


Automated annotations of epithelial cells and stroma in hematoxylin–eosin–stained whole–slide images using cytokeratin re–staining

Tomáš Brázdil¹ , Matej Gallo¹ , Rudolf Nenutil², Andrej Kubanda¹, Martin Toufar¹ and Petr Holub^{3*} 

¹Faculty of Informatics, Masaryk University, Brno, Czech Republic

²Department of Pathology, Masaryk Memorial Cancer Institute, Brno, Czech Republic

³Institute of Computer Science, Masaryk University, Brno, Czech Republic

*Correspondence to: Petr Holub, Institute of Computer Science, Masaryk University, Botanická 68a, 602 00 Brno, Czech Republic.
E-mail: hopet@ics.muni.cz

Abstract

The diagnosis of solid tumors of epithelial origin (carcinomas) represents a major part of the workload in clinical histopathology. Carcinomas consist of malignant epithelial cells arranged in more or less cohesive clusters of variable size and shape, together with stromal cells, extracellular matrix, and blood vessels. Distinguishing stroma from epithelium is a critical component of artificial intelligence (AI) methods developed to detect and analyze carcinomas. In this paper, we propose a novel automated workflow that enables large-scale guidance of AI methods to identify the epithelial component. The workflow is based on re-staining existing hematoxylin and eosin (H&E) formalin-fixed paraffin-embedded sections by immunohistochemistry for cytokeratins, cytoskeletal components specific to epithelial cells. Compared to existing methods, clinically available H&E sections are reused and no additional material, such as consecutive slides, is needed. We developed a simple and reliable method for automatic alignment to generate masks denoting cytokeratin-rich regions, using cell nuclei positions that are visible in both the original and the re-stained slide. The registration method has been compared to state-of-the-art methods for alignment of consecutive slides and shows that, despite being simpler, it provides similar accuracy and is more robust. We also demonstrate how the automatically generated masks can be used to train modern AI image segmentation based on U-Net, resulting in reliable detection of epithelial regions in previously unseen H&E slides. Through training on real-world material available in clinical laboratories, this approach therefore has widespread applications toward achieving AI-assisted tumor assessment directly from scanned H&E sections. In addition, the re-staining method will facilitate additional automated quantitative studies of tumor cell and stromal cell phenotypes.

Keywords: U-Net; artificial intelligence; digital pathology; H&E; immunohistochemistry; deep learning; tissue registration

Received 5 March 2021; Revised 22 September 2021; Accepted 29 September 2021

No conflicts of interest were declared.

Introduction

There has been a rapid onset of deep learning applications in pathology. Although achieving the full potential of artificial intelligence (AI)-assisted histopathology is some distance away, research in this area is expanding and providing promising results. Progress in the digitization of histological samples has resulted in the production of so-called ‘virtual slides’ or ‘whole-slide images’ (WSI). The trend to digitize the whole laboratory workflow [1,2] has the potential to

substantially change the profession, especially when practically useful AI assistance algorithms are developed, validated, and implemented [3,4] to support pathologists in routine tasks, such as finding carcinoma micrometastases in lymph nodes or small foci of prostate carcinoma in core biopsies, or in tasks requiring quantification, such as precise evaluation of different antigens by immunostaining, location, and counting of different cell populations [5,6]. Unfortunately, developing AI algorithms requires large volumes of learning and validation data. For supervised

learning, these are created mostly by manual annotation of histological images, requiring substantial time and effort from experienced pathologists, representing a roadblock for their development, validation, and implementation.

Solid tumors of epithelial origin (carcinomas) represent a major part of the workload in diagnostic histopathology laboratories. These tumors are composed of malignant epithelial cells arranged in more or less cohesive clusters of different sizes and shapes, together with admixed stroma including fibroblasts, adipocytes, endothelium, and inflammatory cells, and variable amounts of extracellular matrix. Various types of inflammatory cells may also infiltrate the epithelial islands.

The manual identification of tumor areas is often a relatively simple process for a trained pathologist and, by default, surrounding stromal areas are also identified. Such manually created data have been useful for developing AI applications, especially in the diagnosis of malignancy *per se*, tumor typing, and prognostication. On the other hand, there are applications that require precise identification and annotation of carcinoma cells and stromal cells within tumor islands, at the tumor/stroma interface, and more distant stromal cell populations devoid of carcinoma cells (e.g. counting mitoses, nuclear morphometry, enumerating intraepithelial and stromal lymphocytes, and evaluating the topographical distributions of specific inflammatory cell populations). These applications require the identification of an ‘epithelial mask’ that accurately delineates the areas of epithelial cells. Due to the complexity of carcinoma structures, manual creation of a precise epithelial mask is challenging if representative areas of tumor are to be identified for a large number of samples, such as required for AI training. This differentiation of cancer cells from stroma is recognized as an important problem, to which solutions using staining of consecutive tissue sections have been proposed [6,7]. A recent paper [8] demonstrated an approach to manual annotations supervised by an epithelial mask based on re-staining of hematoxylin and eosin (H&E)-stained slides using immunohistochemistry (IHC) to visualize cytokeratins, cytoskeletal components specific to epithelial cells.

Aims and scope

The idea behind our project was to develop a simple, rapid, and robust working procedure that is able to create precise epithelial masks in WSIs and to use these masks to train a neural network. Once trained, the procedure can then create ‘virtual’ epithelial masks in

previously unseen H&E-stained WSIs. These virtual masks have immediate utility in AI-mediated diagnostic applications in cancer detection and can also be employed to co-localize carcinoma epithelium/stromal regions to a second WSI produced for any additional marker, obtained by de-staining the original H&E section and re-staining for the marker of choice. In this way, the procedure can be employed both for training AI-based methods with minimal expert annotation to simplify and speed up the learning process, and for further sophisticated analysis of multiple parameters of interest both clinically and in research.

Materials and methods

Experimental design

The study includes (a) development of a method for automated pan-cytokeratin re-staining of previously digitized H&E-stained sections; (b) an automated method for aligning re-stained pan-cytokeratin immunohistological images to the original H&E images, in spite of shifts and nonlinear tissue distortions that occur during re-staining; (c) automated mask generation based on thresholding; and (d) the demonstration that these masks serve as a guide for machine learning to detect and label epithelium in previously unseen H&E-stained tissue sections of carcinomas.

We decided to use tissue microarrays (TMAs) as the source material in phases (a)–(c). These compound blocks containing multiple tissue cores collected from routine diagnostic samples ensure adequate numbers of tumors and with a range of growth patterns for training and validation. In addition, that these samples were obtained and processed under real-world clinical conditions, rather than idealized and highly controlled model systems, ensures easy transition into the clinical setting. In phase (d), we used both TMA sections and whole block tissue sections to test whether the trained method independently identifies tumor/stroma areas in H&E sections, applicable for automated quantitation of additional immunohistochemical or molecular pathology markers of interest for diagnosis and research.

Material

The source material originated from diagnostic cases evaluated at the Department of Pathology, MMCI. All patients provided written informed consent to use their leftover material for research. Staining of breast carcinomas was performed on residual excess slides from

Table 1. Details of breast cancer samples.

Grade	Grade 1	Grade 2	Grade 3		
	1	6	10		
Morphology	Lobular	NST	Mucinous	Micropapillary	Medullary
	1	13	1	1	1
Estrogen receptor	Positive	Negative			
	13	4			
HER2	Positive	Negative			
	4	13			
Format/use	TMA core	Full section scan			
Learning	12	0			
Evaluation	0	5			

NST, No Special Type / No Specific Type.

IHC quality control procedures, using two TMAs representing 12 tumor samples in total. Patient details are provided in Table 1. These tumors included the most frequent histological patterns from dissociated lobular carcinoma, through trabecular and solid, to medullary growth. The five H&E-stained breast cancer WSIs used for performance evaluation are part of routine case documentation in the MMCI Biobank and were selected randomly from the collection. The colon carcinoma TMA samples represent residual material from a retrospective immunohistochemical study of consecutive colon carcinomas. Patient details are given in Table 2. Each of the five TMAs originally contained 24 tumor samples, each represented by two 1.4 mm cores. After evaluation of H&E- and cytokeratin-stained scans, this number was reduced to a total of 85 tumors represented by 141 cores. Four TMAs were used for U-Net learning and the fifth TMA was used

to evaluate the performance of the trained network. These TMAs were manufactured by a homemade tissue arrayer.

Additional scans of full-block H&E sections of one breast carcinoma and three colon carcinomas were selected from the collection of MMCI biobank to evaluate the efficacy of mask generation to demonstrate the U-Net performance online (Figure 1). The materials used at each step of method development are summarized in Figure 2 and details are provided in the supplementary material.

Re-staining protocol

Sections of 5 µm thickness were collected onto positively charged slides for IHC (TOMO, Matsunami Glass IND LTD, Osaka Japan). The H&E staining protocol was performed according to the standard

Table 2. Details of colon cancer samples.

Age	Range	Mean	Median		
	30–92	65.4	66.5		
Sex	Male	Female			
	56	32			
Location	Right colon	Left colon	Indetermined		
	35	47	6		
Clinical stage	Stage I	Stage II	Stage III	Stage IV	
	15	27	35	11	
Grade	Grade 1	Grade 2	Grade 3		
	23	46	19		
Morphology	Tubular	Tubopapillary	Cribriform	Medullary	Dissociated
	16	28	37	5	2
Format/use	TMA core	Full section scan			
Learning	67	0			
Evaluation	18	3			

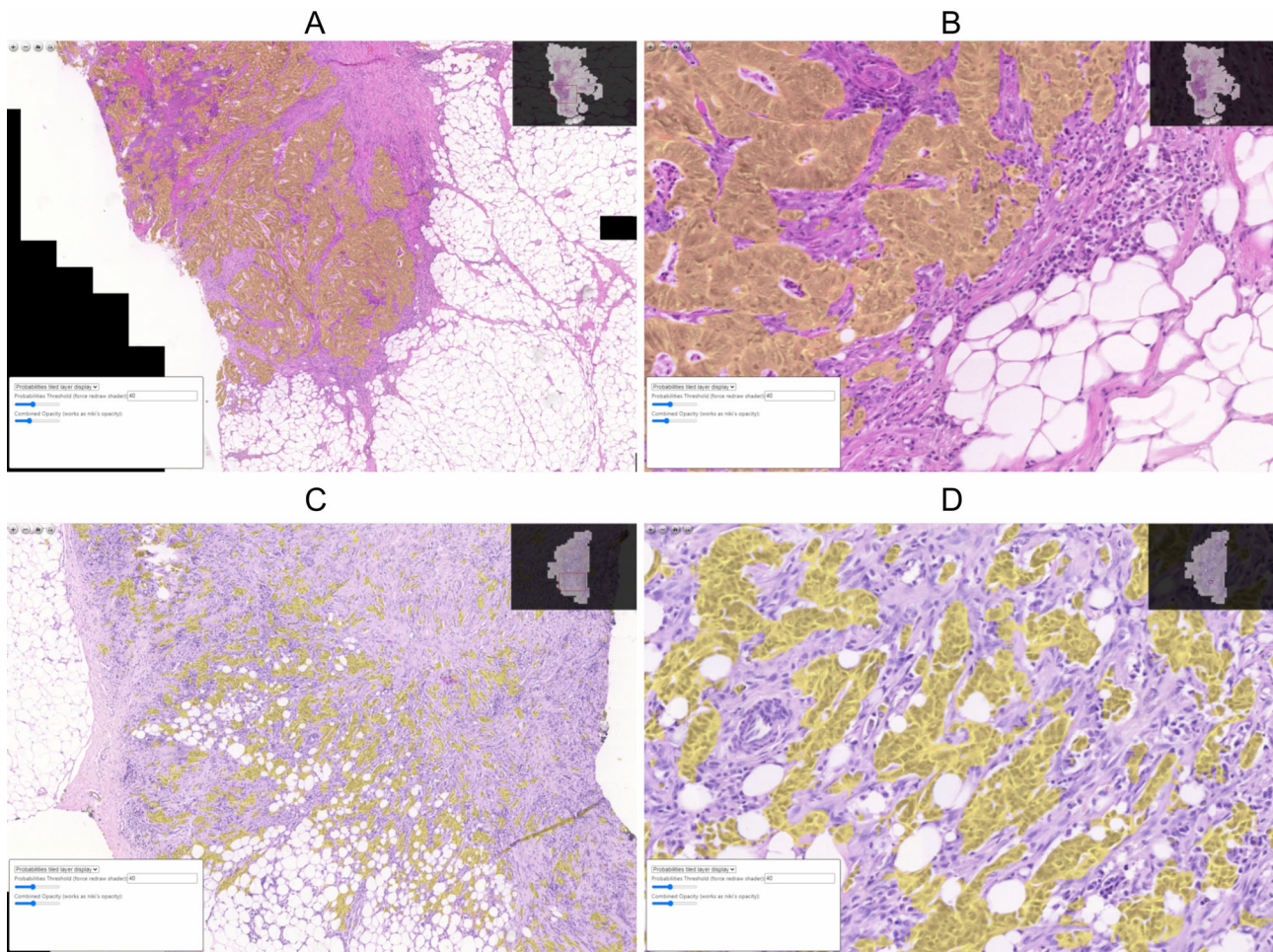


Figure 1. Illustration of WSIs available online including U-Net computed epithelial mask in colon (A–C) and breast (D) carcinoma. The links for online viewing include: http://rationai-vis.ics.muni.cz/iipmooviewer-jiri/matej/index.html?image=Biobanka/rgb/2021_0188-01-T.tif&layer=Biobanka/seg/2021_0188-01-T.tif, http://rationai-vis.ics.muni.cz/iipmooviewer-jiri/matej/index.html?image=Biobanka/rgb/2021_1546-01-T.tif&layer=Biobanka/seg/2021_1546-01-T.tif, http://rationai-vis.ics.muni.cz/iipmooviewer-jiri/matej/index.html?image=Biobanka/rgb/2021_1583-02-T.tif&layer=Biobanka/seg/2021_1583-02-T.tif, and http://rationai-vis.ics.muni.cz/iipmooviewer-jiri/matej/index.html?image=Biobanka/rgb/2021_1768-07-T.tif&layer=Biobanka/seg/2021_1768-07-T.tif.

laboratory protocol using a Leica Autostainer XL5000 (Leica Biosystems, Wetzlar, Germany), followed by cover glass mounting with Solakryl BMX (Draslovka, Kolín, Czech Republic) on a Leica CV500 mounting machine. Details are given in the supplementary material. The slides were left for 1 h to dry and then scanned using a Panoramic[®] MIDI (3DHitech, Budapest, Hungary) with $\times 20$ objective lens at a resolution $0.172 \mu\text{m}/\text{pixel}$. The WSIs were uncompressed in PNG inside Mirax format. After scanning, slides were placed in xylene to remove the coverslips, washed in xylene to remove the remaining mountant, and re-hydrated through graded alcohols before returning to water. IHC with a cocktail of cytokeratin antibodies was performed in a Dako Autostainer Link

48 (Agilent, Santa Clara, CA, USA) using a standard staining procedure including high pH antigen retrieval. The cytokeratin antibody cocktail used to stain breast carcinoma samples consisted of Ba-17, DC10, C-11, and C-43 antibodies (Masaryk Memorial Cancer Institute, Brno, Czech Republic). A pan-cytokeratin cocktail AE1/AE3 (Agilent) was used for colon cancer samples. Antibody binding was detected with Envision HRP-labeled polymer secondary antibody with DAB as chromogen (Agilent), and nuclei were counterstained with hematoxylin before dehydration and mounting. Details are given in the supplementary material. The WSIs were checked in the Automated Slide Analysis Platform (ASAP) (available at <https://computationalpathologygroup.github.io/ASAP/#home>)

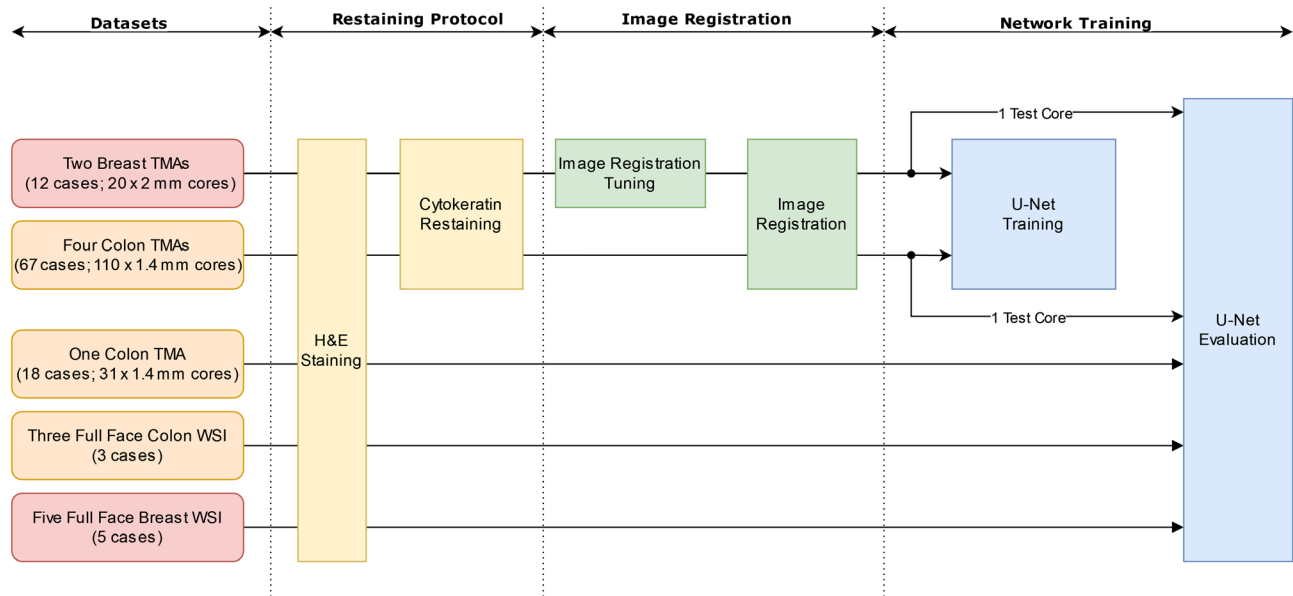


Figure 2. Flowchart of the automated annotation system for cytokeratin re-staining. Colors serve as a visual aid for related parts of the system.

and were manually annotated to exclude unacceptable cores or major artifacts from further analyses.

Dataset access

The dataset is available as raw files stored in Mirax MRXS format compatible with OpenSlide (<https://openslide.org/formats/mirax/>). Annotations used for the evaluation stage are available as XML files compatible with ASAP. The dataset is pseudonymized and access can be requested via BBMRI-ERIC European Research Infrastructure (<https://www.bbmri-eric.eu/>) by following its access policy; the request should be placed via BBMRI-ERIC Negotiator platform to Masaryk Memorial Cancer Institute.

Automated alignment of cytokeratin mask on H&E slides

Software packages and libraries used in this study are provided in the supplementary material. Automated generation of the cytokeratin mask was performed as follows: (1) the H&E and the re-stained slides are split into individual tissue cores from the TMA (Figure 3); (2) cores are registered (aligned) using one of the two methods described below; and (3) the resulting mask is created using adaptive thresholding with subsequent noise filtering. The registration method developed for step (2) utilizes a specific property of the re-staining

procedure to achieve robust results: cell nuclei centroids act as reference points and are visible in both the H&E and the re-stained tissue due to the repeated hematoxylin staining, which is standard procedure for IHC using DAB and other commonly employed chromogens. The developed method was then compared to the previously used state-of-the-art registration systems for alignment of consecutive slides [6,7] and was found both more robust and more accurate.

Nuclei-based registration

We developed a simple yet robust alignment method based on isolation of cell nuclei marked by hematoxylin in the original H&E slides and the re-stained cytokeratin slides. Description of the algorithm refers to the functions defined in algorithm 1; full pseudocode is available in the supplementary material, Section S2.

- Image processing starts with decomposition into hematoxylin and eosin, and hematoxylin and DAB channels [9,10] for each image, respectively. For optimization of separation of the DAB and hematoxylin channels, the sparse non-negative matrix factorization (SNMF) method was used [11].
- Cell nuclei centroids are identified in the hematoxylin channel (DETECT-NUCLEI function), taking into account both roundness of the shape and the

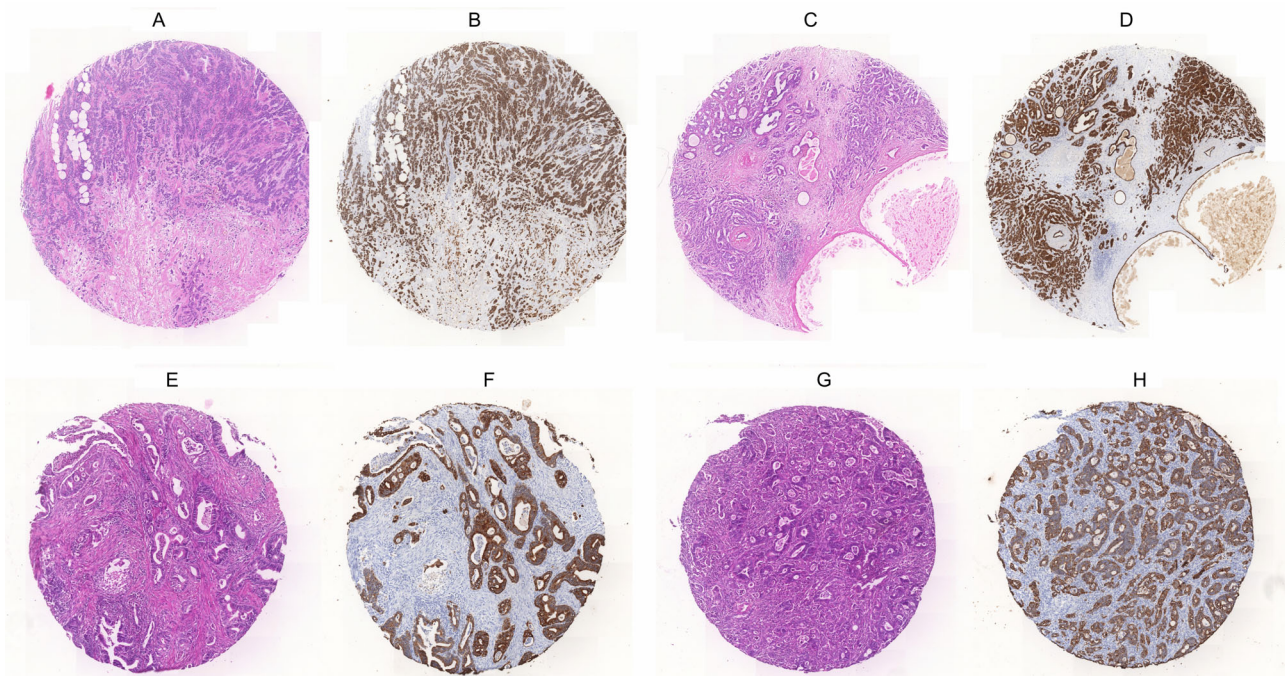


Figure 3. Corresponding H&E and cytokeratin re-stained tissue cores used as reference tissues. (A, B) Tissue core 2; (C, D) core 5; (E, F) core 7; (G, H) core 13. An overview of the full dataset is shown in supplementary material, Section S4.

minimum/maximum size, pre-specified for the given resolution.

- Matching of nuclei centroids in both images is based on searching for minimum square error when performing rigid transformations (TRANSLATION and ROTATION functions) with the gradient descent, and applying nearest neighbor match.
- Once the best possible match is identified, the matched pairs are used to determine shift vectors for nonrigid transformation of the image, which is then performed using warp transformation.

Whole-tissue-based registration

Automatic alignment of consecutive slides using different staining was tackled by the ANHIR challenge in 2019 (<https://anhir.grand-challenge.org/>) [12] and evaluated using landmark validation [13]. This competition did not provide any additional assumptions on staining and was based on consecutive slides being stained with the different methods and not the same slide being stained with both methods, as described here. Hence, the methods had to be more general without being able to use the presumptions that we were able to make for matching cell nuclei. One of the top three ranking methods, and the only one that has been

published as open source [14], was used as the basis for comparison with the method proposed in this paper.

The algorithm combines several methods and proceeds in two stages. First, the initial alignment based on rigid or similarity transformations is computed, and then non-rigid transformations finish the registration. For the initial alignment, the system automatically selects between the two methods. The first uses feature detection algorithms (SURF, SIFT, and ORB) followed by RANSAC for calculating similarity transformations between features. The second method, related to our algorithm, computes centroids in binary versions of both images (obtained by Li thresholding) and uses them to iteratively compute the desired rotation. For nonrigid transformation, the system automatically chooses between the four methods: local affine transformation with local brightness and contrast corrections, two methods based on Thirion's Demons algorithm [15], and a thin plate interpolation applied to all the good matches from the initial alignment procedure.

Automation of cytokeratin mask generation

We tested various methods of thresholding, and the isodata [16,17] and minimum methods [18,19] provided the best results based on comparison with the

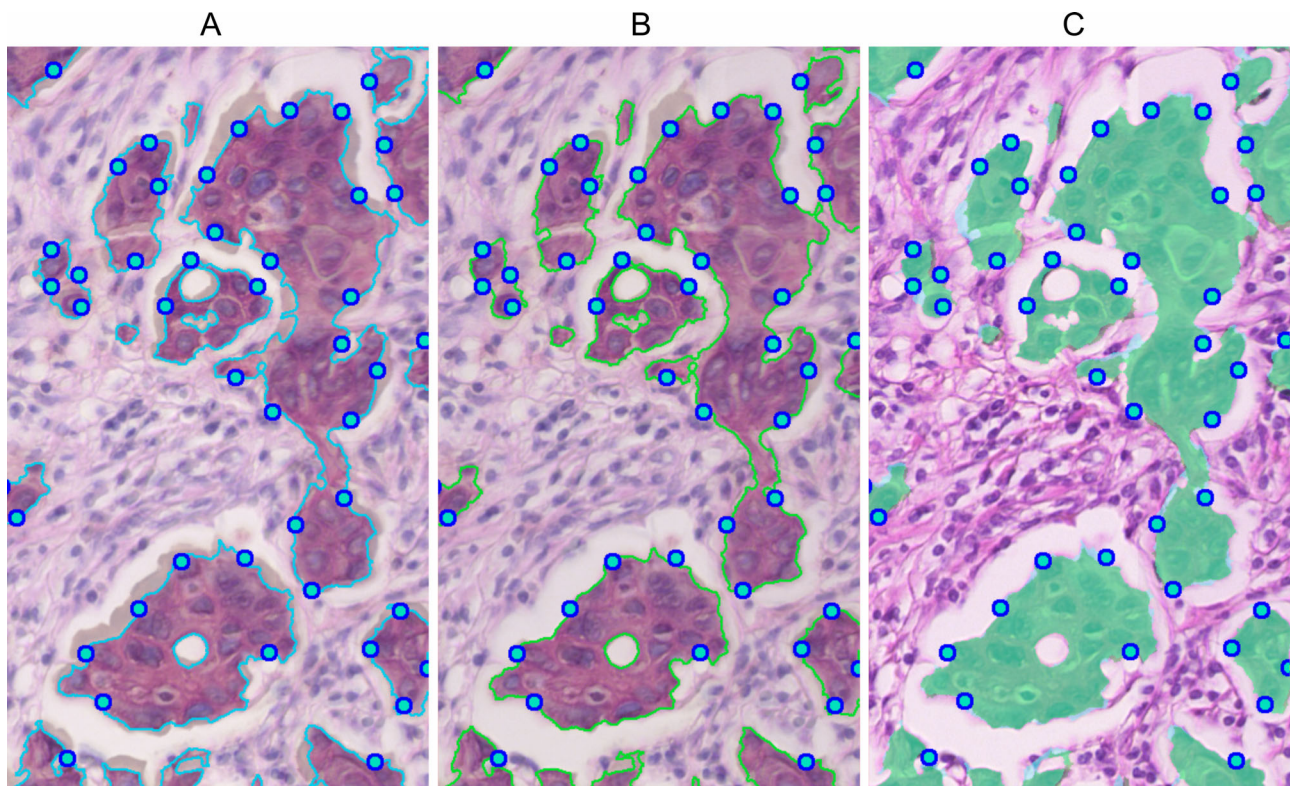


Figure 4. Visualization of registrations and masks generated from part of Figure 3B using the two methods. (A, B) Original H&E-stained slide, overlaid by the registered slide with cytokeratin visualized with DAB. (A) Nuclei-based registration and (B) whole-tissue-based registration. Note that both figures also show outlines of the final mask overlaid on the aligned figures after subtracting white background. (C) Differences between the two masks using overlaid areas. Cyan designates nuclei-based registration, green designates whole-tissue-based registration.

pathologist's expert knowledge, which is consistent with surveys of thresholding methods [18,20,21].

Perhaps surprisingly, the tissue areas in WSIs after immunohistochemical staining are slightly larger than the corresponding regions in the scans of the original H&E section (Figure 4). We presume that this is caused by the antigen retrieval step, involving exposure to an aqueous solution at high temperature and high pH, which produces partial hydrolysis resulting in swelling of the tissue. As part of our nuclei-based registration method, we therefore subtracted areas of the mask where they overlay white background in the H&E-stained image (i.e. areas of the slide without tissue, thereby demarcating the boundary of the tissue). The resulting binary mask is filtered for objects smaller than the fixed threshold (smaller than $10.32 \mu\text{m}/\text{area}$ for our images), which are considered noise.

Two representative areas of two different cores were selected and the border of the epithelium was marked by an experienced pathologist in the H&E image as a series of precisely placed points – these became the reference points. The ℓ_2 distance between the nearest

border point of the automatically computed mask from each of the reference points was taken as a metric of alignment. Pairing of the nearest point and the reference point was also inspected visually.

Automation of stroma/epithelium detection in H&E-stained tissues

To show the feasibility and utility of the automated annotation generation procedure, we developed a simple AI pipeline to detect cytokeratin-rich regions in H&E slides. We used 120 TMA cores to train a segmentation network: 102 colon cores and 18 breast cores. Dimensions of the tissue core images ranged from 2496 to 7104 px, with a median of 4096 px measured as the smallest of the two sides. We used a window slide technique with a step size of 128 px to divide the images into 512×512 px patches.

To avoid extraction of patches from background areas of the image, we filtered out the background. First, we converted the image from RGB (red, green, blue) to HSV (hue, saturation, value/brightness)

Table 3. Evaluation of registration and mask generation for NBR and WTBR.

TC	Method	Mean error (px)	Median error (px)	MSE
2	NBR	2.51	2	11.19
	WTBR	2.23	2	8.83
5	NBR	2.21	1.41	53.31
	WTBR	2.21	1.41	54.12
7	NBR	3.36	1.41	46.88
	WTBR	3.21	1.41	48.14
13	NBR	3.09	2	25.47
	WTBR	2.75	2	21.27

Results are provided individually for tissue cores 2, 5, 7, and 13 (denoted as TC in the table header).

MSE, mean squared error; NBR, nuclei-based registration; WTBR, whole-tissue-based registration.

representation and applied an Otsu's thresholding on the saturation channel of the image. We then applied closing followed by an opening morphological operation using a disk of size 10 [22–24].

Before the patches are passed to the network input, image augmentations are applied randomly on each patch. The following augmentations were used: horizontal and vertical flips with 50% probability each; random 90° rotation; perturbations to brightness in the range (−64; 64); perturbations to contrast in the range (0.7; 1.3); and perturbations to hue and saturation in ranges (−64; 64) and (−10; 10), respectively. As there is no canonical orientation for the tissue samples, we can use the flips to artificially increase the training dataset and thus prevent overfitting. The random perturbations serve as a preventative measure against the network relying on a specific color palette of a slide [25].

We selected U-Net [26] as our architecture of choice. The minimum number of channels was 64 and the maximum was 1024. The Adam optimizer with a combination of dice and binary cross-entropy loss [27] was used to train the network. We initialized the learning rate to 3×10^{-05} . During the training procedure, the learning rate was reduced by a factor of 10 after every 5 epochs with no improvement. The network

was trained with a batch size of 1 for 30 epochs. In addition to image augmentations, ℓ_2 regularization with parameter 1×10^{-05} was used to prevent overfitting. All hyperparameter values were tuned on a validation set.

Results

We measured both the accuracy and the computational performance of cytokeratin mask generation/registration for both the nuclei-based registration and the whole-tissue-based registration on four selected areas of four distinct cores.

Generation of masks

Accuracy

To allow numeric assessment of the masks generated by the automated workflow, we used regions of tissue cores 2, 5, 7, and 13, which were manually annotated by the pathologist with reference points on the border of the epithelial region in the H&E scan. An example is given in supplementary material, Section S3 and Figure S3.1. The numerical accuracy of the generated mask was computed by finding the nearest point on the mask edge to each of these reference points using ℓ_2 distance. The evaluation was manually examined to check if the measurements had caused mismatch of the nearest mask edge point for non-convex mask shapes; no evidence was found for this. The ℓ_2 distance is taken as the measure of error.

As mentioned above, we found that a WSI of cytokeratin DAB after re-staining extends beyond the tissue area of the original H&E WSI. This is directly visible in the results of nuclei-based registration shown in Figure 4A, as this method only uses nuclei as reference points and does not attempt to align the border of the DAB staining to the edges in the H&E image. The whole-tissue-based registration method performs this alignment with inconsistent results

Table 4. Results for tissue cores annotated using a registration method and DAB staining.

TC	Sensitivity	Specificity	Average sensitivity	Average specificity
TC-CRC-1	0.9648	0.9627	0.931 ± 0.056	0.91 ± 0.12
TC-CRC-2	0.9694	0.9756		
TC-CRC-3	0.8894	0.9248	0.794 ± 0.063	0.937 ± 0.013
TC-CRC-4	0.8535	0.9624		
TC-CRC-5	0.9753	0.7011		
TC-MBC-1	0.8386	0.9459		
TC-MBC-2	0.7498	0.9282		

TC, tissue core number.

Table 5. Results for manually annotated colon tissue cores.

TC	Sensitivity	Specificity
KOS01-1	0.8828	0.5138
KOS01-2	0.8567	0.669
KOS01-3	0.9279	0.9139
KOS01-4	0.932	0.9361
KOS01-5	0.9293	0.9437
KOS01-6	0.9286	0.8953
KOS01-7	0.9263	0.6678
KOS01-8	0.9197	0.8093
KOS01-9	0.9363	0.7651
KOS01-10	0.8364	0.8889
KOS01-11	0.933	0.8777
KOS01-12	0.9404	0.8694
KOS01-13	0.9359	0.8482
KOS01-14	0.9292	0.7437
KOS01-15	0.8465	0.8338
KOS01-16	0.8346	0.8375
KOS01-17	0.948	0.6693
KOS01-18	0.9598	0.8349
KOS01-19	0.8818	0.8827
KOS01-20	0.921	0.6967
KOS01-21	0.9449	0.8235
KOS01-22	0.7198	0.9574
KOS01-23	0.9096	0.8924
KOS01-24	0.9222	0.8143
KOS01-25	0.9591	0.7536
KOS01-26	0.973	0.6975
KOS01-27	0.9301	0.6957
KOS01-28	0.9718	0.1184
KOS01-29	0.9621	0.9374
KOS01-30	0.8689	0.9747
KOS01-31	0.9606	0.8634
KOS01-32	0.9429	0.746
KOS01-33	0.8764	0.9483
KOS01-34	0.8994	0.9206
Average	0.913 ± 0.050	0.80 ± 0.16

TC, tissue core number.

(Figure 4B) – sometimes this process succeeds for part of the given region, while it fails for other parts of the same region. However, these issues are handled correctly for both nuclei-based registration and whole-tissue-based registration by subtraction of the mask where it overlaps with the white region on the H&E-stained WSI, as demonstrated in Figure 4C.

Resulting errors are shown in Table 3 and the accuracy of both methods is equivalent, with mean errors of 0.344–0.516 µm. Given that the typical cell size is 15–20 µm and nuclear size is 5.16–6.88 µm, these errors are a fraction of the size of a nucleus. As our goal is to identify groups of epithelial cells, this level of error does not compromise the method’s accuracy for its intended purpose.

The nuclei-based registration procedure is relatively straightforward compared to more sophisticated nonlinear methods. Hence, it is also consistent and robust. Even if tissue damage occurs during cytokeratin re-staining, nuclei-based registration works well as long as there are sufficient numbers of nuclei distributed throughout the tissue. On the other hand, whole-tissue-based registration truly aligning the tissue borders is more problematic if tissue lifting, folding, or loss occurs (supplementary material, Section S3 and Figure S3.7).

For a more detailed comparison of the two methods, see Figure 4. The whole-tissue-based registration method is able to align a mask over the tissue with high accuracy in a well-separated object; nuclei-based registration is agnostic to separation of borders and thus cannot take advantage of clearly demarcated borders. However, if borders are fuzzy (as at the top of

Table 6. Results for manually annotated breast tissue cores and resection slides using a model trained only on breast tissue cores (MBC only) and a model trained using a combination of colon and breast tissue cores (CRC + MBC).

Slide	Region	CRC + MBC		MBC only	
		Sensitivity	Specificity	Sensitivity	Specificity
C-MBC-1	1	0.8359	0.9084	0.8884	0.8815
	2	0.7616	0.9707	0.7633	0.9916
	3	0.9167	0.8927	0.8969	0.9187
	Average	0.838 ± 0.063	0.9239 ± 0.0037	0.850 ± 0.061	0.931 ± 0.046
C-MBC-2	1	0.9068	0.9038	0.9017	0.9163
	2	0.943	0.9629	0.9265	0.9665
	3	0.9162	0.9277	0.906	0.9359
	Average	0.922 ± 0.015	0.932 ± 0.024	0.911 ± 0.011	0.939 ± 0.021
R-MBC-3	1	0.9374	0.7896	0.9414	0.7538
	2	0.8425	0.9421	0.8494	0.941
	3	0.8806	0.8734	0.8054	0.929
	Average	0.887 ± 0.039	0.868 ± 0.062	0.865 ± 0.057	0.874 ± 0.086
R-MBC-4	1	0.8155	0.794	0.7738	0.8127
	2	0.809	0.8769	0.8359	0.8686
	3	0.8335	0.7988	0.7858	0.8687
	Average	0.819 ± 0.010	0.823 ± 0.038	0.798 ± 0.027	0.850 ± 0.026

Figure 4A,B), the error of nuclei-based registration stays consistent, as opposed to whole-tissue-based registration, which incorrectly distorts the mask. However, both these problems are corrected by subtraction of the mask on the white background in the H&E-stained image. This simple (but unexpected) correction aids precise delimitation of the original and re-stained section areas, and consequent alignment of registered nuclei between the two. Thus, distortion within the body of the tissues is also accounted for by consecutive alignment of registered nuclei across the entire tissue images, based on minimal distance between each registered nucleus in the two sets of registered points in the two images.

Results of H&E-stained tissue segmentation

To evaluate the network, we chose pixel-level sensitivity and specificity. We combined all the patches to reconstruct the full mask. The overlapping regions of the patches were combined using arithmetic mean so that the resulting inter-patch borders are smoother. The full maps were thresholded at 0.5. For testing, we used seven TMA cores, five colorectal (TC-CRC) and two breast (TC-MBC), which were annotated using DAB-stained tissues. The results are found in Table 4. Furthermore, we also tested the network on 34 manually annotated colon tissue cores (KOS01), two manually annotated breast biopsies (C-MBC), and two manually annotated full-face breast resection slides (R-MBC).

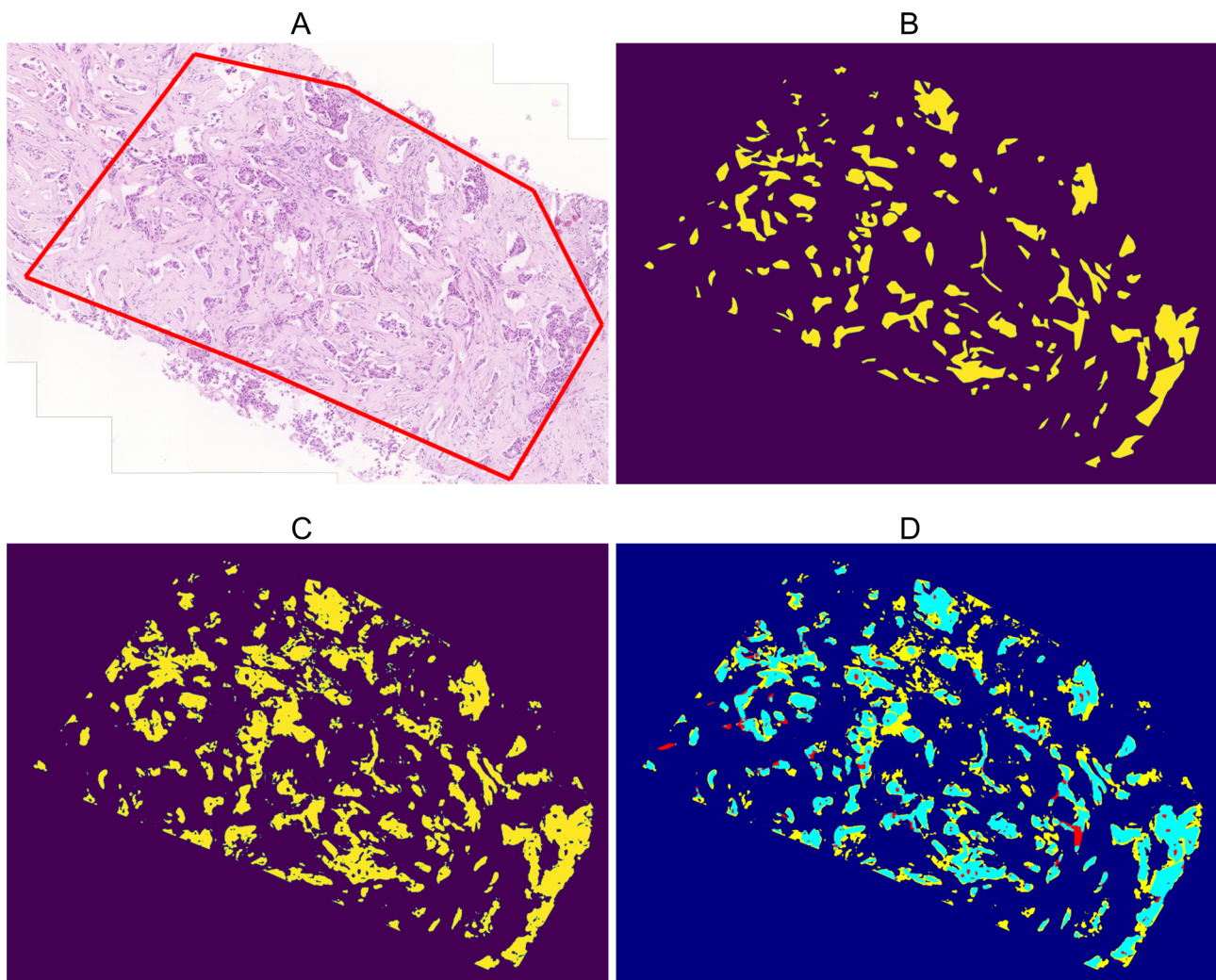


Figure 5. Low-power field from a breast carcinoma biopsy (region 2 from slide M025): (A) the region analyzed; (B) manual annotations made by pathologist; (C) predictions made by neural network; (D) heatmap denoting false positives in yellow, false negatives in red, true positives and true negatives in light and dark blue, respectively.

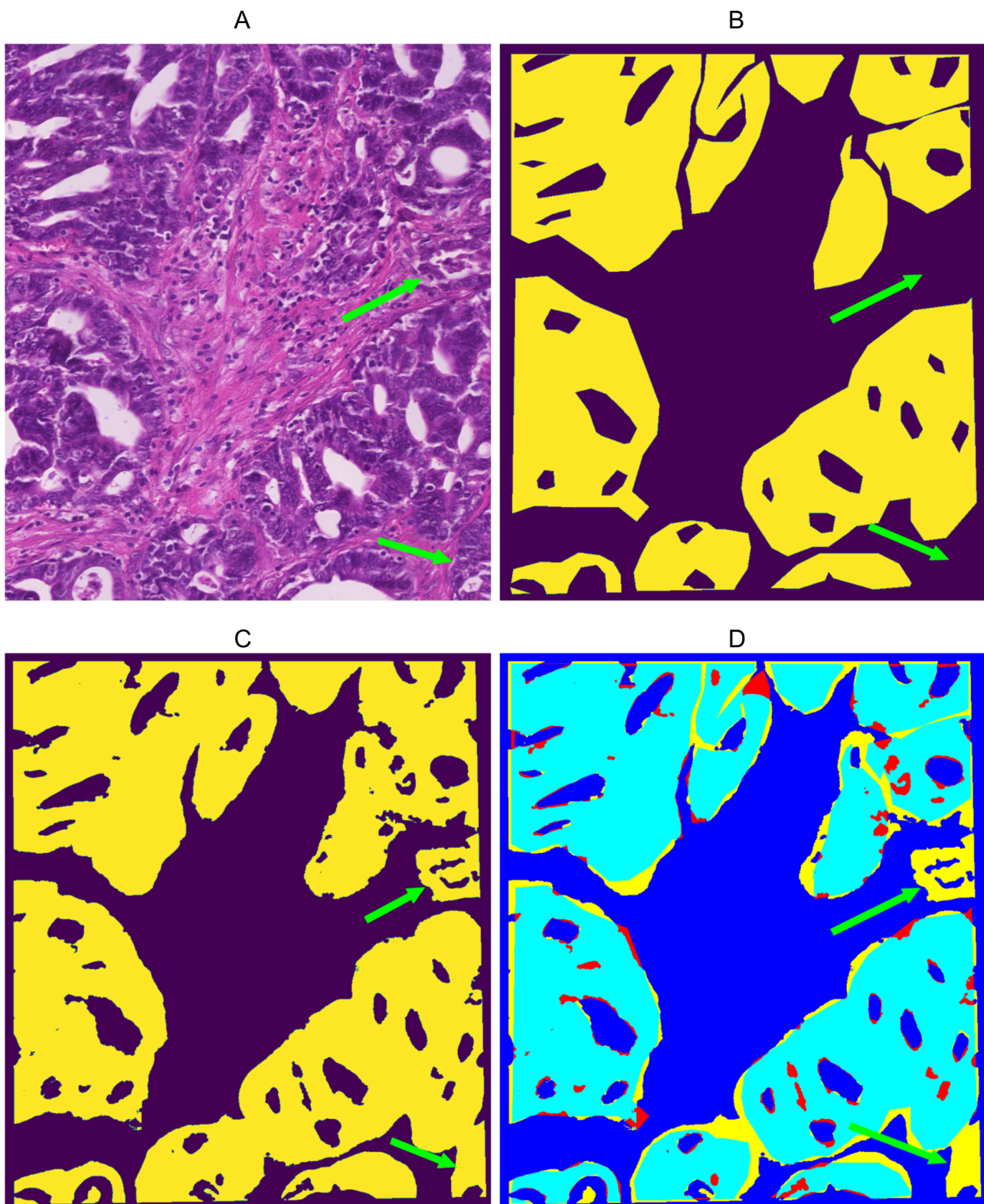


Figure 6. (A) High-power view of H&E-stained KOS01-21 tissue core, cribriform growth pattern; (B) manual annotation; (C) predictions made by neural network; (D) heatmap denoting false positives in yellow, false negatives in red, true positives and true negatives in light and dark blue, respectively. The carcinoma areas omitted by manual annotation are labeled with green arrows.

The results are summarized in Table 5 for colon samples and Table 6 for breast samples. In view of the complexity of the annotation process, only one subregion within each test core (colon) or three subregions in each full-face WSI (breast) were selected, representing the predominant tissue pattern and avoiding section artifacts. For breast tissues, we retrained a U-Net model using only the 18 training breast tissue cores (MBC only) and compared the results with the model trained on a combination of colon and breast tissue cores (CRC + MBC).

In Figure 5, we present the results of the U-Net model on previously unseen core biopsies of breast carcinoma against manual annotation. The results of the model on previously unseen colon cancer cores compared to the subsequently prepared DAB-stained cytokeratin mask are shown in supplementary material, Section S5 and Figure S5.1. These low-magnification images show overall concordance. Figure 6 compares the results of the U-Net model with manual annotations in a selected high-power field of previously unseen colon cancer core. Three more examples can be found in supplementary material, Section S5 and Figures S5.2–5.4. These four cases were selected to represent different histological patterns without knowledge of the U-Net results. Unintentionally, they illustrate how imprecise manual annotation can be. In Figure 6 and supplementary material, Section S5 and Figure S5.2, small areas of carcinoma were omitted by pathological annotation (arrows), but these were accurately recognized by U-Net. The delineation of epithelial islands using polygons is less precise compared to the results of U-Net. In supplementary material, Section S5 and Figures S5.3 and S5.4, two less frequent growth patterns of colon cancer are presented. The dissociating pattern (supplementary material, Figure S5.3) is difficult to annotate manually with precision, and the U-Net result is closer to ground truth. The medullary pattern (supplementary material, Figure S5.4) has less distinct stroma than other types and U-Net reveals some small stromal areas not delineated by manual annotation. Importantly, this pattern was not represented in the learning dataset. As further illustrations of the trained U-net performance to automatically generate virtual epithelial masks from H&E sections without cytokeratin re-staining or manual annotation, we provide four randomly selected WSIs of colon and breast cancer (Figure 1). The figure provides links to allow interactive control of probability and opacity of the automatically generated epithelial masks aligned over the H&E WSIs.

Discussion

Detecting carcinoma cells is a key focus for AI methods in digital pathology, either as an end in itself for cancer diagnosis or as the starting point for further analyses of tumor phenotypes. As such, the ability of AI methods to accurately differentiate stroma from epithelium is an absolute requirement for development and evaluation of AI methods for different cancer types, and for identifying abnormalities in pre-cancers, dysplasias, etc. This paper presents a novel method that overcomes the shortage of manually annotated WSIs to differentiate stroma from epithelium. It must be noted that, in our study, intended as a ‘proof of concept’, we did not attempt to distinguish between neoplastic and non-neoplastic epithelium (both positive for cytokeratins). Some TMA cores in the learning dataset contained admixtures of non-neoplastic epithelium and accurately distinguishing between neoplastic and non-neoplastic cells would require a different construction of learning datasets, containing more non-neoplastic tissue and strictly defined carcinoma areas. This might be a matter of further studies using our approach. In addition, implementation of other AI-based methods employing parameters such as nuclear size, topology, and texture [1–5,28,29] will also be useful in this respect and is to some extent simplified by analyzing only the epithelial components.

By integrating an automated re-staining, scanning, and image processing pipeline (registration, color channel separation, and thresholding), the procedure has potential to generate large amounts of data suitable for machine learning. Another advantage is that the method can use existing H&E-stained material. Hence, well-characterized pathology collections can be used easily for training, without the need to retrieve and recut tissue blocks to provide sections that are consecutive to each other for H&E and corresponding IHC staining. The original H&E-stained slides as well as re-stained sections can be scanned to produce WSIs using any modern scanner. No additional visual cues are needed or assumed by the methods.

In this paper, we have also demonstrated the viability of ingesting the resulting pixel-level annotated images into the machine learning pipeline: we trained a U-Net-based image segmentation method to detect epithelium directly in previously unseen H&E slides. The method of re-staining and alignment can be applied to obtain tumor masks using either full-face tissue sections from clinical samples or, more easily for training purposes, from large numbers of samples using TMAs created from those tissues, as we

demonstrated in breast and colorectal carcinoma, allowing data to be collected rapidly for training on large and diverse sample sets of samples. Once the AI is trained, data derived from applying the virtual mask directly from any H&E-stained section can be transferred to other available software for cell measurements and high throughput analysis of stromal/epithelium compartments, or morphometric analysis can be performed, such as nuclear morphometry, mitotic counts, or epithelial topology. This will undoubtedly aid in the recent development of AI methods that can identify characteristic tumor cell morphologies associated with specific genetic changes [28,29].

In addition, after AI-assisted identification of tumor/stroma compartments in the original diagnostic H&E slide, the re-staining protocol with re-alignment can be applied for automated enumeration of any marker or cell parameter of potential importance clinically or in research (e.g. tumor Ki67 index; PD-L1+ cells in tumor and stroma; tumor-infiltrating CD8 T-cells; proliferation and activation states of endothelium, fibroblasts, or macrophages; etc). Although we used immunoperoxidase with DAB for bright-field microscopy, the re-staining protocol can be adapted to allow fluorescent-based detection or the use of other chromogens, opening up the possibility for analysis of multiple antigens on the same section, coupled with repeated re-staining operations, as required [30–33]. Thus, by correcting for the small scanner shifts and distortions of tissue sections caused by de-staining/re-staining procedures, the nuclei-based co-registration of sequential scans can also be applied to the increasingly evolving multiplexed techniques to allow comprehensive automated analysis of multiple antigens (and DNA/RNA [34]) in a single tissue section from a diagnostic sample. These approaches will therefore allow simplified analysis of both the tumor and the tumor microenvironment in clinical samples. In addition to the simple but important task of identifying tumor from stroma, the re-staining method with alternative markers that identify stromal rather than epithelial antigens can theoretically be used to train similar AI pipelines for automated identification of specific tumor microenvironment components in clinical H&E sections in the future.

Acknowledgements

This work was supported by the Czech Ministry of Health (MMCI 00209805) and the Czech Ministry of Education, Youth and Sports (LM2018125). Computational resources were supplied by the project ‘e-Infrastruktura CZ’

(e-INFRA LM2018140) provided within the Projects of Large Research, Development and Innovations Infrastructures program. The authors thank Philip J Coates for valuable advice on the re-staining method and for revision and editing of the manuscript.

Author contributions statement

RN selected the material, constructed the TMAs, elaborated the re-staining method, performed the scans, and annotated the WSI. TB and PH designed and contributed to implementation of the nuclei-based registration method and mask generation methods. MT and AK contributed to implementation of registration and mask generation methods and performed comparison with the whole-tissue-based registration method. MG designed and implemented the AI pipeline to detect epithelium directly in H&E sections. All authors contributed to the writing and editing of the paper and approved the final version.

References

1. Evans AJ, Bauer TW, Bui MM, *et al.* US Food and Drug Administration approval of whole slide imaging for primary diagnosis: a key milestone is reached and new questions are raised. *Arch Pathol Lab Med* 2018; **142**: 1383–1387.
2. Stathonikos N, Nguyen TQ, Spoto CP, *et al.* Being fully digital: perspective of a Dutch academic pathology laboratory. *Histopathology* 2019; **75**: 621–635.
3. Litjens G, Sánchez CI, Timofeeva N, *et al.* Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. *Sci Rep* 2016; **6**: 26286.
4. Campanella G, Hanna MG, Geneslaw L, *et al.* Clinical-grade computational pathology using weakly supervised deep learning on whole slide images. *Nat Med* 2019; **25**: 1301–1309.
5. Pell R, Oien K, Robinson M, *et al.* The use of digital pathology and image analysis in clinical trials. *J Pathol Clin Res* 2019; **5**: 81–90.
6. Acs B, Pelekanou V, Bai Y, *et al.* Ki67 reproducibility using digital image analysis: an inter-platform and inter-operator study. *Lab Invest* 2019; **99**: 107–117.
7. Bengtsson E, Ranefall P. Image analysis in digital pathology: combining automated assessment of Ki67 staining quality with calculation of Ki67 cell proliferation index. *Cytometry A* 2019; **95**: 714–716.
8. Valkonen M, Isola J, Ylinen O, *et al.* Cytokeratin-supervised deep learning for automatic recognition of epithelial cells in breast cancers stained for ER, PR, and Ki-67. *IEEE Trans Med Imaging* 2020; **39**: 534–542.
9. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. *Anal Quant Cytol Histol* 2001; **23**: 291–299.

10. Alsubaie N, Trahearn N, Raza SEA, *et al.* Stain deconvolution using statistical analysis of multi-resolution stain colour representation. *PLoS One* 2017; **12**: e0169875.
11. Xu J, Xiang L, Wang G, *et al.* Sparse non-negative matrix factorization (SNMF) based color unmixing for breast histopathological image analysis. *Comput Med Imaging Graph* 2015; **46**: 20–29.
12. Borovec J, Munoz-Barrutia A & Kybic J. Benchmarking of image registration methods for differently stained histological slides. In *2018 25th IEEE International Conference on Image Processing (ICIP)*. IEEE: Athens, 2018; 3368–3372.
13. Borovec J. BIRL: benchmark on image registration methods with landmark validation – search results. *arXiv* 2019; 1912.13452v2.
14. Wodzinski M, Skalski A. Automatic nonrigid histological image registration with adaptive multistep algorithm – search results. *arXiv* 2019; 1904.00982.
15. Thirion JP. Image matching as a diffusion process: an analogy with Maxwell's demons. *Med Image Anal* 1998; **2**: 243–260.
16. Ridler TW, Calvard S. Picture thresholding using an iterative selection method. *IEEE Trans Syst Man Cybern* 1978; **8**: 630–632.
17. Velasco FRD. Thresholding using the ISODATA clustering algorithm. *IEEE Trans Syst Man Cybern* 1980; **10**: 771–774.
18. Glasbey CA. An analysis of histogram-based thresholding algorithms. *CVGIP: Graph Models Image Process* 1993; **55**: 532–537.
19. Prewitt JM, Mendelsohn ML. The analysis of cell images. *Ann N Y Acad Sci* 1966; **128**: 1035–1053.
20. Sezgin M, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. *J Electron Imaging* 2004; **13**: 146–165.
21. Xue J-H, Zhang Y-J. Ridler and Calvard's, Kittler and Illingworth's and Otsu's methods for image thresholding. *Pattern Recog Lett* 2012; **33**: 793–797.
22. Kong B, Wang X, Li Z, *et al.* Cancer metastasis detection via spatially structured deep network. In: *Information Processing in Medical Imaging. Lecture Notes in Computer Science, 10265*, Niethammer M, Styner M, Aylward S, *et al.* (Eds). Springer International Publishing: Cham, 2017; 236–248.
23. Li Y, Ping W. Cancer metastasis detection with neural conditional random field. *arXiv* 2018; 1806.07064.
24. Ho DJ, Yarlagadda DVK, D'Alfonso TM, *et al.* Deep multi-magnification networks for multi-class breast cancer image segmentation. *Comput Med Imaging Graph* 2021; **88**: 101866.
25. Liu Y, Gadepalli K, Norouzi M, *et al.* Detecting cancer metastases on gigapixel pathology images. *arXiv* 2017; 1703.02442.
26. Ronneberger O, Fisher P, Brox T. U-Net: convolutional networks for biomedical image segmentation. *arXiv* 2015; 1505.04597.
27. Jadon S. A survey of loss functions for semantic segmentation. In *2020 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE: Via del Mar, 2020.
28. Krause J, Grabsch HI, Kloor M, *et al.* Deep learning detects genetic alterations in cancer histology generated by adversarial networks. *J Pathol* 2021; **254**: 70–79.
29. Yamashita R, Long J, Longacre T, *et al.* Deep learning model for the prediction of microsatellite instability in colorectal cancer: a diagnostic study. *Lancet Oncol* 2021; **22**: 132–141.
30. Du Z, Lin J-R, Rashid R, *et al.* Qualifying antibodies for image-based immune profiling and multiplexed tissue imaging. *Nat Protoc* 2019; **14**: 2900–2930.
31. Ijsselstein ME, Brouwer TP, Abdulrahman Z, *et al.* Cancer immunophenotyping by seven-colour multispectral imaging without tyramide signal amplification. *J Pathol Clin Res* 2019; **5**: 3–11.
32. McGinnis LM, Ibarra-Lopez V, Rost S, *et al.* Clinical and research applications of multiplexed immunohistochemistry and in situ hybridization. *J Pathol* 2021; **254**: 405–417.
33. Remark R, Merghoub T, Grabe N, *et al.* In-depth tissue profiling using multiplexed immunohistochemical consecutive staining on single slide. *Sci Immunol* 2016; **1**: aaf6925.
34. Voith von Voithenberg L, Fomitcheva Khartchenko A, Huber D, *et al.* Spatially multiplexed RNA in situ hybridization to reveal tumor heterogeneity. *Nucleic Acids Res* 2020; **48**: e17.

SUPPLEMENTARY MATERIAL ONLINE

Section S1. Detailed re-staining protocol

Section S2. Detailed nuclei-based registration algorithm

Section S3. Evaluation of mask generation

Section S4. Overview of test samples with generated masks

Section S5. Overview of automated detection of stroma/epithelium samples

Section S6. List of patient samples and scanned H&E sections

Section S7. Software packages and libraries