



**HAL**  
open science

## Detection of glomeruli in renal pathology by mutual comparison of multiple staining modalities

M. Temerinac-Ott, Germain Forestier, J. Schmitz, M. Hermsen, J.H. Brasen,  
F. Feuerhake, C. Wemmert

► **To cite this version:**

M. Temerinac-Ott, Germain Forestier, J. Schmitz, M. Hermsen, J.H. Brasen, et al.. Detection of glomeruli in renal pathology by mutual comparison of multiple staining modalities. 2017 10th International Symposium on Image and Signal Processing and Analysis (ISPA), Sep 2017, Ljubljana, France. pp.19-24, 10.1109/ispa.2017.8073562 . hal-03807632

**HAL Id: hal-03807632**

**<https://hal.science/hal-03807632>**

Submitted on 12 Oct 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Detection of glomeruli in renal pathology by mutual comparison of multiple staining modalities

M. Temerinac-Ott<sup>1</sup>, G. Forestier<sup>2</sup>, J. Schmitz<sup>3</sup>, M. Hermsen<sup>4</sup>, J.H. Bräsen<sup>3</sup>, F. Feuerhake<sup>3</sup> and C. Wemmert<sup>1</sup>

<sup>1</sup>ICube, University of Strasbourg, France

<sup>2</sup>MIPS, University of Haute Alsace, France

<sup>3</sup>RUMC, Radboud University Medical Center, Nijmegen, The Netherlands

<sup>4</sup>Institute for Pathology, MHH, Hannover Medical School, Germany

**Abstract**—We evaluate the detection of glomerular structures in whole slide images (WSIs) of histopathological slides stained with multiple histochemical and immuno-histochemical staining using a convolutional neural network (CNN) based approach. We mutually compare the CNN performance on different stainings (Jones H&E, PAS, Sirius Red and CD10) and we present a novel approach to improve glomeruli detection on one staining by taking into account the classification results from differently stained consecutive sections of the same tissue. Using this integrative approach, the detection rate (F1-score) on a single stain can be improved by up to 30%.

## I. INTRODUCTION

The detection of glomeruli plays an important role in renal biopsy evaluation. As an example, for grading of biopsies after kidney transplantation according to the Banff classification scheme [1] the minimal glomeruli number is seven (marginally representative) or more. Ten glomeruli are adequate and representative. Further examples for the relevance of detecting glomeruli include quantification of morphological changes (e.g., sclerotic glomeruli), vascular pathology, and infiltration by immune cells. Most steps of biopsy evaluation are performed in so-called "routine" staining protocols (e.g., H&E, PAS), and various specific immunohistochemistry (IHC), or histochemistry staining methods are used to delineate different additional relevant structures, such as integrity of the basal lamina (e.g., PAS, Jones), complement expression on vascular endothelium (e.g. C4d), and many others.

The digitization of microscopical slides to whole slide images (WSI) has led to a new research field often referred to as "digital pathology", broadening the technological capabilities and adding to the conventional biopsy evaluation, where tissue sections on glass slides are visually analyzed by trained pathologists using a light microscope. The emerging new digital methods enable storage and analysis of image-based data beyond visual inspection, complementing established workflows, and increasing efficiency in some aspects, such as high-throughput quantification. However, many questions remain open on how to transfer formalized expert knowledge to computer systems and how to increase objectivity and reproducibility of WSI analysis [2], [3].

Formalin-fixed and paraffin-embedded (FFPE) renal tissue was serially sectioned, resulting in 2-3 micrometers thick sections. The renal tissue analyzed for this study was obtained

from kidneys which were taken out due to a tumor (nephrectomies) and small needle biopsies. The tissue parts used from the nephrectomies were tumor-free and showed normal morphology. Due to their big size and thus high numbers of glomeruli (in relation to the biopsies) these nephrectomy samples were used as a first training set for the detection of glomeruli.

The prepared sections from biopsies or from nephrectomies can be stained with different markers for certain cell types like T-cells, B-cells, macrophages or certain tissue or cell subcompartments. Standard staining techniques in the clinical routine are usually limited to one or two IHC markers on one section. This, and the different histochemical staining modalities that highlight distinct compartments, requires evaluation of consecutive sections from the same tissue. The alignment of the different markers is "computed" in the brain of the pathologists while viewing one slide after the other under the microscope, focusing on diagnostically relevant regions. Automatic detection of glomeruli might complement diagnostic procedures by counting normal and sclerotic glomeruli or evaluating immune cell infiltration.

Previous published work on detection of glomeruli in mouse renal samples [4] was done using Histogram of Oriented Gradients (HOG) features and Support Vector Machine (SVM) to first detect windows containing glomeruli candidates. A segmentation algorithm was then applied to extract the precise boundaries of the detected glomeruli. An approach for detecting glomeruli using Fitzgibbon's ellipse fitting method in human tissue was presented in [5] and applied to WSIs of sections stained with the histochemical Masson's trichrome. A k-means classification based approach for glomeruli detection on CD68 and CD34 double stainings was presented in [6]. Deep Learning showed superior results for detecting tumors in WSI [7] and was recently successfully applied to automated detection of specific regions of interest in WSI of breast tissue, specifically lobular areas in the human mammary gland [8].

In this paper, we address three questions that are important for detecting glomeruli: (1) Which of the currently existing methods is best suitable for detecting glomeruli in WSI? (2) What differences can be found for the detection in different stainings? (3) How can we mutually combine the information from multiple stainings to achieve better detection results?

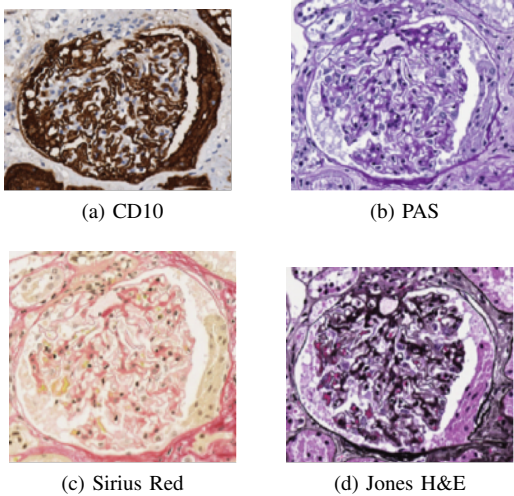


Fig. 1: Same glomerulus in consecutive section, visualized with different stainings.

## II. GLOMERULI DETECTION

### A. Context

A glomerulus consists of a ball-like network of specialized capillaries, representing the filter of the functional renal unit called nephron. The filtrated primary urine is collected in the surrounding cup-like structure named bowman’s capsule and transferred to the tubular apparatus, which concentrates the filtrate from up to 200 liters of primary urine to 1.5 liters of final urine. In Fig. 1 we show how the same glomerulus is represented in several serial sections stained with different staining modalities. While the whole object is delineated, it is obvious that different tissue components are highlighted: CD10 stains the glomerular podocytes, Bowman’s epithelium and proximal tubules; Jones silver and PAS stain glomerular basement membranes, and Sirius Red stains connective tissue (collagen). Therefore, the added information obtained by different techniques is superior to single staining, and beyond visually detecting glomeruli some diagnostic tasks require to combine information from multiple modalities.

Reliable detection of glomeruli in different staining modalities is also relevant for another aspect: Sequential sections of the same 3-dimensional structure may provide additional important information that is lost when the nearly 2-dimensional, only few microns thick sections are evaluated one by one. For example, loss or artificial changes of the morphology (*e.g.*, due to tissue processing or sectioning) can be compensated, if the same glomerular structure is better represented in neighboring cutting levels.

### B. HOG features classifier

Histogram of Oriented Gradients (HOG) [9] is a powerful and fast tool for detecting objects in images. It has already been applied in brightfield pathological images analysis for cell detection of human prostate cancer cells [10].

As a preprocessing step we use morphological operations to separate the tissue from the white background. The training is performed using patches from the tissue only. HOG features are extracted on the annotated glomeruli, using the VLFeat library and the HOG cell size of 8 on rescaled glomeruli images of size  $64 \times 64 \times 3$ . A SVM model is trained using the positive glomeruli examples and applied to all the pixels from the tissue using the sliding window approach, thereby producing a score for each pixel of the tissue. The detection is computed for multiple scales ( $2^{-1}, 2^{-0.75}, \dots, 2^{1.75}, 2^2$ ) of the image. We use hard negative mining to select the negative examples and iterate the training of the SVM model for seven rounds. For testing, the learned SVM model is applied to a new image and the top 200 detections with the highest score are returned.

### C. CNN classifier

Deep learning has proven its efficiency in WSI analysis [7]. A Convolutional Neural Network (CNN) [11] has the ability to learn a hierarchical description of visual patterns from a set of annotated examples, and then make accurate predictions for new visual inputs. By combining patch-based image analysis with CNNs, it is possible to automatically detect glomeruli in renal histological images.

For patch-based image analysis, we assume that it is possible to predict the class of a pixel by observing its neighboring pixels. We settled on a square neighborhood of size  $128 \mu\text{m}$  which is enough to cover the major part of the cross-section of a glomerulus with its surroundings (glomeruli are approximately  $0.2 \text{ mm}$  in diameter; size of glomeruli in sections firstly depends on cutting plane through the three-dimensional ball structure, secondly shrinkage due to tissue fixation leads to smaller diameter in the sections). It is then possible to generate, for each pixel, a description of the patch centered on this pixel (*i.e.* a square RGB patch described by  $32 \times 32 \times 3$  values). The resolution of the patch is  $4 \mu\text{m}/\text{pixel}$ . In images annotated by a pathologist, each patch can be considered *positive* or *negative* according to its central pixel: if a pixel is in an annotated glomerulus, then its patch is positive, otherwise it is negative.

Our CNN is a function that can predict the class of a patch (positive or negative) based on its featureless description. The associated deep network architecture (Table I) is designed to perform multiclass classification of small RGB images [12], which is conveniently similar to the task of patch classification. We initialize the network using random weights drawn from a zero-mean Gaussian distribution.

### D. Combining image information from multiple stainings by mutual comparison

In order to confirm the detection of glomeruli in consecutive sections of the same tissue, which were stained using different stainings, we propose to first co-register the sections using object based registration. The results from the CNN are converted to objects by thresholding followed by connected component analysis. The values returned by the CNN are in

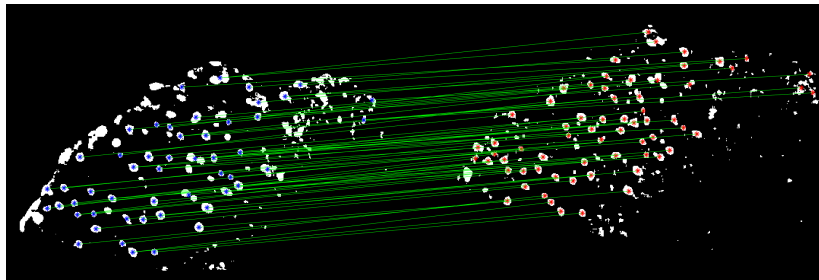


Fig. 2: Registered objects from CNN predictions for PAS and CD10 for patient 12.

Input	$32 \times 32 \times 3$
Convolution layer	kernels: 32; size: $5 \times 5$ ; padding: 2
Max-pooling layer	size: $3 \times 3$ ; stride: 2
Activation layer	Rectified Linear Unit (ReLU)
Convolution layer	kernels: 32; size: $5 \times 5$ ; padding: 2
Activation layer	ReLU
Average-pooling layer	size: $3 \times 3$ ; stride: 2
Convolution layer	kernels: 64; size: $5 \times 5$ ; padding: 2
Activation layer	ReLU
Average-pooling layer	size: $3 \times 3$ ; stride: 2
Fully connected layer	neurons: 64
Fully connected layer	neurons: 2

TABLE I: Network architecture adapted from *CIFAR10 Quick* example to two classes.

[0, 1], the threshold value is 0.8. We remove objects that are too small or too big (detected object area  $< 40\mu m$  or detected object area  $> 400\mu m$ ). Each detected object in the image is represented by the coordinates of its centroid. Therefore, the registration between stains reduces to the registration between point sets. We use rotation and translation invariant features [13] to describe the neighborhood of each point. Objects are matched between images, if their point descriptors agree (Fig. 2). This rigid registration step does not account for deformations due to preparation and sectioning of the tissue. It provides a first rough matching between the same objects in the different stainings.

All detected objects are given a global and a local identity and are stored in a table. Common objects share the same global identity, while the local identity is unique for each individual object. In Fig. 3,  $i$  is the global identity of the object and  $id$  is the local identity. If the local identity is 0, this means, that no corresponding object is found, however it is possible to compute the position of the potential correspondence by applying the computed transform from the registration to the coordinates of the common object in the reference stain.

In order to confirm detections using the information from different stainings, we apply two mutual comparison steps:

- **Mutual comparison 1:** We keep only the detections which are supported by at least two other stainings (majority vote). Using this consensus step, we can only reduce the number of false positives (FP), however not increase the number of true positives (TP) (See Fig. 4 (a,b)).
- **Mutual comparison 2:** We add objects that were initially

not detected as matching objects, but are in a radius of the joint detections suggested from the three other stainings. For this step we consider also objects having classification score  $> 0.5$ . From the computed registration parameters we know where the object should be in the staining and we search for a detected object in the radius of  $40\mu m$  around the suggested position. If a detected object is found in that radius, it is added as a confirmed object to the set of detected glomeruli in that staining (See Fig. 3).

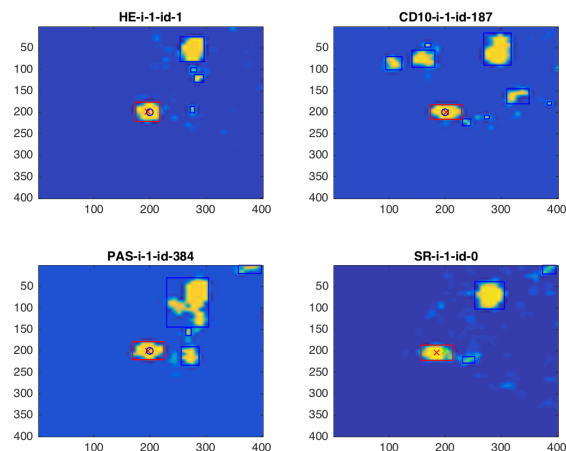
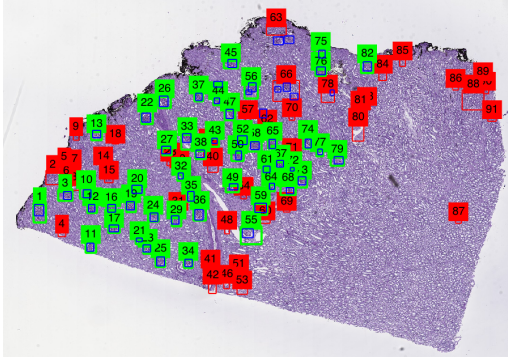


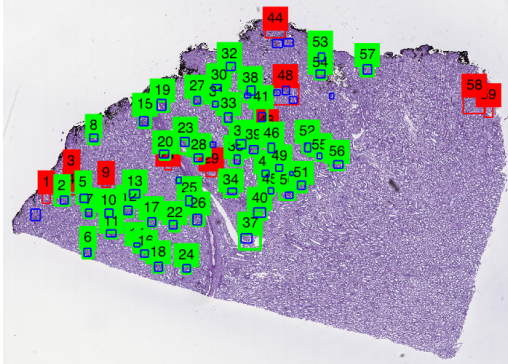
Fig. 3: Added object from Mutual comparison 2: The object in the SR staining was not found as a corresponding object using rigid registration. However, after searching in a small radius around the suggested position from the other three correspondences, it was added as a corresponding object. blue boxes: All detected objects (CNN score  $> 0.5$ ), red box: corresponding objects, green cross: suggested position of the registered object, red cross: center of detected corresponding object, blue circle: center of corresponding objects detected using Mutual comparison 1.

### III. EXPERIMENTS AND RESULTS

In order to determine which method is best suitable for glomeruli detection on our dataset, we applied the CNN and HOG classifiers on each of the four stainings individually performing a three-fold cross-validation for the first three



(a) CNN object detections.



(b) Consensus of CNN detections after Mutual comparison 1.

Fig. 4: (a) Detections of the CNN without comparison (green box: TP, red box: FP, blue box: ground truth) (b) After combining the predictions by using Mutual comparison 1 the number of FP decreases.

patients (Table III). Since the CNN performed better in our experiments, we chose to apply it to the larger data set of six patients and evaluate the comparative detection using the CNN results (Table IV). For the CNN, we randomly chose 5,000 positive patches and 10,000 negative patches from each image in the training set.

The accuracy is calculated on object detection only (not on the pixels), since we want to count the number of detected glomeruli and analyze intact objects in the further processing steps. A true positive (TP) is found if the bounding box of the detected object and the bounding box of the annotated glomerulus have an overlap of  $> 20\%$ . We use the F1-score to compute accuracy:

$$F1 = 2 \cdot TP / (2 \cdot TP + FP + FN) \quad (1)$$

where FP are false positives and FN false negatives.

#### A. Description of the data

To validate this work, we collected FFPE tissue blocks derived from renal tissue that can be considered normal. We used kidney tissue obtained from six tumor nephrectomies, selecting specimens with the largest possible distance from

the tumor. Four of those tumor-free sections per patient were selected for digitization. WSIs were acquired by an Aperio AT2 scanner at 40X magnification in the proprietary .svs format, with a resolution of  $0.253 \mu\text{m}/\text{pixel}$ . For the formal comparison, four different stainings were selected: Jones H&E, CD10, Sirius Red and PAS. Jones H&E is a combination of the H&E common stain used to delineate the nucleus and the cytoplasm of the cells, and methenamine which stains basement membrane material (in the peripheral capillary wall as well as mesangium). CD10 is present on specific cell types such as podocytes and proximal tubular epithelium (see above). For each WSI outlines were drawn over the digital images to annotate each glomerulus using the software tools ASAP and Cytomine[14] and subsequently verified by an experienced nephropathologist (J.H.B.). Table II shows the size of each WSI and summarizes the number of annotated glomeruli for each case and each staining.

#### B. Results on nephrectomies

Table III shows the F1-scores (three-fold cross validation) for all methods for each individual image. The HOG detector performed by 10–20% worse than the CNN. Only in one case P11-PAS, it showed a slightly better performance. The CNN results for Jones H&E, CD10 and PAS are on average 70%, while Sirius Red has only 52%. There is a large variation in the classification results between the individual patients, that can be explained by the heterogeneity of the tissue samples. Patient 10 with the largest number of glomeruli has a score greater than 70% for all stainings and performed by 20% higher than the other patients for Jones H&E, CD10 and Sirius Red staining.

Table IV shows the F1-scores (six-fold cross validation) for the CNN based methods for each individual image. By using three more patients for training the individual CNN accuracies could be significantly improved. On average, best detection results could be achieved on Jones H&E staining (81.32%), while detection on PAS performed the worst on average (66.38%). After applying Mutual comparison 1, the F1-score can decrease, since potentially true positives are removed if they are not supported by at least two other stainings. Mutual comparison 2 always increases the score with respect to Mutual comparison 1. The final score after applying Mutual comparison 2 is higher for almost all of the stainings and patients. On average the F1-score increases for all patients and the standard deviation decreases after applying the comparative method.

Further steps in the comparative process should include: the separation of adjacent glomeruli, which are detected as one big object. These appear frequently in the nephrectomies, when two glomeruli are very close to each other (Fig. 5). Since the detection of two glomeruli as one object is too big, it is discarded (false negative) or considered as a false positive (the overlap between ground truth box and detected box is too small). Furthermore, it can happen that a glomerulus was detected successfully, but does not have a correspondence in

TABLE II: Data description: size of the images and number of glomeruli annotated for each patient for each staining.

Patient	Jones H&E		CD10		Sirius Red		PAS	
	size	#glom.	size	#glom.	size	#glom.	size	#glom.
10	95,615×72,293	151	93,623×75,420	165	95,615×73,610	162	95,615×72,293	151
11	69,719×84,475	124	69,719×90,070	115	73,703×82,828	122	73,703×85,955	113
12	37,847×25,401	65	37,847×34,946	63	37,847×28,199	60	37,847×27,705	58
13	107,567 ×61,759	356	109,559×60,113	360	109,559 ×63,569	363	107,567 ×59,126	380
14	93,623×85,952	169	91,631 ×82,661	159	97,607 ×82,002	158	95,615 ×83,812	179
15	73,703 ×52,878	206	73,703 ×53,207	202	71,711 ×52,384	205	71,711 ×50,081	207

the majority of the other stainings. For this case the confidence of the detection could be considered.

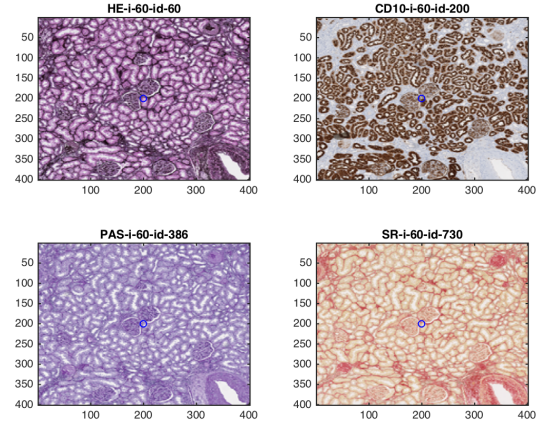
TABLE III: Comparison of HOG and CNN: F1-score for each stain for each method.

Staining	Method	P10	P11	P12	Average
Jones H&E	HOG	59.73	56.79	35.43	50.65
Jones H&E	CNN	<b>75.00</b>	<b>72.36</b>	<b>67.46</b>	<b>71.61</b>
CD10	HOG	48.77	40.64	31.94	40.45
CD10	CNN	<b>89.78</b>	<b>54.36</b>	<b>63.64</b>	<b>69.26</b>
PAS	HOG	64.69	<b>62.82</b>	37.65	55.05
PAS	CNN	<b>76.28</b>	57.64	<b>65.77</b>	<b>66.56</b>
Sirius Red	HOG	60.94	36.14	39.22	45.43
Sirius Red	CNN	<b>70.68</b>	<b>43.61</b>	<b>42.42</b>	<b>52.24</b>

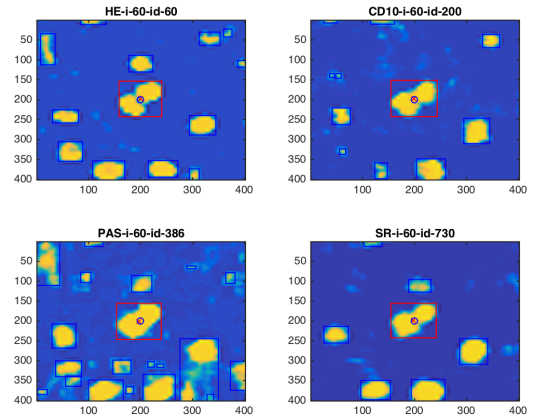
### C. Results on biopsies

We applied the CNN detection framework on a set of 20 needle biopsies from transplanted kidneys. The biopsies were stained with Jones H&E, Sirius Red and PAS and WSI were acquired by an Aperio AT2 scanner at 20X magnification in the proprietary .svs format, scanned images having a resolution of  $0.502 \mu\text{m}/\text{pixel}$ . The number of annotated glomeruli per WSI is between 0 and 27. We randomly selected 4 patients for testing and the remaining 16 patients were used for training. We conducted three detection experiments: (1) We used the CNN trained on nephrectomies to classify glomeruli in biopsies. (2) The CNN from the nephrectomies was used as initialization for training a CNN with the data from the biopsy training set. (3) Random initialization was used for training a CNN with the data from the biopsy training set. For (2) and (3) we used 500 positive and 5000 negative patches chosen randomly from the tissue.

Table V illustrates that applying the CNN trained on nephrectomies (1) directly on the biopsies failed, especially for Sirius Red staining. Using the trained CNN from the nephrectomies as an initialization step for training the CNN using biopsies (2) did improve the results, however it did not perform substantially better than training a CNN on the biopsies using random initialization (3). That really demonstrates that the two problems are distinct (not the same color, not the same shape or texture in the glomeruli of nephrectomies and biopsies, different distributions of glomeruli: in biopsies a higher ratio of glomeruli are located at the outer edge of the tissue and thus more susceptible to artifacts like squeezing than in the bigger tissue parts from nephrectomies). The very high detection rate for P17 and Sirius Red can be explained by the small number of glomeruli (see Table VI), that are all detected well. However, if there is a small number of glomeruli, the



(a) Registered ROIs.



(b) Registered CNN detections.

Fig. 5: The ROIs are correctly registered (a), however the adjacent glomeruli are detected as one object in all 4 stainings (b).

chances are also high to produce very low detection rates (P18 PAS or P19 and P20 Sirius Red), if most or all of the glomeruli are missed.

## IV. CONCLUSION

This work is a first step towards comparative detection of glomeruli using multiple stainings. We have shown that the CNN has better detection performance for glomeruli than the HOG descriptor. The method is applicable on various stainings; however, the performance is better on Jones H&E, Sirius Red and PAS than in CD10 staining. Our co-registration

TABLE IV: Results using the CNN and after applying Mutual comparison 1 and Mutual comparison 2 from the comparative method.

Staining	Method	P10	P11	P12	P13	P14	P15	Average (Std. Deviation)
Jones H&E	CNN	83.94	80.58	77.86	<b>84.59</b>	71.11	<b>89.83</b>	81.32 ( $\sigma = 6.43$ )
Jones H&E	Mutual comparison 1	81.63	78.4	75.81	80.49	68.73	89.18	79.04 ( $\sigma = 6.76$ )
Jones H&E	Mutual comparison 2	<b>84.55</b>	<b>83.3</b>	<b>78.46</b>	81.66	<b>73.01</b>	89.50	81.75 ( $\sigma = 5.61$ )
CD10	CNN	66.33	<b>78.23</b>	64.41	<b>81.54</b>	69.06	82.92	73.75 ( $\sigma = 8.11$ )
CD10	Mutual comparison 1	70.56	69.42	66.07	77.46	67.82	84.29	72.60 ( $\sigma = 6.93$ )
CD10	Mutual comparison 2	<b>71.62</b>	70.59	<b>67.23</b>	77.96	<b>71.71</b>	<b>84.38</b>	73.92 ( $\sigma = 6.19$ )
PAS	CNN	50.26	82.03	62.81	77.50	40.93	84.74	66.38 ( $\sigma = 18.03$ )
PAS	Mutual comparison 1	68.69	82.24	77.03	85.56	69.66	84.29	77.91 ( $\sigma = 7.37$ )
PAS	Mutual comparison 2	<b>69.86</b>	<b>82.55</b>	<b>78.26</b>	<b>85.83</b>	<b>71.80</b>	<b>85.35</b>	78.94 ( $\sigma = 6.86$ )
Sirius Red	CNN	80.53	71.69	70.59	<b>74.58</b>	51.23	88.95	72.92 ( $\sigma = 12.61$ )
Sirius Red	Mutual comparison 1	82.57	67.26	65.49	71.78	76.76	90.08	75.66 ( $\sigma = 9.44$ )
Sirius Red	Mutual comparison 2	<b>83.12</b>	<b>74.63</b>	<b>74.38</b>	72.56	<b>77.44</b>	<b>90.45</b>	78.76 ( $\sigma = 6.81$ )

TABLE V: F1-score of glomeruli detection results on biopsies. Methods are (1) CNN trained on nephrectomies directly applied to biopsies (2) CNN from the nephrectomies used as initialization for training a CNN with data from the biopsies (3) CNN trained on the biopsies data with random initialization.

Staining	Method	P17	P18	P19	P20
Jones H&E	(1)	63.63	57.14	36.36	47.06
Jones H&E	(2)	86.96	<b>80.0</b>	<b>80.0</b>	76.19
Jones H&E	(3)	<b>95.65</b>	<b>80.0</b>	<b>80.0</b>	<b>85.71</b>
PAS	(1)	62.50	36.36	61.54	64.52
PAS	(2)	<b>86.96</b>	<b>61.54</b>	66.67	<b>82.76</b>
PAS	(3)	78.57	42.11	<b>72.73</b>	82.35
Sirius Red	(1)	14.29	57.14	0	0
Sirius Red	(2)	<b>100.0</b>	<b>66.67</b>	50.0	50.0
Sirius Red	(3)	<b>100.0</b>	50.0	<b>66.67</b>	<b>72.73</b>

TABLE VI: Number of annotated glomeruli in the biopsy test set.

Staining	P17	P18	P19	P20
Jones H&E	11	3	5	10
PAS	12	4	5	8
Sirius Red	3	3	2	6

approach allows to correct errors in the difficult staining modalities by superimposing correctly detected glomerular structures from other nearly consecutive sections.

Transferring knowledge from nephrectomies to biopsies proved to be difficult, since the glomerular structures differ due to preparation, embedding and staining. From our experiments, we conclude that learning structures on the nephrectomies should be separated from learning structures on the biopsies.

Tissue segmentation [15] and incorporation of spectral unmixing of multiplex stainings [16], [17] can be useful as a pre-processing step to the CNN training data. Further WSI registration techniques [18], [19] will be considered to enhance our method in future research. Furthermore, we will apply segmentation to separate neighboring objects detected as one object.

#### ACKNOWLEDGEMENT

This work has been funded by the ERACoSysMed project "SysMIFTA" under the grant agreement No. 031L0085 (BMBF, Projektträger Jlich, Germany), No. 9003035004

(NWO, The Netherlands), and No. ANR-15-CMED-0004-03 (ANR, France).

#### REFERENCES

- [1] L. C. Racusen *et al.*, "The banff 97 working classification of renal allograft pathology," *Kidney International*, vol. 55, pp. 713–723, 1999.
- [2] F. Ghaznavi, A. Evans, A. Madabhushi, and M. Feldman, "Digital imaging in pathology: Whole-slide imaging and beyond," *Annu. Rev. Pathol. Mech. Dis.*, vol. 8, pp. 331–59, 2013.
- [3] M. T. McCann, J. A. Ozolek, C. A. Castro, B. Parvi, and J. Kovačević, "Automated histology analysis: Opportunities for signal processing," *IEEE Signal Processing Magazine*, 2015.
- [4] T. Kato, R. Relator, H. Ngouv, Y. Hirohashi, O. Takaki, T. Kakimoto, and K. Okada, "Segmental hog: new descriptor for glomerulus detection in kidney microscopy image," *BMC Bioinformatics*, vol. 16, pp. 316–, 2015.
- [5] R. Marée, S. Dallongeville, J.-C. Olivo-Marin, and V. Meas-Yedid, "An approach for detection of glomeruli in multisite digital pathology," in *In IEEE International Symposium on Biomedical Imaging (ISBI)*, April 2016, pp. 1033–1036.
- [6] A. Sicard, V. Meas-Yedid, M. Rabeyrin, A. Koenig, S. Ducreux, F. Dijoud, V. Hervieu, L. Badet, and E. M., "Computer-assisted topological analysis of renal allograft inflammation adds to risk evaluation at diagnosis of humoral rejection," *Kidney International*, vol. 92, pp. 214–226, 2017.
- [7] G. Litjens, C. I. Sánchez, N. Timofeeva, M. Hermsen, I. Nagtegaal, I. Kovacs, C. Hulsbergen-van de Kaa, P. Bult, B. van Ginneken, and J. van der Laak, "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis," *Scientific reports*, vol. 6, no. 26286, 2016.
- [8] G. Apou, N. S. Schaadt, B. Naegel, G. Forestier, R. Schönmeier, F. Feuerhake, C. Wemmer, and A. Grote, "Detection of lobular structures in normal breast tissue," *Computers in Biology and Medicine*, vol. 74, pp. 91–102, 2016.
- [9] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *In IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2005, pp. 886–893.
- [10] T. Tikkanen, P. Ruusuvaari, L. Latonen, and H. Huttunen, "Training based cell detection from bright-field microscope images," in *In International Symposium on Image and Signal Processing and Analysis (ISPA)*, Sept 2015, pp. 160–164.
- [11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.
- [12] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei, "Imagenet large scale visual recognition challenge," *Internat J Comput Vision*, vol. 115, pp. 211–252, 2015.
- [13] M. Temerinac-Ott, M. Keuper, and H. Burkhardt, "Evaluation of a new point clouds registration method based on group averaging features," in *In IEEE International Conference on Pattern Recognition (ICPR)*, August 2010, pp. 2452–2455.

- [14] R. Mare, L. Rollus, B. Stvens, R. Hoyoux, G. Louppe, R. Vandaele, J.-M. Begon, P. Kainz, P. Geurts, and L. Wehenkel, "Collaborative analysis of multi-gigapixel imaging data using cytomine," *Bioinformatics*, vol. 32, no. 9, pp. 1395–1401, 2016.
- [15] G. Apou, B. Naegel, G. Forestier, F. Feuerhake, and C. Wemmert, "Fast segmentation for texture-based cartography of whole slide images," in *In International Conference on Computer Vision Theory and Applications (VISAPP)*, 2014, pp. 309–319.
- [16] C. Wemmert, J. Kruger, G. Forestier, L. Sternberger, F. Feuerhake, and P. Gañarski, "Stain unmixing in brightfield multiplexed immunohistochemistry," in *In IEEE International Conference on Image Processing (ICIP)*, vol. 13, 2013, pp. 1125–1129.
- [17] A. Bell, D. Friedrich, K. Chaisaowong, T. Braunschweig, R. Knüchel-Clarke, and T. Aach, "Fast virtual destaining of immunocytological specimens," in *In IEEE International Symposium on Biomedical Imaging (ISBI)*, March 2011, pp. 914–917.
- [18] A. Sarkar, Q. Yuan, and C. Srinivas, "A robust method for inter-marker whole slide registration of digital pathology images using lines based features," in *In IEEE International Symposium on Biomedical Imaging (ISBI)*, April 2014, pp. 762–765.
- [19] J. Lotz, J. Olesch, B. Müller, T. Polzin, P. Galuschka, J.M.Lotz, S. Heldmann, H. Laue, A. Warth, B. Lahrman, N. Grabe, O. Sedlacek, K. Breuhahn, and J. Modersitzki, "Patch-based nonlinear image registration for gigapixel whole slide images," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 9, pp. 1812–1819, 2015.