Is the Vascular Network Discriminant Enough to Classify Renal Cell Carcinoma?

Alexis Zubiolo, Eric Debreuve, Damien Ambrosetti, Philippe Pognonec,
Xavier Descombes

Is the Vascular Network Discriminant Enough to Classify Renal Cell Carcinoma?

Alexis Zubiolo[1,4], Eric Debreuve[2,4], Damien Ambrosetti[5], Philippe Pognonec[6], Xavier Descombes[3,4]

[1] *University Nice Sophia Antipolis, Sophia Antipolis, France. Email: alexis.zubiolo@gmail.com*
[2] *CNRS, Sophia Antipolis, France. Email: eric.debreuve@cnrs.fr*
[3] *Inria, Sophia Antipolis, France. Email: xavier.descombes@inria.fr*
[4] *Team Morpheme, Inria/CNRS/University Nice Sophia Antipolis*
[5] *Laboratoire Central d'Anatomo Pathologie, CHU Nice / Hôpital Pasteur, Nice, France.*
Email: ambrosetti.d@chu-nice.fr
[6] *Team TIRO, CEA/University Nice Sophia Antipolis, Nice, France. Email: philippe.pognonec@unice.fr*

Abstract

The renal cell carcinoma (RCC) is the most frequent type of kidney cancer (between 90% and 95%). Twelve subtypes of RCC can be distinguished, among which the clear cell carcinoma (ccRCC) and the papillary carcinoma (pRCC) are the two most common ones (75% and 10% of the cases, respectively). After resection (i.e., surgical removal), the tumor is prepared for histological examination (fixation, slicing, staining, observation with a microscope). Along with protein expression and genetic tests, the histological study allows to classify the tumor and define its grade in order to make a prognosis and to take decisions for a potential additional chemotherapy treatment. Digital histology is a recent domain, since routinely, histological slices are studied directly under the microscope. The pioneer works deal with the automatic analysis of cells. However, a crucial factor for RCC classification is the tumoral architecture relying on the structure of the vascular network. For example, coarsely speaking, ccRCC is characterized by a "fishnet" structure while the pRCC has a tree-like structure. To our knowledge, no computerized analysis of the vascular network has been proposed yet. In this context, we developed a complete pipeline to extract the vascular network of a given histological slice and compute features of the underlying graph structure. Then, we studied the potential of such a feature-based approach in classifying a tumor into ccRCC or pRCC. Preliminary results on patient data are encouraging.

1 Introduction

1.1 Medical context

The renal cell carcinoma (RCC) represents about 90% of kidney cancers, with several hundreds of thousands of new cases each year worldwide [11]. In some situations, surgery (i.e., partial or total nephrectomy) is required. It consists in ablating the tumor together with part or all of the kidney. Once removed, the tumor is analyzed in order to provide a diagnosis (namely, the histological subtype of the tumor [21] and its grade), and therefore an accurate prognosis. This analysis is based on four criteria:

- the cellular morphology of the tumor [2],

- the architecture of the tumor, mainly corresponding to the vascular organization of the tissue [19],

- the protein expression [20],

- the genetic expression [7].

The histological type allows to establish a prognosis [12] and to propose a treatment adapted to the patient. In 2004, the World Health Organization (WHO) has established a categorization into twelve subtypes of renal cell tumors [15]. Only two of them are benign (papillary adenoma and oncocytoma). Among the malignant tumor types, the most common ones are the clear cell renal cell carcinoma (ccRCC, 75% of the cases), the papillary carcinoma (pRCC, 10%), and the chromophobe carcinoma (5%). Note that there is also an official "unknown type" (4-6%), illustrating the fact that biological states are not 100% clear-cut.

1.2 Previous works and proposed viewpoint

Digital histopathology is a technique at the intersection of histopathology and medical/biological imaging. Scanning conventional glass slides of histological slices with a motorized microscope is referred to as Whole Slide Imaging (WSI). It is gaining popularity among physicians and it brings new problems to image processing. For kidney tumors, a large part of the literature deals with the segmentation of the various parts of the tumor [9, 17], or with the analysis of cell and nucleus morphology [10, 22]. However, as mentioned in Section 1.1, cell shape is not the only morphological feature that helps to establish a diagnosis. Indeed, during tumor formation, the vascular network grows. This phenomenon is called angiogenesis (see Fig. 1). Clearly, the histological type is also characterized by the topology and geometry of this network, which is therefore a crucial diagnostic criterion [19]. It is also a factor linked to the grade of the tumor. When the network is regular, it suggests that the tumor develops in a structured way, probably indicating a low grade. Conversely, when the network is chaotic, the development may be seen as aggressive, probably indicating a high grade, malignant tumor. The goal of this work is to extract the vascular network and to perform a feature-based analysis of its structure at a large scale in order to study relationships between vascular structure and histological types. To our knowledge, this kind of feature-based approach has not been proposed yet.
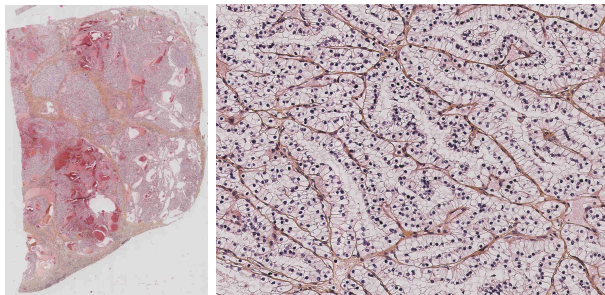

Figure 1: One slice of a tumor: *(left)* the whole slice; *(right)* a 1400×1000 close-up view. On the right image, the brownish network is the vascular network; the blue/violet dots are cell nuclei.

2 Data description and pre-processing

2.1 Histological data

Once the tumor has been extracted, it is prepared and sliced into 3-micron-thick slices. The slices are then stained using a standard procedure. The most commonly used dyes are eosin, an acid that stains pink the cytoplasm and the cell matrix, and hematoxylin, a base that turns the nuclei blue/violet(see Fig. 1). Our database is composed of seven slides scanned by a Leica SCN400 and stored in "scn" format (RGB encoding with 8 bits per channel). The images have a resolution of 0.25 μm and they are around 100 000×100 000 pixels in size. However, not the whole tumor area is of interest. For example, it contains adipose tissue. Hence, the slides were first manually segmented to extract the regions of interest (ROIs), a step that should be automatized in the future. Collectively, the ROIs are, on average, equivalent in size to a 40 000×40 000-image, which is still very large (see Section 2.2).

2.2 Computational burden

Since the slides/ROIs have huge dimensions, they could not be processed in a reasonable time using typical iterative variational or stochastic approaches. In other words, we were constrained to build the pro-
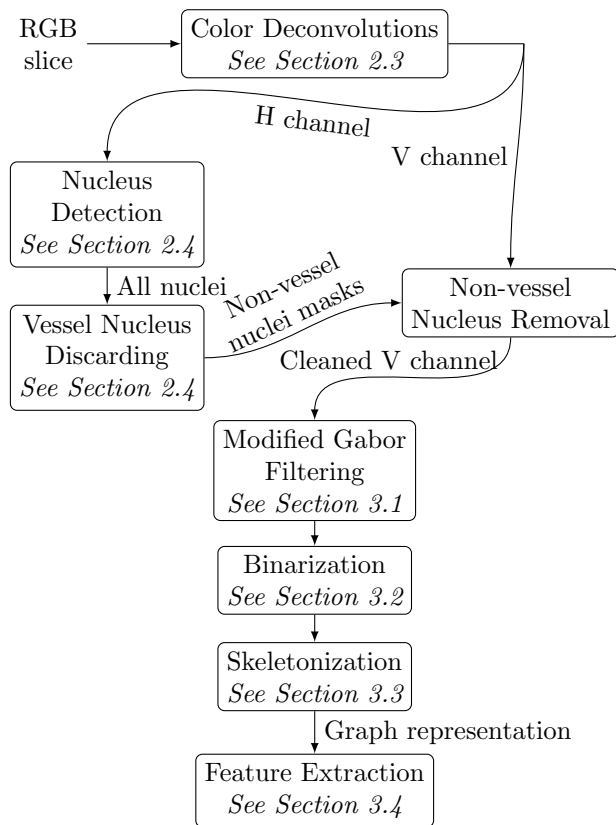
Figure 2: Global processing pipeline: From original slice to feature extraction on the graph model of the vascular network.

cessing pipeline with more basic ingredients, namely, filtering, thresholding, and mathematical morphological tools (see Fig. 2 for an overview of the processing pipeline). Besides the computational load, memory usage is also a concern. As a consequence, the slides were split into slightly overlapping sub-images, and each sub-image was processed independently. The overlap between neighboring sub-images simplified the recombination of the individual results. Note that the result images presented in the following are close-ups of whole slides, and are never larger than 1000×1000 pixels.

2.3 Channel combination (a.k.a. color deconvolution)

In order to make the image elements (here vascular network and nuclei) appear more clearly in the images, color channels must be combined together, a pre-processing step known as color deconvolution. A simple combination that emphasizes the vascular network is the average of the red and blue channels, resulting in the V channel (see Fig. 3):

$$V = 0.5 \frac{R + B}{\sqrt{R^2 + G^2 + B^2}}. \tag{1}$$

Note that more sophisticated solutions exist, for example using a principal component analysis (PCA) [18].More recently, empirical formulas corresponding to markers used in histology (including hematoxylin and eosin) have been proposed [1]. They are based on the contrast invariant channels proposed in [6]. Nuclei are more clearly seen on a pure hematoxylin channel (the H channel; See Fig. 3). It is defined as:

$$H = \frac{B}{C_3} \quad \text{where} \quad C_3 = \arctan\left(\frac{B}{\max(R, G)}\right). \tag{2}$$

2.4 Detection and removal of non-vessel nuclei

A nuclei detection step (on the H channel) may seem as an out-of-scope processing step for studying the vascular network. Actually, the nuclei appear quite distinctively in the V channel (see Fig. 4). Nevertheless, this channel is dedicated to the vascular network extraction, which is based on the (natural) assumption that the network has a tubular structure. The presence of nearby nuclei limits the accuracy of this assumption. Hence, detecting the nuclei on the H channel is only used as a mean to remove them efficiently from the V channel (set the pixels of the nuclei masks to zero). Note however that the vessels also contain nuclei. Their removal can create discontinuities in the extracted vascular network. Therefore, once the nuclei are detected, they must be classified into vessel and non-vessel nuclei, with only the non-vessel nuclei to be removed from the V channel. The
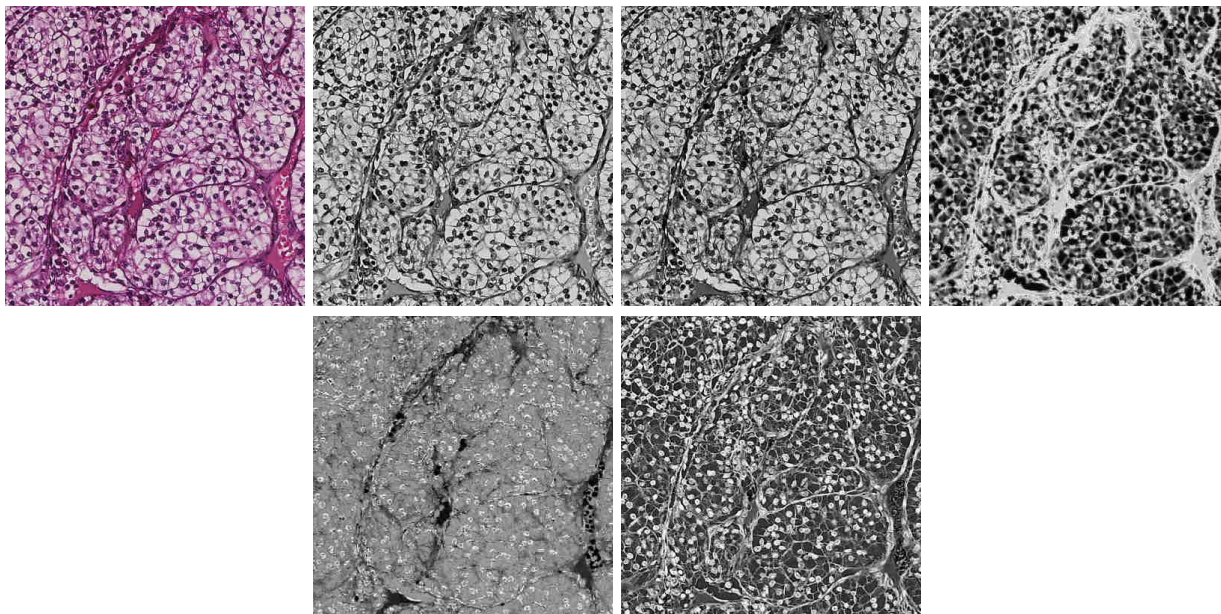
Figure 3: Color deconvolution. ***(Top row)*** Vascular network (V channel): *(from left to right)* RGB image, red channel, blue channel, V channel.
(Bottom row) Nuclei (H channel): *(from left to right)* C_3 channel, H channel (for hematoxylin). All images are 1000×1000 pixels.
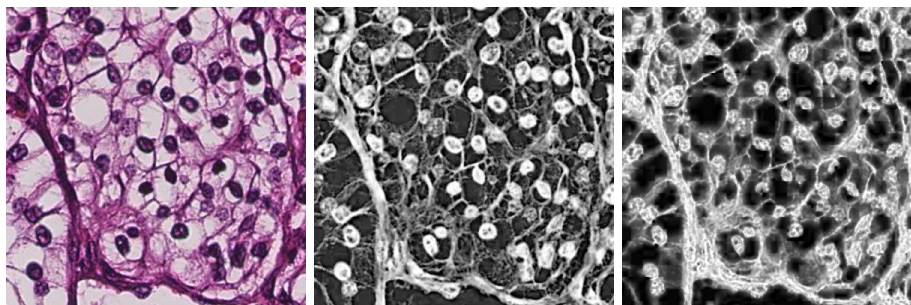

Figure 4: Nuclei visibility on different channels. *(From left to right)* RGB image, H channel (where the nuclei have the brightest and sharpest appearance), V channel (used for vascular network extraction; however, the nuclei also appear quite distinctively). All images are 400×400 pixels.

distinction between the two types of nuclei can be performed based on morphology: non-vessel nuclei have a round shape while vessel nuclei are elongated (see Fig. 5).

The nuclei detection procedure relies on mathe-matical morphology tools.The knowledge of the image resolution and of the typical size of a nucleus allows to set the procedure parameters appropriately to detect bright blobs with a size in the correct range. Specifically, a morphological opening, a thresholding,
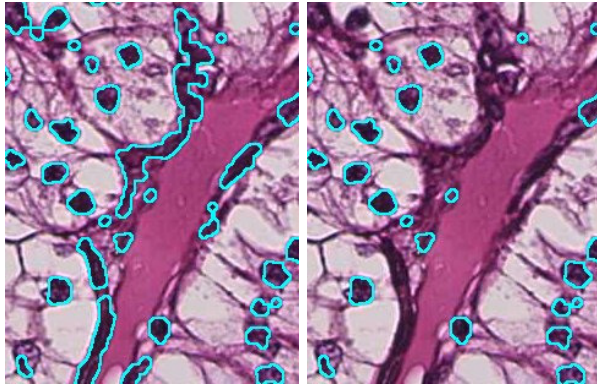
4

Figure 5: Nuclei detection in the H channel. *(Left)* Detection of all the nuclei (non-vessel and vessel nuclei). *(Right)* Result after keeping only non-vessel nuclei (the elongated, vessel nuclei have been discarded from the detection). Images are 250×300 pixels.
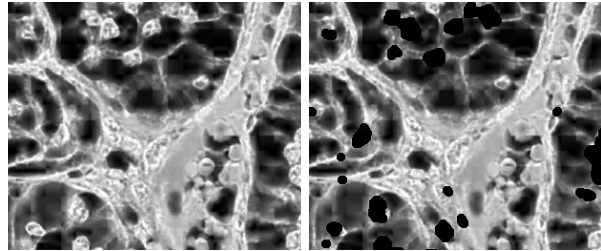

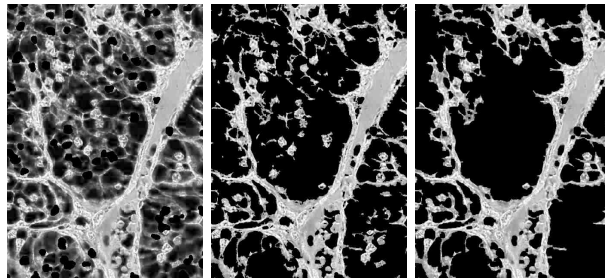Figure 6: Nuclei removal in the V channel. Before *(left)* and after *(right)* removal. Images are 300×250 pixels.

Figure 7: V channel thresholding. *(From left to right)* *(a)* Original channel, *(b)* after thresholding (a), *(c)* after removing tiny components in (b). Images are 450×640 pixels.

and a morphological closing are successively applied to the H channel. At this point, the vessel and non-vessel nuclei are detected (see Fig. 5).

To distinguish between vessel and non-vessel nuclei, two criteria are used: size and shape. Basically, a detected blob larger than a certain threshold (chosen equal to 5 times the average nucleus size) corresponds to a vessel nucleus or an aggregation of vessel nuclei. Moreover, a blob shape criterion is computed as the eccentricity of the ellipse having the same moments of order 2. The blobs with a large eccentricity are also classified as vessel nuclei. The remaining, non-vessel nuclei form an image of masks applied to the V channel to force the corresponding pixels to zero (see Fig 6). The vascular network extraction of Section 3 is performed on the resulting, cleaned V channel.

3 Vascular network extraction and analysis

3.1 Filtering of the V channel

A thresholding step is necessary to reduce the noise around vessels. An hysteresis thresholding [3] is more suitable since the high threshold allows to retrieve the main pieces of the network while the low threshold allows to maintain most of them connected. Tiny, unconnected components are finally removed. This procedure is illustrated in Fig. 7.

Several filters have been proposed to detect tubular structures [4, 13].Among them, the Gabor filtering is one of the most popular. In the 2D case, Gabor filters are defined by:

$$G(x,y) = \cos\left(\frac{x'\pi}{\sqrt{2}\sigma}\right) \exp\left(-\frac{x'^2 + y'^2}{2\sigma^2}\right) \quad (3)$$

where $x' = x\cos\theta + y\sin\theta$, $y' = x\sin\theta - y\cos\theta$, and σ and θ are respectively the extent and the orientation of the filter (see Fig. 8). Filtering an image by a Gabor set amounts to computing the convolution of the image with Gabor filters of predefined values of σ and θ, then taking the maximum of the filtering responses pixelwise. Instead of the classical form (3),
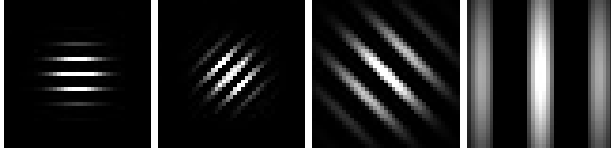
5

Figure 8: Gabor filters for various values of σ and θ.

we propose to use a modified version [16]. The idea is to cut the Gabor filters along their axis of symmetry into two half filters, formally:

$$G^+(x,y) = \begin{cases} G(x,y) & \text{if } x \geq 0 \\ 0 & \text{otherwise,} \end{cases} \qquad (4)$$

$$G^-(x,y) = \begin{cases} G(x,y) & \text{if } x \leq 0 \\ 0 & \text{otherwise.} \end{cases} \qquad (5)$$

Using such half filters allows to avoid false alarms along contours that do not delimit tubular structures. For given values of σ and θ, the pixelwise minimum of the respective responses to the filters G^+ and G^- is computed. Then, just like in the case of the classical Gabor filters, the pixelwise maximum of the responses corresponding to predefined combinations of extents and orientation is retained (see Fig. 9). In the following, this whole filtering procedure will be referred to as the modified Gabor filter bank (mGFB) processing step.

3.2 Binarization of the modified Gabor filtering

Once the mGFB applied, the next step is to segment (here, binarize) the resulting response. Again, a hysteresis thresholding is adapted to our goal: the high threshold highlights areas of strong response to the mGFB (which correspond to tubular structures), then the low threshold allows to keep the main network pieces connected. In order to get a cleaner result, a few post-processing steps are applied (see the final result in Fig. 9):

- a morphological closing for a smoother segmentation,

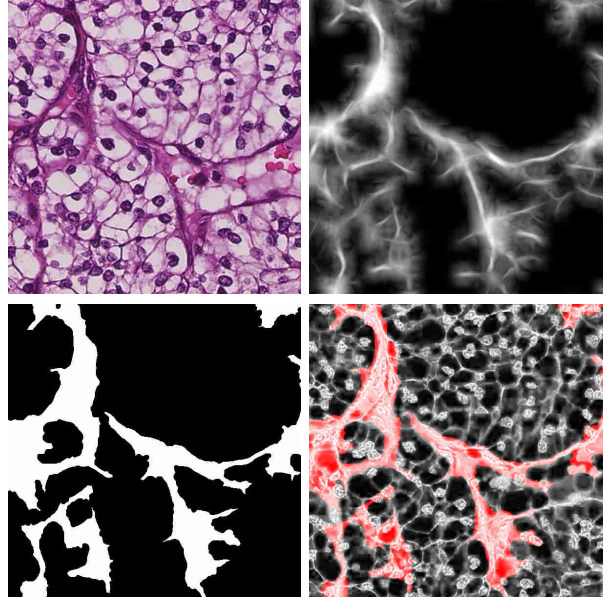- a hole filling procedure for holes in the segmen-

Figure 9: Gabor filtering of the V channel. *(In lexicographical order) (a)* Original channel, *(b)* modified Gabor filter bank (mGFB) applied to (a), *(c)* binarization of (b), *(d)* (c) superimposed on (a). All images are 600×600 pixels.

tation with a size below a given threshold,[1]

- the removal of tiny components corresponding to segmentation noise.

3.3 Graph-based model

The vascular network can be represented by a graph. This graph can be built from the skeleton of the segmentation obtained in Section 3.2. Many skeleton extraction algorithms have been proposed. Among them, we chose a thinning method [14] (see Fig. 10), a framework which preserves the topological properties of the network and is considered robust.

The resulting, raw skeleton must be post-processed to remove small terminal branches that do not cor-

---

[1]This step is important because the skeleton of the segmentation will be computed later in the process (see Section 3.3), and holes create loops in the skeleton that are not related to the underlying vascular network.
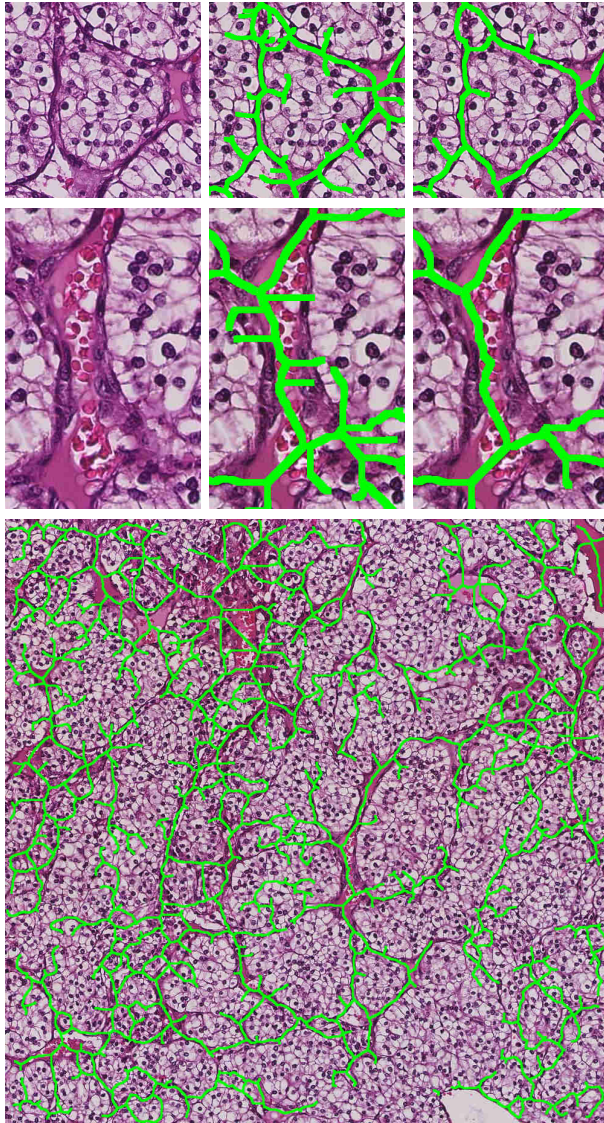
Figure 10: Skeleton of the vascular network segmentation. ***(From left to right in the first 2 rows)*** Original image, raw skeleton, post-processed skeleton. ***(Bottom image)*** Post-processed skeleton on a larger area. Image sizes: first row=500×500, second row=320×500, bottom image=1000×1000.

respond to actual branches of the vascular network[2]

---

[2]In other words, it is a pruning step.

but are rather due to either:

- the roughness of the segmentation contours (despite the smoothing described in Section 3.2),

- or the segmentation being too thick (either simply because some vessels are thick, or at the intersections of vessels).[3]

First, the terminal branches of the skeleton are extracted [8]. Then, among such branches, the ones that are either too small or composed of purely vertical, horizontal, or diagonal pieces are removed. This removal procedure is iterated until the resulting skeleton remains unchanged (see Fig. 10). In practice, at most 5 iterations are necessary. This final skeleton corresponds to our graph model of the vascular network.

3.4 Feature extraction

The obtained graph model is used to calculate objective features of the vascular network. As mentioned in Section 1.1, the two most common carcinomas are the clear cell renal cell carcinoma (ccRCC) and the papillary carcinoma (pRCC). Coarsely speaking, ccRCC is characterized by a "fishnet" vascular structure while the pRCC has a tree-like structure. Therefore, we propose to compute features that should allow to distinguish between these two cases and that would permit to apprehend the regularity or the chaotic pattern[4] of this network:

- the number of terminal (or end) branches (NE),

- the average length of the terminal branches (LE),

- the number of junctions (NJ),

- the average length of the non-terminal branches (LJ).

As a summary, the global processing pipeline, from original slice to feature extraction, was introduced in Fig. 2.

---

[3]In this case, thinning algorithms often create piecewise vertical, horizontal, or diagonal branches to *cover* the whole surface of the segmentation.

[4]which, as already mentioned, is linked to the tumor grade

Data description		Features					
Patient	Type	NE	LE	NJ	LJ	NE/NJ	LE/LJ
1	ccRCC	17794	61.1	139032	101.3	0.12	0.60
2	ccRCC	49945	67.8	118645	105.2	0.42	0.64
2	ccRCC	26850	63.7	76513	100.6	0.35	0.63
2	ccRCC	6204	62.4	25354	101.4	0.24	0.62
2	ccRCC	28489	56.1	68264	88.9	0.41	0.63
3	pRCC	13933	72.9	23724	101.6	0.58	0.72
3	pRCC	23164	67.5	50085	100.6	0.46	0.67
Average	ccRCC	25856	62.2	85562	99.5	0.31	0.62
Average	pRCC	18548	70.2	36904	101.1	0.52	0.69

Table 1: Data description (patient number and histological type) and extracted features for seven slices. The last two rows present average features for each histological type.

3.5 Vascular network analysis

Table 1 presents the features that have been extracted on seven slices taken from tumors of three patients. All three tumors have been graded 3 on the Fuhrman scale [5].[5] Although the number of cases is small, the results suggest two tendencies: the NE/NJ ratio is generally higher for pRCC than for ccRCC (on average, about 68% higher), and the average length of the terminal branches (LE) is larger for pRCC than for ccRCC (on average, LE/LJ ratio about 11% higher). This is in accordance with their respective tree-like (pRCC; more, longer terminal branches) versus "fishnet" (ccRCC; more junctions, shorter terminal branches) vascular structures.

4 Conclusion

We proposed an automatic processing pipeline to model (as a graph) and analyze the vascular network of RCCs. The preliminary results obtained on routinely acquired tumor slices suggest that features such as the NE/NJ and LE/LJ ratios are potential biomarkers to distinguish between pRCC and ccRCC tumor types. This points toward a positive answer to our initial question: "Is the vascular network discriminant enough to classify renal cell carcinoma?". Naturally, it is essential to process much more patient data. In particular, the complete pipeline de-

pends on a significant number of parameters (thresholds, mGFB parameters. . .) that were tuned manually on the same images as the ones used for testing. A larger database would allow (i) to estimate them using cross-validation and (ii) to properly separate tuning and testing.

Regarding biomarkers, it can be interesting to study the potential of other markers such as the number of "cycles" of the vascular graph/network, their size, the size of the vessels. . . Regarding additional perspectives, let us mention that, when the vascular network grows, some cells tend to align to it. Hence, cell arrangement and vascular structure represent complementary histological information. Therefore, cell detection could help extracting the vascular network more accurately and efficiently. Moreover, as far as analysis is concerned, it seems legitimate to try to combine the proposed approach with the cell morphology studies mentioned in Section 1.2.

References

[1] Maya Alsheh Ali. *Analyse statistique de populations pour l'interprtation d'images histologiques*. PhD thesis, Université Paris-Descartes, 2015.

[2] John C Cheville, Christine M Lohse, Horst Zincke, Amy L Weaver, and Michael L Blute. Comparisons of outcome and prognostic features among histologic subtypes of renal cell carcinoma. *The American journal of surgical pathology*, 27(5):612–624, 2003.

[3] Alexandru-Paul Condurache and Til Aach. Vessel segmentation in angiograms using hysteresis thresholding. In *IAPR Conference on Machine Vision Applications*. Citeseer, 2005.

[4] Alejandro F Frangi, Wiro J Niessen, Koen L Vincken, and Max A Viergever. Multiscale vessel enhancement filtering. In *Medical Image Computing and Computer-Assisted InterventationMICCAI98*, pages 130–137. Springer, 1998.

[5] Susan A Fuhrman, Larry C Lasky, and Catherine Limas. Prognostic significance of morphologic parameters in renal cell carcinoma. *The American journal of surgical pathology*, 6(7):655–664, 1982.

[6] Theo Gevers and Arnold WM Smeulders. Pictoseek: Combining color and shape invariant features for im-

---

[5]The Fuhrman grade is an integer value between 1 and 4 characterizing the malignity of the tumor, 1 being the most benign and 4 the most malignant.

age retrieval. *IEEE Trans. Image Process.*, 9(1):102–119, 2000.

[7] John PT Higgins, Rajesh Shinghal, Harcharan Gill, Jeffrey H Reese, Martha Terris, Ronald J Cohen, Michael Fero, Jonathan R Pollack, Matt van de Rijn, and James D Brooks. Gene expression patterns in renal cell carcinoma assessed by complementary dna microarray. *The American journal of pathology*, 162(3):925–932, 2003.

[8] Nicholas R Howe. Document binarization with automatic parameter tuning. *International Journal on Document Analysis and Recognition (IJDAR)*, 16(3):247–258, 2013.

[9] Chao-Hui Huang, Antoine Veillard, Ludovic Roux, Nicolas Loménie, and Daniel Racoceanu. Time-efficient sparse analysis of histopathological whole slide images. *Computerized medical imaging and graphics*, 35(7):579–591, 2011.

[10] Humayun Irshad, Antoine Veillard, Ludovic Roux, and Daniel Racoceanu. Methods for nuclei detection, segmentation, and classification in digital histopathology: A review – current status and future potential. *Biomedical Engineering, IEEE Reviews in*, 7:97–114, 2014.

[11] Eric Jonasch, Jianjun Gao, and W Kimryn Rathmell. Renal cell carcinoma. *BMJ*, 349, 2014.

[12] Kirk A Keegan, Clayton W Schupp, Karim Chamie, Nicholas J Hellenthal, Christopher P Evans, and Theresa M Koppie. Histopathology of surgically treated renal cell carcinoma: survival differences by subtype and stage. *The Journal of urology*, 188(2):391–397, 2012.

[13] Karl Krissian, Grégoire Malandain, Nicholas Ayache, Régis Vaillant, and Yves Trousset. Model-based detection of tubular structures in 3d images. *Computer vision and image understanding*, 80(2):130–171, 2000.

[14] Louisa Lam, Seong-Whan Lee, and Ching Y Suen. Thinning methodologies-a comprehensive survey. *IEEE Trans. Pattern Anal. Mach. Intell.*, 14(9):869–885, 1992.

[15] Antonio Lopez-Beltran, Marina Scarpelli, Rodolfo Montironi, and Ziya Kirkali. 2004 who classification of the renal tumors of the adults. *European urology*, 49(5):798–805, 2006.

[16] Alejandro Mottini. *Axon Morphology Analysis: from Image Processing to Modelling*. Theses, Universite de Nice Sophia-Antipolis (UNS), September 2014.

[17] Nandita Nayak, Hang Chang, Alexander Borowsky, Paul Spellman, and Bahram Parvin. Classification of tumor histopathology via sparse feature learning. In *Biomedical Imaging (ISBI), IEEE International Symposium on*, pages 410–413. IEEE, 2013.

[18] Yahui Peng, Yulei Jiang, Laurie Eisengart, Mark A Healy, Francis H Straus, and Ximing J Yang. Computer-aided identification of prostatic adenocarcinoma: Segmentation of glandular structures. *Journal of pathology informatics*, 2, 2011.

[19] Edmond Sabo, Albina Boltenko, Yanina Sova, Avi Stein, Shira Kleinhaus, and Murray B Resnick. Microscopic analysis and significance of vascular architectural complexity in renal cell carcinoma. *Clinical Cancer Research*, 7(3):533–537, 2001.

[20] Cecilia Sarto, Alessandro Marocchi, Jean-Charles Sanchez, Daniela Giannone, Séverine Frutiger, Olivier Golaz, Marc R Wilkins, Giancarlo Doro, Francesco Cappellano, Graham Hughes, et al. Renal cell carcinoma and normal kidney protein expression. *Electrophoresis*, 18(3-4):599–604, 1997.

[21] Stephan Störkel, John N. Eble, K. Adlakha, Mahul Amin, Michael L. Blute, David G. Bostwick, M. Darson, Brett Delahunt, and K. Iczkowski. Classification of renal cell carcinoma. *Cancer*, 80(5):987–989, 1997.

[22] Fang-Cheng Yeh, Anil V Parwani, Liron Pantanowitz, and Chien Ho. Automated grading of renal cell carcinoma using whole slide imaging. *Journal of pathology informatics*, 5, 2014.