provides new insights into kidney histopathology – a proof-of-concept study

Laxmi Gupta[1,*]**, Barbara Mara Klinkhammer**[2]**, Claudia Seikrit**[2, 3]**, Nina Fan**[1]**, Nassim Bouteldja**[1,2]**, Philipp Gräbel**[1]**, Michael Gadermayr**[1,4]**, Peter Boor**[2]**, and Dorit Merhof**[1]

[1]Institute of Imaging & Computer Vision, RWTH Aachen University, Aachen, Germany
[2]Institute of Pathology, University Hospital Aachen, RWTH Aachen University, Aachen, Germany
[3]Division of Nephrology and Clinical Immunology, RWTH Aachen University, Aachen, Germany
[4]Salzburg University of Applied Sciences, Puch/Salzburg, Austria
[*]corresponding author: laxmi.gupta@lfb.rwth-aachen.de

A Supplementary Material
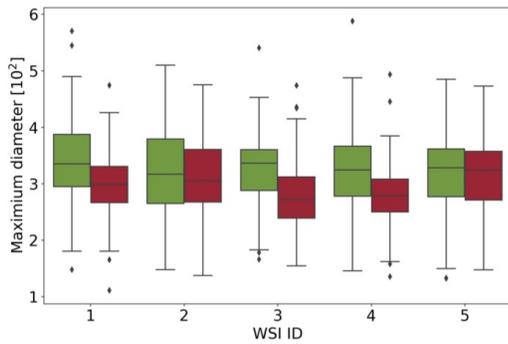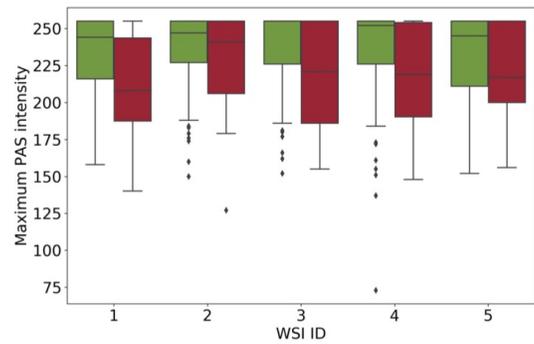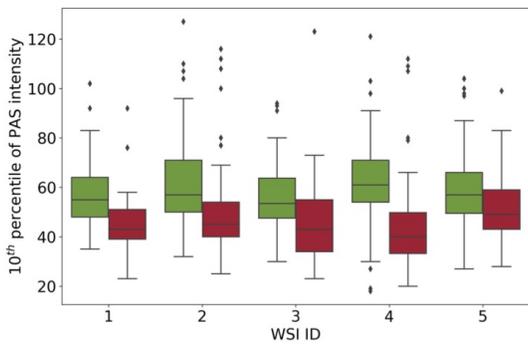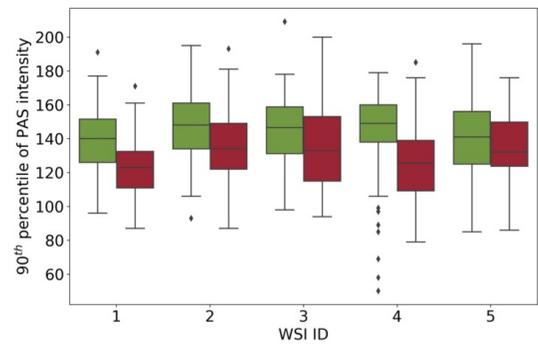
Supplementary Figure 1. Individual features per object. The X-axis shows the five WSIs and the corresponding feature values are shown on the Y-Axis, with green representing healthy and red, UUO. Examples shown include the BMC shape feature maximum diameter (a), glomeruli PAS intensity features, namely maximum intensity (b), intensity range (c), 10th percentile (d), 90th percentile (e), and the glomeruli color feature namely percentage of light PAS area.
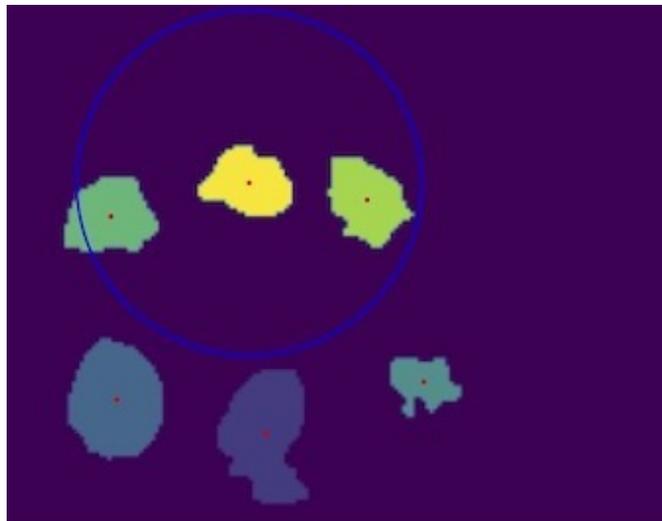
Supplementary Figure 2. Illustration of nuclei neighbour calculation. Nuclei neighbours are defined as the number of nuclei located within a certain radius of each nucleus as its neighbours. The radius is defined as two times the average of the lengths of the major axes of all the nuclei in a ROI.

Figure	Morphological feature	Equation	Description
	Area	-	Number of pixels within contour. Area is shown in green color in the figure.
	Perimeter	-	Length of contour. Perimeter is shown as green boundary in figure.
	Convex area	-	Number of pixels within convex contour. Convex area is shown in green color in the figure.
	Convex perimeter	-	Length of convex contour. Convex perimeter is shown as green boundary in figure.
	Minor axis	-	Length of minor axis of bounding ellipse. Minor axis is shown in blue color for the ellipse shown in green in figure.
	Major axis	-	Length of major axis of bounding ellipse. Major axis is shown in red color for the ellipse shown in green in figure.
	Solidity	$\frac{\text{area}}{\text{convex area}}$	Ratio of object area to its convex hull area. Describes to what extent the object is concave or convex.
	Convexity	$\frac{\text{perimeter}}{\text{convex perimeter}}$	Ratio of convex hull perimeter to the contour perimeter.
	Rectangularity	$\frac{\text{area}}{\text{bounding box area}}$	Ratio of object area to area of minimum bounding box. In the figure, the minimum bounding box is shown in green.
	Equivalent Diameter	$\sqrt{\frac{4 * \text{object area}}{\pi}}$	Diameter of the circle with the same area as the object.
	Compactness	$\frac{4 * \pi * \text{area}}{\text{perimeter}^2}$	Ratio between object area to area of a circle with same perimeter.
	Elongation	$1 - \frac{\text{minor axis}}{\text{major axis}}$	Describing the length of the object.
	Eccentricity	$\frac{\sqrt{\text{major axis}^2 - \text{minor axis}^2}}{\text{major axis}}$	The ratio of distance between major and minor axis. Takes values in [0,1].
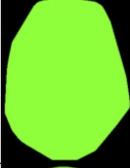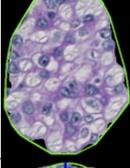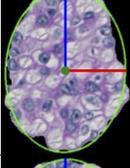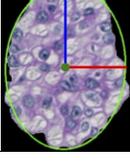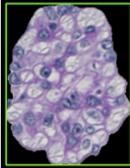
Supplementary Table 1. Definitions of morphological features

Color feature	Description
Color area	Number of pixels of a given color within the object
Color density	Color area divided by the object area

Nuclei feature	Description
Nuclei count	Number of nuclei within the object.
Nuclei density	Number of nuclei divided by the object area.
Mean number of neighbours	Mean value of the number of neighbours around each nucleus within the object.
Max number of neighbours	Maximum value of the number of neighbours per nucleus within the object.
Min number of neighbours	Minimum value of the number of neighbours per nucleus within the object.
Median number of neighbours	Mean value of the number of neighbours per nucleus within the object.
STD	Standard deviation of the distribution of nuclei neighbours within the object.

Supplementary Table 2. Definitions of color and nuclei-related features.

B Tables showing all features per object and the results after feature selection

Here, tables showing results per feature type are presented. The features that are significantly different in healthy and corresponding UUO WSIs are shown in purple. If no significant differences are found, features are highlighted in light blue. The number in each cell shows the trend for the feature for the WSI pair. If the median of the features from a healthy WSI is higher than that of the corresponding UUO WSI, it is marked as 1, otherwise as -1. The features that significantly different for at least 3 WSI pairs and follow the same trend for all five of them, are finally chosen after the feature selection step. These are outlined in black.

Morphological features

Object	WSI ID	Area	Perimeter	Convex Area	Convex Perimeter	Major Axis Length	Minor Axis Length	Solidity	Convexity	Rectangularity	Equivalent Diameter	Circularity	Elongation	Eccentricity
Glomerulus	1	1	1	1	1	1	1	-1	1	1	1	-1	-1	-1
	2	1	1	1	1	1	1	-1	1	-1	1	-1	-1	-1
	3	1	1	1	1	1	1	-1	1	1	1	-1	-1	-1
	4	1	1	1	1	1	1	-1	1	1	1	-1	1	1
	5	-1	-1	-1	-1	-1	-1	-1	1	-1	1	-1	1	1
BMC	1	-1	1	1	1	1	1	-1	1	-1	-1	-1	1	1
	2	1	1	1	1	1	1	-1	1	1	1	-1	1	1
	3	1	1	1	1	1	1	-1	1	1	1	-1	1	1
	4	1	1	1	1	1	1	-1	1	1	1	-1	1	1
	5	-1	1	-1	-1	-1	-1	-1	1	-1	-1	-1	1	1
Artery	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	2	1	1	1	1	1	1	-1	1	1	1	-1	1	1
	3	1	1	1	1	1	-1	-1	-1	1	1	-1	1	1
	4	1	1	1	1	1	-1	-1	1	-1	1	-1	1	1
	5	1	1	1	1	1	1	1	-1	1	1	1	1	1
Lumen	1	-1	-1	-1	-1	-1	-1	1	-1	-1	-1	1	-1	-1
	2	1	1	1	1	1	1	-1	1	1	1	-1	1	1
	3	-1	1	-1	1	-1	1	1	-1	-1	-1	-1	1	1
	4	1	1	1	1	1	1	-1	1	1	1	-1	-1	-1
	5	-1	-1	-1	-1	-1	-1	1	-1	-1	-1	1	1	1
Tubule	1	1	1	1	1	1	1	1	-1	1	1	1	-1	-1
	2	1	1	1	1	1	1	1	1	1	1	-1	1	1
	3	1	1	1	1	1	1	1	-1	1	1	1	1	1
	4	1	1	1	1	1	1	-1	1	1	1	-1	1	1
	5	1	1	1	1	1	1	1	1	1	1	-1	1	1
Interstitium	1	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	2	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	3	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	4	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
	5	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA

Supplementary Figure 3. Morphological features

Shape features

Object	WSI ID	Elongation	Major Axis Length	Maximum Diameter	Mesh Surface Area	Minor Axis Length	Perimeter	Perimeter Surface Ratio	Pixel Surface Area	Sphericity
Glomerulus	1	1	1	1	1	1	1	-1	1	-1
	2	1	1	1	1	1	1	1	1	-1
	3	1	1	1	1	1	1	-1	1	-1
	4	-1	1	1	1	1	1	-1	1	-1
	5	-1	-1	-1	-1	-1	-1	1	-1	-1
BMC	1	-1	1	1	1	1	1	-1	1	-1
	2	-1	-1	1	1	1	1	-1	1	1
	3	1	1	1	1	1	1	-1	1	1
	4	1	1	1	1	1	1	-1	1	1
	5	-1	-1	1	1	-1	-1	-1	1	1
Artery	1	-1	1	1	1	1	-1	-1	1	1
	2	-1	1	1	1	1	1	1	1	-1
	3	-1	1	1	1	-1	1	-1	1	1
	4	-1	1	1	1	-1	1	1	1	-1
	5	-1	1	1	1	-1	1	-1	1	1
Lumen	1	1	-1	-1	-1	1	-1	-1	-1	1
	2	1	1	1	1	1	1	-1	1	-1
	3	-1	1	-1	-1	-1	-1	1	-1	-1
	4	-1	1	1	1	1	1	1	1	-1
	5	1	-1	-1	-1	-1	-1	1	-1	-1
Tubule	1	1	1	1	1	1	1	-1	1	1
	2	-1	1	1	1	1	1	-1	1	-1
	3	-1	1	1	1	1	1	-1	1	1
	4	-1	1	1	1	1	1	-1	1	-1
	5	-1	1	1	1	1	1	-1	1	-1
Interstitium	1	NA	NA	NA	NA	NA	NA	NA	NA	NA
	2	NA	NA	NA	NA	NA	NA	NA	NA	NA
	3	NA	NA	NA	NA	NA	NA	NA	NA	NA
	4	NA	NA	NA	NA	NA	NA	NA	NA	NA
	5	NA	NA	NA	NA	NA	NA	NA	NA	NA

Supplementary Figure 4. Shape features

Color features

Object	WSI ID	Haemmatoxylin Area	Light Magenta Area	Dark PAS Area	Unstained Area	Hamatoxylin Area Percentage	Light Magenta Area Percentage	Dark PAS Area Percentage	Unstained Area Percentage
Glomerulus	1	1	1	1	1	-1	1	-1	1
	2	-1	1	1	-1	-1	1	1	-1
	3	1	1	1	-1	1	1	1	-1
	4	-1	1	1	1	-1	1	1	1
	5	1	-1	-1	-1	1	1	1	-1
BMC	1	1	1	1	-1	-1	1	1	-1
	2	1	1	1	-1	-1	1	1	-1
	3	1	1	1	1	-1	1	-1	-1
	4	1	1	1	1	-1	1	-1	-1
	5	-1	1	1	-1	1	1	-1	-1
Artery	1	-1	1	-1	1	-1	1	-1	1
	2	-1	1	-1	1	-1	1	-1	1
	3	1	1	1	-1	1	1	-1	-1
	4	-1	1	-1	1	-1	1	-1	1
	5	1	1	1	1	-1	1	-1	1
Lumen	1	1	-1	-1	-1	1	-1	-1	1
	2	1	1	1	1	1	1	-1	-1
	3	1	-1	-1	-1	1	-1	-1	1
	4	1	-1	-1	1	-1	-1	-1	1
	5	1	-1	1	-1	1	1	-1	-1
Tubule	1	1	1	1	1	-1	1	-1	-1
	2	-1	1	1	-1	-1	1	-1	-1
	3	1	1	1	-1	-1	1	-1	-1
	4	-1	1	1	1	-1	1	-1	-1
	5	-1	1	1	-1	-1	1	-1	-1
Interstitium	1	-1	-1	-1	-1	-1	-1	-1	1
	2	-1	-1	-1	-1	-1	-1	-1	1
	3	-1	-1	-1	-1	-1	-1	-1	1
	4	-1	-1	-1	1	-1	-1	-1	1
	5	-1	-1	-1	-1	1	-1	-1	1

Supplementary Figure 5. Color features

Nuclei features

Object	WSI ID	Nuclei Count	Nuclei Density	Nuclei Neighbour Standard Deviation	Maximum Nuclei Neighbour	Minimum Nuclei Neighbour	Mean Nuclei Neighbour	Median Nuclei Neighbour
Glomerulus	1	1	-1	-1	1	-1	-1	-1
	2	1	-1	1	1	1	-1	-1
	3	1	-1	-1	1	1	-1	1
	4	-1	-1	-1	-1	-1	-1	-1
	5	1	1	1	1	-1	1	1
BMC	1	-1	-1	-1	-1	1	-1	-1
	2	-1	-1	-1	-1	1	-1	1
	3	1	-1	-1	-1	1	-1	-1
	4	-1	-1	-1	-1	1	-1	-1
	5	-1	-1	-1	1	1	-1	-1
Artery	1	NA	NA	NA	NA	NA	NA	NA
	2	NA	NA	NA	NA	NA	NA	NA
	3	NA	NA	NA	NA	NA	NA	NA
	4	NA	NA	NA	NA	NA	NA	NA
	5	NA	NA	NA	NA	NA	NA	NA
Lumen	1	NA	NA	NA	NA	NA	NA	NA
	2	NA	NA	NA	NA	NA	NA	NA
	3	NA	NA	NA	NA	NA	NA	NA
	4	NA	NA	NA	NA	NA	NA	NA
	5	NA	NA	NA	NA	NA	NA	NA
Tubule	1	1	-1	-1	-1	-1	-1	1
	2	-1	-1	-1	-1	-1	-1	1
	3	1	-1	1	1	1	-1	1
	4	1	-1	-1	1	-1	-1	-1
	5	-1	-1	-1	1	-1	-1	1
Interstitium	1	-1	-1	-1	-1	1	-1	-1
	2	-1	-1	-1	-1	1	-1	-1
	3	-1	1	-1	-1	1	-1	-1
	4	-1	-1	-1	-1	1	-1	-1
	5	-1	1	-1	-1	1	-1	-1

Supplementary Figure 6. Nuclei features

Supplementary Figure 7. Intensity features

Supplementary Figure 8. GLCM-based texture features

GLRLM-based texture features
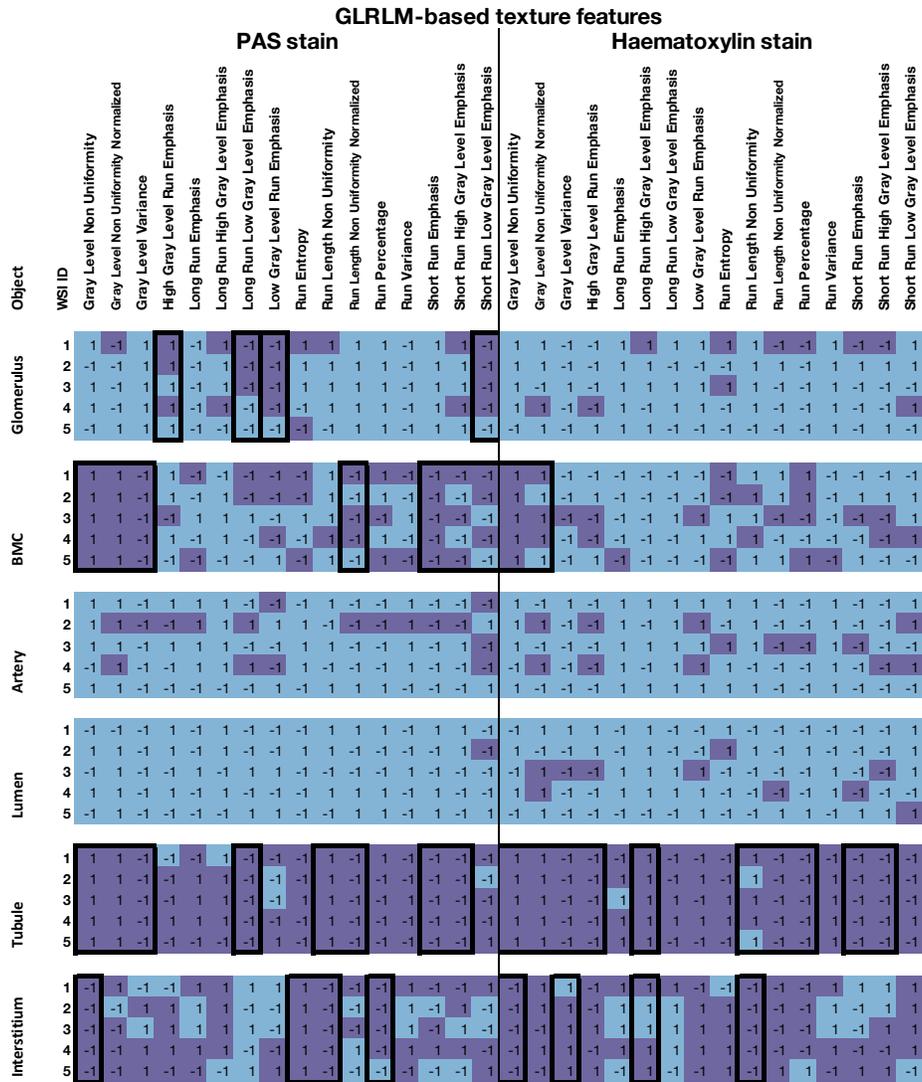
Supplementary Figure 9. GLRLM-based texture features

GLSZM-based texture features

PAS stain　　　　**Haematoxylin stain**

Object

WSI ID

(Feature columns, both PAS stain and Haematoxylin stain:)
Gray Level Non Uniformity, Gray Level Non Uniformity Normalized, Gray Level Variance, High Gray Level Zone Emphasis, Large Area Emphasis, Large Area High Gray Level Emphasis, Large Area Low Gray Level Emphasis, Low Gray Level Zone Emphasis, Size Zone Non Uniformity, Size Zone Non Uniformity Normalized, Small Area Emphasis, Small Area High Gray Level Emphasis, Small Area Low Gray Level Emphasis, Zone Entropy, Zone Percentage, Zone Variance

Row groups: Glomerulus (1–5), BMC (1–5), Artery (1–5), Lumen (1–5), Tubule (1–5), Interstitium (1–5)

Supplementary Figure 10. GLSZM-based texture features

Supplementary Figure 11. GLDM-based texture features

eltab:gldm$_f eatures$

NGTDM-based texture features

Object	WSI ID	PAS stain					Haematoxylin stain				
		Busyness	Coarseness	Complexity	Contrast	Strength	Busyness	Coarseness	Complexity	Contrast	Strength
Glomerulus	1	-1	-1	1	-1	1	1	-1	-1	-1	-1
	2	-1	-1	1	-1	1	1	-1	1	-1	-1
	3	1	-1	1	-1	-1	1	-1	-1	-1	-1
	4	-1	-1	1	-1	-1	1	-1	1	-1	-1
	5	-1	1	1	-1	1	-1	1	1	-1	1
BMC	1	1	-1	1	1	-1	1	-1	1	-1	-1
	2	1	-1	-1	-1	-1	1	-1	1	-1	-1
	3	1	-1	-1	-1	-1	1	-1	1	-1	-1
	4	1	-1	-1	-1	-1	1	-1	1	-1	-1
	5	1	-1	1	1	-1	1	-1	1	1	-1
Artery	1	1	-1	1	-1	-1	1	-1	1	-1	1
	2	1	-1	-1	-1	-1	1	-1	-1	1	-1
	3	1	-1	-1	-1	-1	1	-1	-1	-1	-1
	4	1	-1	-1	-1	-1	1	1	-1	1	1
	5	1	-1	-1	1	-1	1	-1	-1	1	-1
Lumen	1	-1	1	1	1	-1	-1	1	-1	1	-1
	2	1	-1	1	-1	-1	1	-1	1	1	-1
	3	1	1	-1	-1	-1	1	1	-1	-1	-1
	4	1	-1	-1	-1	-1	1	-1	-1	-1	-1
	5	-1	1	-1	1	1	-1	1	-1	1	1
Tubule	1	1	-1	1	-1	-1	1	-1	-1	-1	-1
	2	1	-1	-1	-1	-1	1	-1	-1	-1	-1
	3	1	-1	-1	-1	-1	1	-1	-1	-1	-1
	4	1	-1	-1	-1	-1	1	-1	-1	-1	-1
	5	1	-1	-1	1	-1	1	-1	-1	-1	-1
Interstitium	1	-1	1	-1	1	1	-1	1	-1	-1	1
	2	-1	1	-1	-1	1	-1	1	-1	-1	1
	3	-1	1	-1	1	1	-1	1	-1	1	1
	4	-1	1	-1	1	1	-1	1	-1	-1	1
	5	-1	1	-1	1	1	-1	1	-1	1	1

Supplementary Figure 12. NGTDM-based texture features