

Patrick Wurzel*, Jörg Ackermann, Hendrik Schäfer, Sonja Scharf, Martin-Leo Hansmann and Ina Koch

Detection of follicular regions in actin-stained whole slide images of the human lymph node by shock filter

<https://doi.org/10.1515/hsz-2020-0178>

Received May 11, 2020; accepted December 2, 2020;

published online December 24, 2020

Keywords: computer vision; digital pathology; human lymph node; morphological filtering; shock filter; whole slide image.

Abstract: Human lymph nodes play a central part of immune defense against infection agents and tumor cells. Lymphoid follicles are compartments of the lymph node which are spherical, mainly filled with B cells. B cells are cellular components of the adaptive immune systems. In the course of a specific immune response, lymphoid follicles pass different morphological differentiation stages. The morphology and the spatial distribution of lymphoid follicles can be sometimes associated to a particular causative agent and development stage of a disease. We report our new approach for the automatic detection of follicular regions in histological whole slide images of tissue sections immuno-stained with actin. The method is divided in two phases: (1) shock filter-based detection of transition points and (2) segmentation of follicular regions. Follicular regions in 10 whole slide images were manually annotated by visual inspection, and sample surveys were conducted by an expert pathologist. The results of our method were validated by comparing with the manual annotation. On average, we could achieve a Zijbendos similarity index of 0.71, with a standard deviation of 0.07.

Introduction

A human organism needs a functional immune system to defend against bacteria, viral structures and tumor cells. In this concept an integral part of the adaptive immune system is the structural and cellular flexibility of lymph nodes (Oswald et al. 2019). Important and sometimes critical components of the immunologically flexible lymph nodes are especially the lymphoid follicles which can be seen by light microscopy as spherical structures (Willard-Mack 2006). During an ongoing infection and the associated antigenic contact, lymphoid follicles pass different stages of development and change their morphology (Victoria and Nussenzweig 2012). Many known diseases, as, e.g., follicular lymphoma (Broyde et al. 2009), and different benign lymphadenopathies (Weiss and O'malley 2013), are associated with specific morphological findings of lymphoid follicles. For example, rheumatoid arthritis is characterized by a follicular hyperplasia in which lymphoid follicles become heavily enlarged (Nosanchuk and Schintzier 1969).

Computer-aided diagnosis will get an increasing importance in future diagnostics. Novel cutting-edge technologies facilitate a differentiated phenotyping of disease entities and generate a huge amount of data. An example is histology and immunohistochemistry using multi-staining and digitalization that leads to gigabytes of data for an individual patient. To handle data in future routine diagnostics, there is a strong need to develop and establish computer-aided diagnosis (Doi 2007).

Nowadays, lymphoid follicles are visually inspected by a pathologist using a light microscope in diagnostic routine. Various algorithms for the detection of lymphoid follicles have been presented. Belkacem-Boussaid et al. (Belkacem-Boussaid et al. 2011) have applied a region-based approach using curve evolution to segment lymphoid follicles. Tosta et al. (Tosta et al. 2017, 2018) have proposed unsupervised

*Corresponding author: Patrick Wurzel, Goethe-Universität Frankfurt am Main, Molecular Bioinformatics, Institute of Computer Science, Robert-Mayer-Str. 11-15, 60325 Frankfurt am Main, Germany; and Frankfurt Institute for Advanced Studies, Ruth-Moufang-Straße 1, 60438 Frankfurt am Main, Germany, E-mail: wurzel@fias.uni-frankfurt.de. <https://orcid.org/0000-0002-2366-9124>

Jörg Ackermann and Ina Koch, Goethe-Universität Frankfurt am Main, Molecular Bioinformatics, Institute of Computer Science, Robert-Mayer-Str. 11-15, 60325 Frankfurt am Main, Germany

Hendrik Schäfer, Hospital of the Goethe University Frankfurt, Theodor-Stern-Kai 7, 60590 Frankfurt am Main, Germany

Sonja Scharf, Goethe-Universität Frankfurt am Main, Molecular Bioinformatics, Institute of Computer Science, Robert-Mayer-Str. 11-15, 60325 Frankfurt am Main, Germany; and Frankfurt Institute for Advanced Studies, Ruth-Moufang-Straße 1, 60438 Frankfurt am Main, Germany

Martin-Leo Hansmann, Frankfurt Institute for Advanced Studies, Ruth-Moufang-Straße 1, 60438 Frankfurt am Main, Germany

segmentation methods for the detection of neoplastic nuclei in histological whole slide images (WSI) to identify indicative regions. A worthwhile method is supervised machine learning, however, a large number of annotated training images are required. A challenge for all segmentation methods are the broad variety of the morphological structures and unsharp edges of lymphoid regions, as, e.g., follicles, cortex, paracortex and medulla (Willard-Mack 2006).

In the 1990s, Osher and Rudin (Osher and Rudin 1990) have proposed a hyperbolic partial differential equation (PDE) for sharpening and restoring of images called shock filter. The iterative application of the shock filter produces sections of constant functions that show fracture points at inflection points (Belean et al. 2015a,b; Osher and Rudin 1990). In the following decades, their method, known as shock filter method, has found many applications in the field of computer vision (Goyal et al. 2020; Huo et al. 2016; Money and Kang 2008; Varatharajan et al. 2018). An exemplary application is the segmentation of spots on microarrays. Belean et al. (Belean et al. 2015a) have applied the shock filter method to compute the border points to envelope microarray spots of various morphologies.

The aim of this work was the development of a method to detect lymphoid regions and substructures in human biopsies. We concentrated on the detection of lymphoid follicles in tissue sections that are stained for actin. In actin-stained tissue lymphoid follicles are well-differentiated and structures are distinctly pronounced. The staining of actin enables even the detection of multiple lymphoid substructures (Dominguez and Holmes 2011). Here, we abstained from studying the detection of complex substructures as, e.g., sinus and vascular structures.

In our approach, we focused on the applicability of the shock filter method to detect transition points between lymphoid regions. The transition points marked out the border lines of lymphoid follicles. The QuickHull-algorithm (Barber et al. 1996) transformed transition points in the tissue section to an annotation of the individual lymphoid follicles. The quality of the segmentation outperformed standard methods of computer vision, i.e., watershed and Otsu thresholding. We discussed the limitations of our approach and its potential application for measuring the spatial structure of lymph nodes.

Results

The solid line in Figure 1 shows values of the Zijdenbos similarity index (ZSI) of the shock filter approach (FLDetect) for 10 images, see Material and Methods. The values of the ZSI are in the range of $0.6 \leq ZSI \leq 0.8$ with a mean of $\overline{ZSI} = 0.71 \pm 0.07$. The values are excellent, i.e., $ZSI \geq 0.7$, for six out of 10 WSI and less favorable, i.e., $ZSI \leq 0.7$, for four WSI (Belkacem-Boussaid et al. 2011). The ZSI is highest, i.e., $ZSI \approx 0.8$, for WSI-7, WSI-8, and WSI-9.

The shock filter approach yielded lowest values of ZSI, i.e., $ZSI \approx 0.6$, for WSI-1, WSI-2, and WSI-10. By visual inspection, we identified image 1 and 10 to have the lowest overall contrast, see Table in Figure 10. Image two contained, beside local regions of low contrast, a number of follicles whose annotation posed a special challenge for automatic approaches. Exemplary, Figure 2 shows a subsection of WSI-2. In the subsection several oval structures (light gray) of lymphoid follicles are arranged around an elongated top-down structure of a vessel (bright region) and surrounded by tissue (dark gray). Black crosses denote

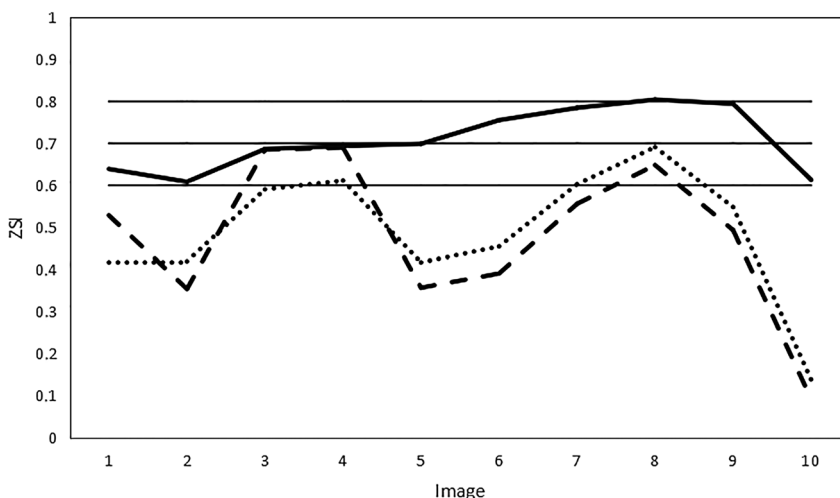


Figure 1: Zijdenbos similarity index (ZSI) for 10 whole slide images (WSI). The shock filter approach (solid line) yields values of the ZSI better than 0.6 for all images and excellent values, i.e., $ZSI \geq 0.7$, for six WSI. Averaged overall images the shock filter approach improves the ZSI by 45% when compared to watershed segmentation (dashed line) and Otsu thresholding (dotted line).

transition point pairs which fulfill Eqs. (15)–(18) and (19)–(22), respectively. The transition point pairs are located on the outsides of follicles but do not completely encircle individual follicles as, e.g., shown for a well-differentiated and isolated lymphoid follicle in Figure 9. As a result, a transition point pair may mark the border lines of two different lymphoid follicles. In Figure 2, horizontal line (a) indicates a transition point pair that spans the borders of two lymphoid follicles in x -direction. Vertical line (b) indicates a transition point pair that spans the borders of two lymphoid follicles in y -direction. Relevant intensity jumps on the border of the lymphoid follicles are ignored by our approach because of threshold conditions (15) and (19), respectively. Because of the ignored intensity jumps, several lymphoid follicles got aggregated. The resulting cluster of lymphoid follicles got filtered because of its unrealistic, large size.

On each image of our data set, the shock filter approach outperformed the standard methods Otsu thresholding and the watershed approach, see Dataset in Materials and Methods. The dashed line and the dotted line in Figure 1 show corresponding values of the ZSI obtained with watershed segmentation and Otsu thresholding, respectively. For the watershed segmentation and Otsu thresholding, the ZSI vary in a broad range of $0.1 \leq ZSI \leq 0.7$ with mean values of $\overline{ZSI} = 0.48 \pm 0.19$ and $\overline{ZSI} = 0.49 \pm 0.16$, respectively. The shock filter approach improved the mean ZSI by 45% compared to these standard approaches. The performance of all three methods

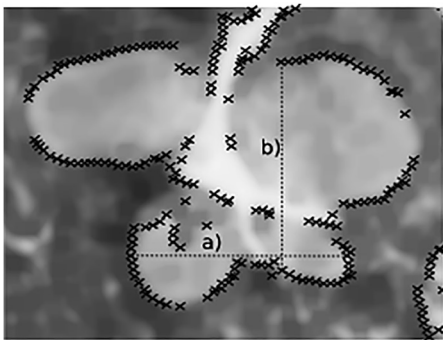


Figure 2: Subsection of WSI-2 with several oval structures (light gray) of lymphoid follicles that are arranged around an elongated top-down structure of a vessel (bright region) and surrounded by tissue (dark gray). Black crosses denote transition point pairs located on the outsides of follicles and the vessel. The transition point pairs do not completely encircle individual follicles. A transition point pair may mark the border lines of two different lymphoid follicles. Horizontal line (a) indicates a transition point pair that spans the borders of two lymphoid follicles in x -direction. Vertical line (b) indicates a transition point pair that spans the borders of two lymphoid follicles in y -direction.

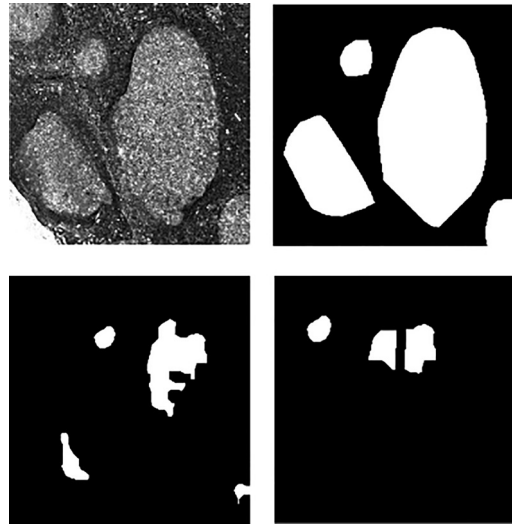


Figure 3: A subsection of WSI-10 with several lymphoid follicles (top left) and the according binary masks of segmentation that are obtained by shock filter approach (top right), Otsu thresholding method (bottom left), and watershed approach (bottom right).

suffered from low contrast in local regions of images. In comparison of three methods, the shock filter approach turned out to be most robust against intensity fluctuation and low contrast. Exemplary, Figure 3 shows a subsection of WSI-10 with several lymphoid follicles and binary masks of segmentations that are obtained by shock filter approach, Otsu thresholding method, and watershed approach. The shock filter identifies correctly three follicles of which Otsu thresholding method and watershed approach identify only a minor fraction of the relevant regions.

Whereas Otsu thresholding and watershed turned out to be sensitive to a lack of homogeneity in the brightness of the WSI, the shock filter enhanced and utilized local intensity jumps and identified the border of structures. In two subsequent steps, the shock filter applied a set of filter rules and an adapted cluster algorithm to facilitate an effective detection and segmentation of lymphoid follicles.

Discussion

We evaluated the shock filter iteration to identify lymphoid follicles in images of actin-stained tissue sections. Our approach is fully automated. We generated a gold standard annotation by careful manual visual inspection of 10 WSIs. We evaluated the quality of the results of the shock filter approach by a comparison with the gold standard annotation. The results of the shock filter approach were excellent for six out of 10 test images. For four test images,

the results were less favorable, i.e., the ZSI drops below 70% and reaches the minimal value of 60% for two images. The lower quality was caused by an increasing number of follicles with less defined borders. This was often correlated with a high density of surrounding unstained substructures.

We observed similar deficiencies for the Otsu's method and watershed segmentation. On our 10 test images, the shock filter approach outperformed Otsu's method and watershed segmentation. A final conclusive comparison of the shock filter with standard methods may require an evaluation of a larger series of WSIs. Prerequisite is the availability of an appropriate gold standard, i.e., an annotation of lymphoid follicles in all WSI. A larger data set may facilitate the worthwhile inclusion of machine learning approaches (Abadi et al. 2015; Pedregosa et al. 2011) in the comparison.

The performance of all three methods, Otsu's method, watershed segmentation, and shock filter, turned out to drop down for WSIs that had no ideal homogeneous brightness, low contrast, or contained vaguely separated structures. Inhomogeneous staining, local low contrast, and intertwined substructures are common in WSI of lymph nodes. Standard operating procedures of quality control address the requirements of clinical routine. By the visual inspection and diagnosis, a pathologist may be able to compensate for inhomogeneous staining, low contrast, local deviation of intensity, and additional artifacts. The three computational approaches, Otsu's method, watershed segmentation, and shock filter are, however, not designed to compensate unsharp borders. The diverse quality of WSIs remains a challenge for the correct automatic detection of lymphoid follicles.

The results of the shock filter approach are promising. In the future, the shock filter method may be enhanced by considering additionally the texture of regions of lymphoid follicles. Additionally, other methods for the transformation of transition points to annotations of lymphoid follicles, like randomized algorithms may allow a better approximation of follicular morphologies (De Berg et al. 2008; Guibas et al. 1992).

Regions of follicles contain more actin positive pixels than the surrounding sinus structure or a vessel. The challenge to separate an individual follicle in a cluster of intertwined structures arose from an incomplete enclosure by transition point pairs; relevant transition point pairs were missed. The shock filter iterations yielded a hierarchy of positions of intensity jumps. The approach ignored the majority of these positions by applying simple selection rules with threshold values that were determined and set

for the entire WSI. Exploring the entire hierarchy of intensity jumps may improve the performance of a shock filter approach significantly. Methods in the field of geometric point patterns (Kendall 1989) and spatial networks (Barthélemy 2011) may contribute to future improvements.

Materials and methods

Dataset

Ten actin-stained tissue sections, shown in Figure 10 were provided by the *Dr. Senckenbergisches Institut für Pathologie*. The tissue sections were various slices of one human lymph node. The slices were digitized with an *Aperio ScanScope XT* scanner. The maximal resolution of an image was 0.25 μm per pixel. We adopted a lower resolution of 7.8944 μm in our approach. Figure 4 shows a WSI of an exemplary slice. The size of the image is about 35.2 \times 28.5 mm. The prominent red color in the image labels actin which is stained by a fuchsin dye. The blue color shows cells that express the programmed cell death protein 1 (PD1) as, e.g., a subset of mainly T cells. Lymphoid follicles are round objects with low density of actin illustrated by the bright ellipsoid objects in the tissue. The overlay shows a magnification of an individual lymphoid follicle. We applied the color deconvolution method of Ruifrok and Johnston (Ruifrok and Johnston 2001) to extract a grayscale image of the actin staining. We identified the tissue (region of interest, ROI) by thresholding the grayscale image. Figure 5 depicts an individual lymphoid follicle in a subsection of a grayscale image. Before applying our shock filter method (FDLdetect), we analyzed each image gray-scale by three standard methods.

- (1) *Manual annotation*. We manually annotated follicular regions inside every WSI with the *VGG Image Annotator (VIA)* (Dutta and Zisserman 2019; Dutta et al. 2016). Sample surveys were conducted by an expert pathologist.
- (2) *Otsu's method* (Fan and Zhao 2007; Vala and Baxi 2013; Xu et al. 2011; Zhang and Hu 2008; Zhu et al. 2009). We used Otsu's method to assign the pixels of each image to one of two classes. A pixel of high intensity above the Otsu threshold was assigned to the class "follicular". A pixel of low intensity below the Otsu threshold was assigned to the class "non-follicular". A region growing with a four-point neighborhood yielded connected components of follicular pixels. Each individual connected component identified the region of a lymphoid follicle. We filtered the regions for size and shape. Regions were deleted if they were not compact enough, i.e., if their form factors

$$F^* = \frac{4 \times \pi \times \text{Area}}{\text{Perimeter}^2} \quad (1)$$

were smaller than the threshold of 0.71. Too small and too large regions were deleted. Regions were accepted only if they covered an area between 340 pixels and 42,000 pixels. We implemented the Otsu pipeline by applying the algorithms of the Python libraries *scipy* (Virtanen et al. 2020) and *scikit-image* (Van der Walt et al. 2014). The Otsu pipeline included the preprocessing steps of the section Preprocessing.

- (3) *Watershed segmentation* (Ng et al. 2006; Serra 1983; Soille 2013). We expanded the *Otsu* pipeline with the watershed algorithm of

Neubert and Protzel (Neubert and Protzel 2014) to improve the gold standard pipeline with the dissociation of clustered follicles. We identified local maxima inside the thresholded image and defined each peak as the centroid of one single follicle. We used the watershed segmentation to separate the covered regions. The separated regions were filtered for size and shape. The regions were deleted if their form factors drop below 0.71. Regions outside the range of 340 pixels and 42,000 pixels were deleted. We expanded the *Otsu* pipeline with existing implementations of the peak identification algorithm and watershed algorithm provided by the Python libraries *scipy* (Virtanen et al. 2020) und *scikit-image* (Van der Walt et al. 2014).

We accepted the manual annotation as gold standard reference to measure the quality of the results of Otsu's method, the watershed

segmentation, and the shock filter approach. The quality of the three methods were evaluated by comparisons pixel by pixel.

The Zijdenbos similarity index (ZSI), also called Dice coefficient or F1 score

$$ZSI = 2 \frac{|V \cap R|}{|V| + |R|} \quad (2)$$

is an effective measure of the pixel-wise quality of an annotation (Belkacem-Boussaid et al. 2011). V and R are the sets of follicular pixels in the annotation under consideration and in the reference gold standard, respectively. The ZSI is zero for disjunct sets, $V \cap R = \emptyset$, and becomes one for perfect agreement, $V = R$. A ZSI better than 70% is considered as excellent agreement (Bartko 1991; Zijdenbos et al. 1994).

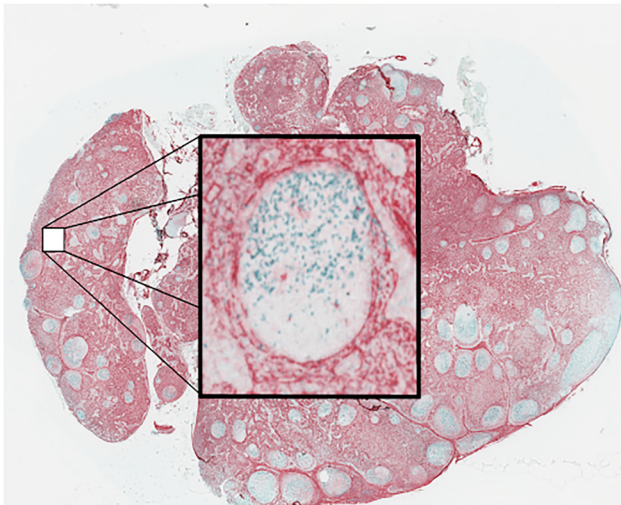


Figure 4: Tissue section of a human lymph node. The tissue section was primarily stained against PD1 and secondarily against actin. PD1 is used to visualize T cells and their ligands, whereas the secondary staining displays underlying sinus cells. A sample section of a lymphoid follicle is magnified. The size of the image is in round numbers 35.2×28.5 mm.

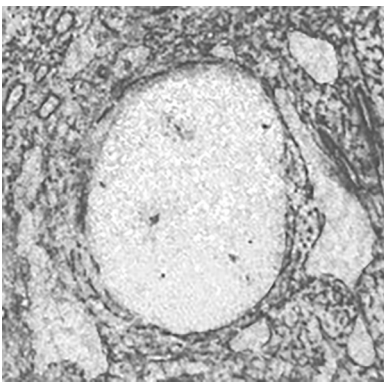


Figure 5: Subsection of a grayscale image of the actin staining. The bright ellipsoid object in the tissue is a lymphoid follicle. The size of the image is in round numbers 1.5×1.5 mm.

Preprocessing

We identified the tissue (ROI) by thresholding the grayscale image. For noise suppression and contrast enhancement, we adopted the concept of preprocessing of Belean et al. 2015 (Belean et al. 2015a). We applied a Gaussian filter with $\sigma = 5$ pixel $\approx 39.5 \mu\text{m}$ (Deng and Cahill 1993). We transformed the intensity values of the grayscale image, $I_{\text{raw}}(x, y)$,

$$I'(x, y) = \log_2(1 + I_{\text{raw}}(x, y)) \quad (3)$$

and determined the minimum and maximum intensity values, I'_{\min} and I'_{\max} , in each scaled image.

A linear transformation

$$\tilde{I}(x, y) = \frac{I'(x, y) - I'_{\min}}{I'_{\max} - I'_{\min}} \times 255 \quad (4)$$

scaled the intensity values of each image in the range of 0–255. Finally, a morphological operation of closing (Haralick et al. 1987) with a 3×3 pixel window eliminated interrupted surface structures. This sequence of standard methods resulted in an image optimized for noise and contrast. In Figure 6, the gray line depicts the original intensity profile

$$i_{\text{raw}}(x) = I_{\text{raw}}(x, y = 0.87 \text{ mm}) \quad (5)$$

along a constant y -value in Figure 5. The black line depicts the corresponding intensity profile

$$i(x) = I(x, y = 0.87 \text{ mm}) \quad (6)$$

after preprocessing. $I(x, y)$ represents the field of intensities of the preprocessed image. After preprocessing, the range of intensities is stretched, the profile is smooth, and distinct maxima and minima become visible. Notice that, the plateau of high intensities in the region of $0.4 \text{ mm} \leq x \leq 1 \text{ mm}$ corresponds to the bright area of the follicular region inside the center of Figure 5.

Detection of transition points by shock filter

In the next step, we want to determine plateaus of high intensity in profiles, such as shown in the black line of Figure 6. The border of a plateau is characterized by points at which the intensity drops significantly. We call such a point of intensity drop a transition point.

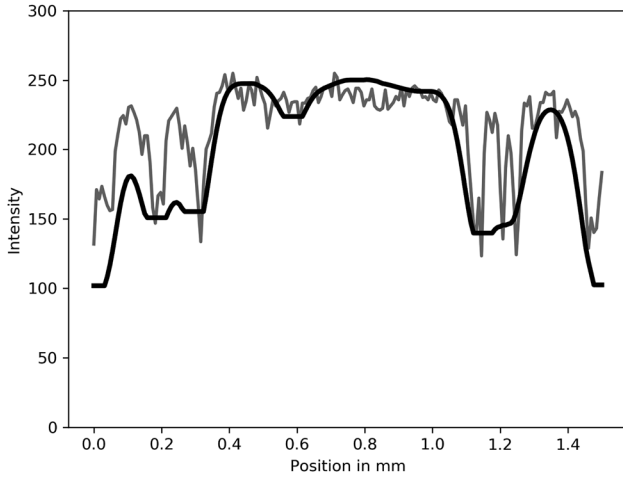


Figure 6: Intensity profile before and after the preprocessing. The gray line shows an initial intensity profile of the subsection in Figure 5, $i_{\text{raw}}(x) = I_{\text{raw}}(x, y = 0.87 \text{ mm})$. The black line depicts the intensity profile after preprocessing, $i(x) = I(x, y = 0.87 \text{ mm})$. The intensity profile is smoothed, and plateaus of high and low intensities get pronounced.

For determining transition points, we used two families of intensity profiles. Each intensity profile was averaged over slices of 10 pixels, i.e., $78.944 \mu\text{m}$,

$$v_{y=a}(x) := \frac{1}{10} \sum_{y=a}^{a+10} I(x, y) \quad (7)$$

and

$$h_{x=b}(y) := \frac{1}{10} \sum_{x=b}^{b+10} I(x, y), \quad (8)$$

respectively. The usage of averaged vertical and horizontal profiles, $v_{x=a}(y)$ and $h_{y=b}(x)$, resulted in a reduction of noise and a speedup of the computation. The intensity profiles $v_{x=a}(y)$ and $h_{y=a}(x)$ were located at the grid positions $a, b = i \times 5 \text{ pixel} \approx i \times 39.5 \mu\text{m}$ for $i = 0, 1, 2, \dots, n$.

We applied a shock filter to the profiles $v_{x=a}(y)$ and $h_{y=a}(x)$. The shock filter is a finite number of iterations:

$$p^{t+1} = p^t - \text{sign}(\Delta p^t) |\nabla p^t| \quad \text{with } t = 0, 1, 2, 3, 4 \dots t_{\text{max}} \quad (9)$$

for each of the starting conditions, $p^0 = v_{x=a}(y)$ and $p^0 = h_{y=a}(x)$. The Laplace operator was spatially discretized by the three-point estimation:

$$\Delta p(k) \doteq p(k+1) - 2p(k) + p(k-1). \quad (10)$$

The Nabla operator was computed as the minimum of the left and right derivative:

$$\nabla p(k) \doteq \min(\Delta_l p(k), \Delta_r p(k)). \quad (11)$$

If the left and right derivatives had the same sign, we estimated them by the two-point estimates:

$$\begin{aligned} \nabla_l p(k) &:= p(k+1) - p(k) \\ \nabla_r p(k) &:= p(k) - p(k-1). \end{aligned} \quad (12)$$

For different signs of the left and right derivatives, we replaced the iteration by the identity operation:

$$p^{t+1}(i) = p^t(i). \quad (13)$$

We performed $t_{\text{max}} = 20$ iterations. Figure 7 shows an intensity profile before (gray line) and after (black line) the application of the shock filter iterations. The iterative application of the shock filter produces a piecewise constant function. A smooth transition from low to high intensity in the gray line becomes an abrupt jump in the intensity of the black line. The position of the jump in the profile becomes well-defined after the application of the shock filter iterations. We call the position of an intensity jump a transition point. The jumps of intensity at transition points may differ in their heights. We computed the increment of intensity ΔI for the transition points of all vertical and horizontal profiles, $v_{x=a}(y)$ and $h_{y=b}(x)$, of an image. We determined a threshold S :

$$S = Q_3 + 1.5 \times IQR, \quad (14)$$

where Q_3 and IQR are the third quartile and the interquartile range, respectively, of the increments $\Delta I_i, i = 1, 2, \dots, k$. For each vertical and horizontal profile $v_{x=a}(y)$ and $h_{y=b}(x)$, we computed pairs of start and end values, $(y_{\text{start}}, y_{\text{end}})$ and $(x_{\text{start}}, x_{\text{end}})$, respectively. The start and end values fulfill the conditions of a plateau of high intensity

$$v_{x=a}(y_{\text{start}}) - v_{x=a}(y_{\text{start}} - 1) > S, \quad (15)$$

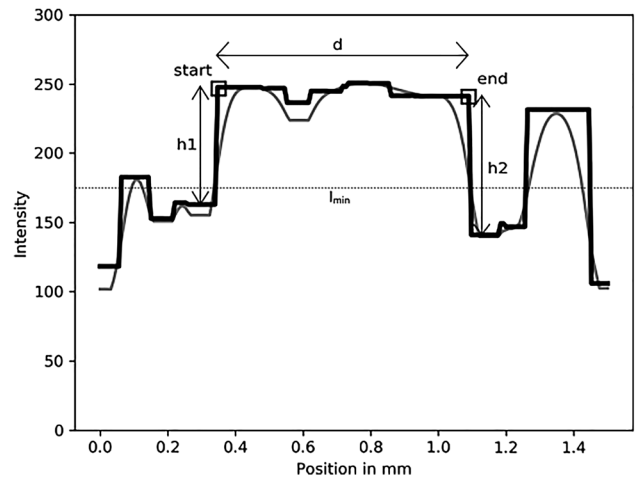


Figure 7: Evolution of an intensity profile through the application of 20 shock filter iterations. Here, we apply the shock filter to the intensity profile, $i(x) = I(x, y = 0.87 \text{ mm})$. The two lines show the intensity profile before (gray) and after (black) the application of the shock filter. The generation of constant functions out of a smooth intensity profile through the application of shock filter iterations is demonstrated. The boxes *start* and *end* represent an example for a transition point pair. Validation: The height h_1 of the initiating intensity jump, as well as the height h_2 of the closing intensity jump, is higher than the threshold $S = 42$. In the range of *start* and *end*, no intensity value drops below $I_{\text{min}} = 175$. The distance d between *start* and *end* is in between 0.018 and 1.579 mm .

$$v_{x=a}(y_{\text{end}}) - v_{x=a}(y_{\text{end}} + 1) > S, \quad (16)$$

$$\forall y \in [y_{\text{start}}, y_{\text{end}}] : v_{x=a}(y) > I_{\text{min}}, \quad (17)$$

$$118 \mu\text{m} < y_{\text{end}} - y_{\text{start}} < 1579 \mu\text{m}, \quad (18)$$

and

$$h_{y=b}(x_{\text{start}}) - h_{y=b}(x_{\text{start}} - 1) > S, \quad (19)$$

$$h_{y=b}(x_{\text{end}}) - h_{y=b}(x_{\text{end}} - 1) > S, \quad (20)$$

$$\forall x \in [x_{\text{start}}, x_{\text{end}}] : h_{y=b}(x) > I_{\text{min}}, \quad (21)$$

$$118 \mu\text{m} < x_{\text{end}} - x_{\text{start}} < 1579 \mu\text{m}, \quad (22)$$

The table in Figure 10 gives the threshold S and the minimal intensity value of the plateau, I_{min} , for each image. Each pair of values $(y_{\text{start}}, y_{\text{end}})$ of the profile $v_{x=a}(y)$ yields a pair of points $(s = (a + 5, y_{\text{start}}), e = (a + 5, y_{\text{end}}))$ in the 2D image. Each pair of values $(x_{\text{start}}, x_{\text{end}})$ of the profile $h_{y=b}(x)$ yields a line $(s' = (x_{\text{start}}, b + 5), e' = (x_{\text{end}}, b + 5))$ in the 2D image. For an image, we computed a set of vertical lines $V = \{(s_1, e_1), (s_2, e_2), \dots, (s_n, e_n)\}$ and a set of horizontal lines $H = \{(s'_1, e'_1), (s'_2, e'_2), \dots, (s'_m, e'_m)\}$.

Segmentation of lymphoid follicles

The start point and the end point of a line in V and H are located on the borderline of the region of high intensity. A vertical line and a horizontal line with an intersection run across the same plateau of high intensity. We aggregated the lines in V and H in clusters

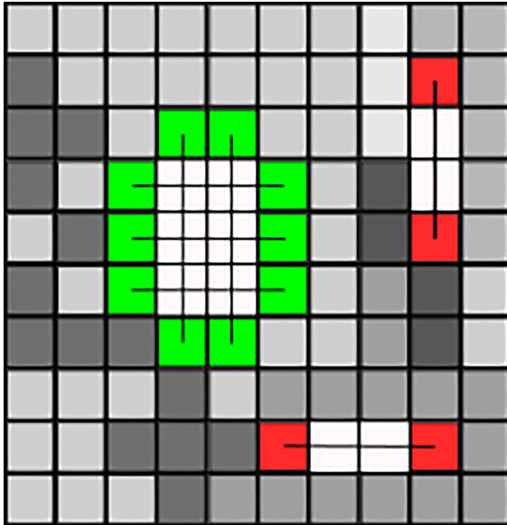


Figure 8: Scheme of the clustering algorithm. One box represents one pixel in an image. If two boxes are linked, they represent potential transition point pairs. Transition point pairs with crossing links get clustered. If the convex hull of a cluster fulfills predefined mathematical conditions, it will be accepted as a valid lymphoid follicle, here shown in green. A transition point pair without an intersecting analog gets rejected, here shown in red.

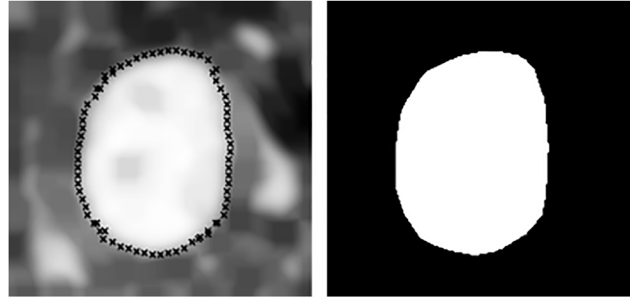


Figure 9: Computation of the convex hull. The left part displays the sample image with marked transition point pairs. The right part is the resulting binary mask of the lymphoid follicle segmentation based on the sample section.

$C_i \subset V \cup H$, $i = 1, 2, 3, \dots, n$. A vertical line $v \in V$ and a horizontal line $h \in H$ with an intersection were assigned to the same cluster C_i . Each cluster C_i represents a plateau of high intensity and is a candidate of a follicular region. Clusters were deleted if they contain only a single line, i.e., an isolated line without any intersection to another line, see Figure 8.

For each cluster C_i , we computed the convex hull (Barber et al. 1996) and its shape parameters. Clusters were deleted if the form factor of their convex hulls were smaller than the threshold of 0.71. Too small and too large clusters were deleted. Clusters were accepted only if the area of their convex hulls is in the range of 340 pixels and 42,000 pixels. Note that, the thresholds for the minimal and maximal size corresponds to circles with diameters d in the range of $0.104 \text{ mm} \leq r \leq 1.024 \text{ mm}$. The left part of Figure 9 shows the transition points of a cluster. Right part of Figure 9 shows the binary mask of the convex hull of a cluster.

Supplementary

The implementation is available under the following link: <https://sourceforge.net/projects/fldetect/>.

Author contributions: All the authors have accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: This investigation was supported by the Wilhelm Sander Stiftung No. 2018.101.1, BMBF, COMPLS2-087, Patho234.

Conflict of interest statement: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Appendix



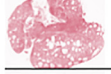



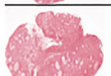

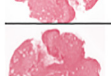
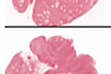
| Thumbnail | ID | Width (Pixel) | Height (Pixel) | S | I_{min} | Q_3 | IQR |
|---|----|---------------|----------------|----|-----------|-------|-----|
|  | 1 | 2171 | 1763 | 42 | 205 | 19 | 15 |
|  | 2 | 2201 | 1748 | 38 | 160 | 18 | 14 |
|  | 3 | 2141 | 1742 | 42 | 172 | 19 | 15 |
|  | 4 | 2231 | 1804 | 42 | 175 | 19 | 15 |
|  | 5 | 1903 | 1757 | 44 | 149 | 18 | 16 |
|  | 6 | 1814 | 1834 | 43 | 130 | 20 | 16 |
|  | 7 | 1903 | 1803 | 45 | 151 | 21 | 17 |
|  | 8 | 1933 | 1798 | 40 | 158 | 19 | 15 |
|  | 9 | 1993 | 1783 | 41 | 141 | 19 | 15 |
|  | 10 | 1993 | 1845 | 42 | 141 | 19 | 15 |

Figure 10: Dataset and the corresponding values per image. The image size is given in pixels, as the algorithm-specific values S , I_{min} , Q_3 and IQR per image. S defines the threshold of an intensity jump of interest. I_{min} defines the minimal intensity value of a plateau of high intensity. Q_3 is the third quartile of the boxplot statistic of all intensity jumps. IQR is the interquartile range of the boxplot statistic of all intensity jumps.

References

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al. (2015). TensorFlow: large-scale machine learning on heterogeneous systems. Software available at: tensorflow.org.
- Barber, C.B., Dobkin, D.P., and Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Trans. Math Software* 22: 469–483.
- Barthélemy, M. (2011). Spatial networks. *Phys. Rep.* 499: 1–101.
- Bartko, J.J. (1991). Measurement and reliability: statistical thinking considerations. *Schizophr. Bull.* 17: 483–489.
- Belean, B., Borda, M., Ackermann, J., Koch, I., and Balacescu, O. (2015a). Unsupervised image segmentation for microarray spots with irregular contours and inner holes. *BMC Bioinf.* 16: 412.
- Belean, B., Terebes, R., and Bot, A. (2015b). Low-complexity PDE-based approach for automatic microarray image processing. *Med. Biol. Eng. Comput.* 53: 99–110.
- Belkacem-Boussaid, K., Samsi, S., Lozanski, G., and Gurcan, M.N. (2011). Automatic detection of follicular regions in H&E images using iterative shape index. *Comput. Med. Imag. Graph.* 35: 592–602.
- Broyde, A., Boycov, O., Strenov, Y., Okon, E., Shpilberg, O., and Bairey, O. (2009). Role and prognostic significance of the Ki-67 index in non-Hodgkin's lymphoma. *Am. J. Hematol.* 84: 338–343.
- De Berg, M., Cheong, O., Van Kreveld, M., and Overmars, M. (2008). Computational geometry: introduction. In: *Algorithms and applications*. Springer, Berlin, Heidelberg, https://doi.org/10.1007/978-3-540-77974-2_1.
- Deng, G. and Cahill, L.W. (1993). An adaptive Gaussian filter for noise reduction and edge detection. In: *Nuclear science symposium and medical imaging conference, 1993, 1993 IEEE conference record*. IEEE, New York, pp. 1615–1619.
- Doi, K. (2007). Computer-aided diagnosis in medical imaging: historical review, current status and future potential. *Comput. Med. Imag. Graph.* 31: 198–211.
- Dominguez, R. and Holmes, K.C. (2011). Actin structure and function. *Annu. Rev. Biophys.* 40: 169–186.
- Dutta, A., Gupta, A., and Zissermann, A. (2016). VGG image annotator (VIA). Version: 2.0.0, Available at: <http://www.robots.ox.ac.uk/vgg/software/via/> (Accessed 07 September 2018).
- Dutta, A. and Zisserman, A. (2019). The VIA annotation software for images, audio and video. In: *Proceedings of the 27th ACM international conference on multimedia, MM '19*. ACM, New York, NY, USA.
- Fan, J.-l. and Zhao, F. (2007). Two-dimensional Otsu's curve thresholding segmentation method for gray-level images. *Acta Electron. Sin.* 35: 751.
- Goyal, B., Dogra, A., Agrawal, S., Sohi, B., and Sharma, A. (2020). Image denoising review: from classical to state-of-the-art approaches. *Inf. Fusion* 55: 220–244.
- Guibas, L.J., Knuth, D.E., and Sharir, M. (1992). Randomized incremental construction of Delaunay and Voronoi diagrams. *Algorithmica* 7: 381–413.
- Haralick, R.M., Sternberg, S.R., and Zhuang, X. (1987). Image analysis using mathematical morphology. In: *IEEE transactions on pattern analysis and machine intelligence*. IEEE, New York.
- Huo, Y., Liu, T., Liu, H., Ma, C.Y., and Wang, X.Z. (2016). In-situ crystal morphology identification using imaging analysis with application to the L-glutamic acid crystallization. *Chem. Eng. Sci.* 148: 126–139.
- Kendall, D. G. (1989). A survey of the statistical theory of shape. *Stat. Sci.* 148: 87–99.
- Money, J.H. and Kang, S.H. (2008). Total variation minimizing blind deconvolution with shock filter reference. *Image Vis Comput.* 26: 302–314.
- Neubert, P. and Protzel, P. (2014). Compact watershed and preemptive slic: on improving trade-offs of superpixel segmentation algorithms. In: *2014 22nd international conference on pattern recognition*. IEEE, New York, pp. 996–1001.
- Ng, H., Ong, S., Foong, K., Goh, P., and Nowinski, W. (2006). Medical image segmentation using K-means clustering and improved watershed algorithm. In: *2006 IEEE southwest symposium on image analysis and interpretation*. IEEE, New York, pp. 61–65.

- Nosanchuk, J.S. and Schintzier, B. (1969). Follicular hyperplasia in lymph nodes from patients with rheumatoid arthritis. a clinicopathologic study. *Cancer* 24: 343–354.
- Osher, S. and Rudin, L.I. (1990). Feature-oriented image enhancement using shock filters. *SIAM J. Numer. Anal.* 27: 919–940.
- Oswald, M.S., Wurzel, P., and Hansmann, M.-L. (2019). 3D analysis of morphological alterations of the fibroblastic reticular cells in reactive and neoplastic human lymph nodes. *Acta Histochem.* 121: 769–775.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., et al. (2011). Scikit-learn: machine learning in python. *J. Mach. Learn. Res.* 12: 2825–2830.
- Ruifrok, A.C. and Johnston, D.A. (2001). Quantification of histochemical staining by color deconvolution. *Anal. Quant. Cytol. Histol.* 23: 291–299.
- Serra, J. (1983). *Image analysis and mathematical morphology*. Academic Press, Inc, Cambridge, Massachusetts.
- Soille, P. (2013). *Morphological image analysis: principles and applications*. Springer Science & Business Media, Berlin, Heidelberg.
- Tosta, T.A., de Faria, P.R., Neves, L.A., and do Nascimento, M.Z. (2018). Fitness functions evaluation for segmentation of lymphoma histological images using genetic algorithm. In: *International conference on the applications of evolutionary computation*. Springer, Berlin, Heidelberg, pp. 47–62.
- Tosta, T.A.A., Faria, P.R., Neves, L.A., and do Nascimento, M.Z. (2017). Computational method for unsupervised segmentation of lymphoma histological images based on fuzzy 3-partition entropy and genetic algorithm. *Expert Syst. Appl.* 81: 223–243.
- Vala, H.J. and Baxi, A. (2013). A review on Otsu image segmentation algorithm. *Int. J. Adv. Res. Comput. Eng. Technol.* 2: 387–389.
- Van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Goullart, E., and Yu, T. (2014). scikit-image: image processing in python. *PeerJ* 2: e453.
- Varatharajan, R., Manogaran, G., and Priyan, M. (2018). A big data classification approach using LDA with an enhanced SVM method for ECG signals in cloud computing. *Multimed. Tool. Appl.* 77: 10195–10215.
- Victoria, G.D. and Nussenzweig, M.C. (2012). Germinal centers. *Annu. Rev. Immunol.* 30: 429–457.
- Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods* 17: 261–272.
- Weiss, L.M. and O'malley, D. (2013). Benign lymphadenopathies. *Mod. Pathol.* 26: S88.
- Willard-Mack, C.L. (2006). Normal structure, function, and histology of lymph nodes. *Toxicol. Pathol.* 34: 409–424.
- Xu, X., Xu, S., Jin, L., and Song, E. (2011). Characteristic analysis of Otsu threshold and its applications. *Pattern Recogn. Lett.* 32: 956–961.
- Zhang, J. and Hu, J. (2008). Image segmentation based on 2D Otsu method with histogram analysis. In: *2008 international conference on computer science and software engineering*, Vol. 6. IEEE, New York, pp. 105–108.
- Zhu, N., Wang, G., Yang, G., and Dai, W. (2009). A fast 2D Otsu thresholding algorithm based on improved histogram. In: *2009 Chinese conference on pattern recognition*. IEEE, New York, pp. 1–5.
- Zijdenbos, A.P., Dawant, B.M., Margolin, R.A., and Palmer, A.C. (1994). Morphometric analysis of white matter lesions in MR images: method and validation. *IEEE Trans. Med. Imag.* 13: 716–724.