BULLETIN OF THE POLISH ACADEMY OF SCIENCES
TECHNICAL SCIENCES, Vol. 66, No. 6, 2018
DOI: 10.24425/bpas.2018.125932

Deep learning for damaged tissue detection and segmentation in Ki-67 brain tumor specimens based on the U-net model

Z. SWIDERSKA-CHADAJ[1]*, T. MARKIEWICZ[1,2], J. GALLEGO[3], G.BUENO[3], B. GRALA[2], and M. LORENT[2]

[1]Warsaw University of Technology, Faculty of Electrical Engineering, Warsaw, Poland
[2]Military Institute of Medicine, Department of Pathomorphology, Warsaw, Poland
[3]VISILAB Group, Universidad de Castilla-La Mancha, E.T.S.I.I, Ciudad Real, Spain

Abstract. The pathologists follow a systematic and partially manual process to obtain histological tissue sections from the biological tissue extracted from patients. This process is far from being perfect and can introduce some errors in the quality of the tissue sections (distortions, deformations, folds and tissue breaks). In this paper, we propose a deep learning (DL) method for the detection and segmentation of these damaged regions in whole slide images (WSIs). The proposed technique is based on convolutional neural networks (CNNs) and uses the U-net model to achieve the pixel-wise segmentation of these unwanted regions. The results obtained show that this technique yields satisfactory results and can be applied as a pre-processing step for automatic WSI analysis in order to prevent the use of the damaged areas in the evaluation processes.

Key words: damaged tissue regions detection, artifacts detection, deep learning, Ki-67 staining specimens.

1. Introduction

Deep learning (DL) techniques based on neural networks represent a major area inside of the machine learning field. The dynamic development of the DL in the previous years has been related to the increase of the computer processing performance and GPU parallel processing. Thanks to these technological improvements, building complex deep learning networks and their application to real problems are now possible. So far, DL applications have been used in distinct areas where machine learning can assist humans in their daily tasks. Some examples are: self-driving cars [1], advertising [2] and the topic of this research, medical diagnosis [3, 4].

In the last group, digital pathology applies image processing and machine learning techniques to automate and/or enhance traditional pathology tasks. In their work, pathologists apply different stains to tissue sections to visualize a special type of structures, cells or biological processes inside the tissue slides. Hematoxylin and Eosin (H&E), PAS and Ki-67 are examples of the most common stains used in the laboratories. In our work, we focus on Ki-67 stain, which is used as a cell proliferation biomarker. This method allows for the detection of tumor proliferation areas in cancer specimens.

In order to apply digital techniques to pathological tasks, the stained tissue slides are digitized by means of whole slide scanners, such as Aperio ScanScope XT, 3dhistech, Hamamatsu Photonics or Olympus Scanner. The resultant high-resolution images are so-called whole slide images (WSIs). WSIs are used to analyze the specimens on the screen, and in the case of DL analysis, to train DL models for automating quantitative evaluations and pathological diagnosis. Madabhushi et al. [3] showed machine learning methods applied to different digital pathology challenges. Meanwhile, the authors in [5, 6] presented the possibilities and limitations of DL techniques in digital pathology.

DL methods based on neural networks such as convolutional neural networks (CNNs) are becoming an important trend in assisting pathologists in performing a fast and robust diagnosis [4–7]. They allow a semantic image segmentation to associate each image pixel to a class label with a relatively simple configuration thanks to:

- Unsupervised feature generation.
- The availability of greater datasets to train the model.
- The possibility to import pre-trained models to partially set up the neural networks.

The main inconveniences of CNNs are both the huge amount of computational power and the amount of time necessary to train the networks. Chen et al. presented a survey of semantic segmentation methods by means of Neural Networks in [8].

In cancer diagnosis, the specimen analysis is central to estimate different parameters of the tumor regions and its prognosis.

Dong et al. [9] proposed a fully automatic method for brain tumor segmentation based on deep convolutional networks working with magnetic resonance images. They obtained an efficient segmentation in a database composed of 220 high-grade brain tumor and 54 low-grade tumor cases. Xue et al. [10] developed a method for semantic segmentation of microvascular morphological types on narrow-band images to aid clinical examination of esophageal cancer. They proposed the application of double-label to obtain pixel-wise predictions. The proposed solution achieved 90% accuracy.

Isaksson et al. [7] presented an algorithm based on convolutional neural networks devoted to semantic segmentation of microscopic H&E prostate tissue images. The method achieved

*e-mail: swidersz@ee.pw.edu.pl

Manuscript submitted 2018-01-30, revised 2018-04-26, initially accepted for publication 2018-04-30, published in December 2018.

80% accuracy. In [11], the authors proposed a fully automated algorithm to segment nuclei from histopathology image data by using deep neural networks trained from a set of manually annotated images. BenTaieb et al. [12] designed a multi-objective learning method that optimizes a single unified deep fully convolutional neural network with two distinct loss functions. They focused on colon adenocarcinomas and showed how glands classification can facilitate their segmentation by adding class-specific spatial priors. Finally in [13], Gallego et al. proposed a DL approach based on pre-trained CNNs to automate the glomeruli classification and detection from digitized kidney slide segments.

Regarding the CNN models, the network configuration is defined by the number and type of layers used. Its design conditions the success of the resultant classification. One solution, which is garnering recognition and is being used in many practical applications, includes using pre-designed CNN models tested for generic applications. Some examples are: AlexNet from the ImageNet challenge [14], GoogleNet [15] and U-net, developed by Ronneberger et al. [16]. While AlexNet and GoogleNet allow for image regions classification, U-net model architecture allows to enhance the classification in the case of limited data and performs a pixel-wise classification, thus obtaining the corresponding region segmentation.

1.1. Problem statement. Previously to the digitization of the tissue slides into WSIs, pathologists follow the regular systematic process used in any pathology laboratory to obtain the tissue samples [17]. The process is characterized by:

- Formalin-fixed paraffin-embedded (FFPE) tissue samples. They are obtained by following these steps: tissue fixation, specimen transfer to cassettes, dehydration, clearing and embedding.
- Section thickness of 4 μm by using a microtome tool.

This process is far from being automatic and still presents some significant challenges regarding the quality of the slides obtained [18]. Those are:

- Errors in the manual creation of the tissue sections, which include architecture damages like distortions, deformations or even folds and tissue breaks.
- Color distortions due to the staining process.
- Large size of WSIs (GigaByte images), which makes manipulation and processing difficult.

Both errors in the manual creation and color differences are present in the physical slide and are transferred to the digitized WSIs. In these areas, the specimens present erroneous architecture which should be excluded during specimen evaluation or structure detections. Therefore, in order to enhance the analysis of these WSIs, a system to detect the affected regions inside the WSIs is needed to avoid these non-informative areas that can lead to errors in the final specimen evaluation.

Nowadays, many tumors located in the Central Nervous System (NC) are sectioned with the use of laser equipment. This method offers a precise tumor excision with (thermal) closing of the cut-off plane. Unfortunately, the borders of the resected tumor tissue are also exposed to high temperatures, which leads to a partially thermal damaging of the histological specimens. In the immunohistochemical staining, the damaged regions can accumulate a positive stain substance, so the results of antibody evaluation can be falsified.

In this paper, we propose a DL method based on CNNs – U-net model [16], to automate the detection and segmentation of the mentioned damaged regions present in the WSIs. The classification of these areas is a difficult and complex task since they do not present coherent borders and can have diverse stain intensities and shapes. Figure 1 presents some examples of damaged specimens.

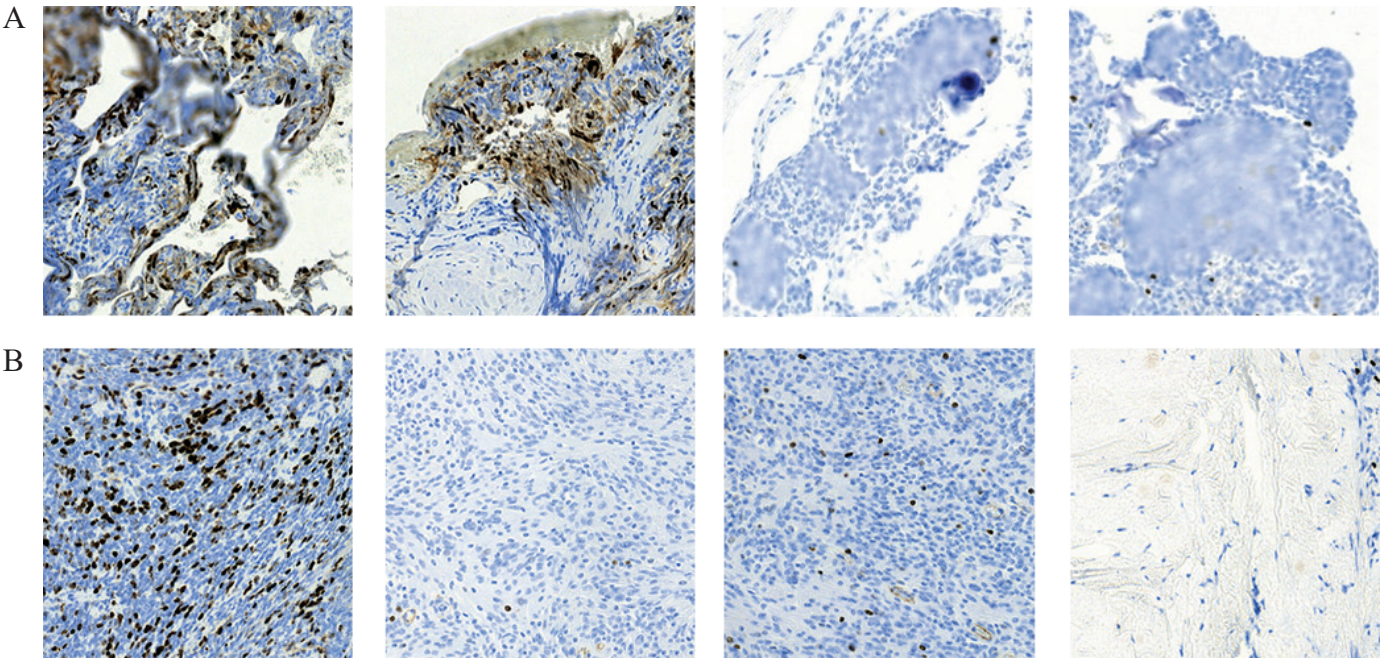

Fig. 1. Examples of damaged (A) and regular (B) parts of specimens

The topic of this study is novel, and to the best of our knowledge, in the literature no similar methods devoted to the detection of architecture damages in WSIs stained with KI-67 have been presented. In order to perform our study, we worked with brain cancer specimens stained with the Ki-67 biomarker.

The proposed solution can be applied as a pre-processing step for WSI analysis in order to prevent the use of the damaged areas in automatic evaluations.

2. Materials

This study involved digital slides from the archives of the Department of Pathology at the Military Institute of Medicine in Warsaw, Poland. The dataset used in the experiments is composed of 34 brain tissue cohorts corresponding to brain tumor areas (meningiomas and oligodendrogliomas) extracted by expert pathologists. The images were collected for a study, and the authors do not have access to any identifying patient information linked to the images. Each of the specimen has been stained with Ki-67 (3,3'-diaminobenzidine tetrahydrochloride, DAB marker, brown color), counterstained with hematoxylin (blue) and was digitized using the Pannoramic 250 Flash II (3DHISTECH, Budapest, Hungary) whole slide scanner equipped with a 20× objective. The WSIs have been saved in MRXS file format with pixel size 0.38895 μm × 0.38895 μm.

The database has been divided into two groups: training and testing. The training dataset includes 10 WSIs, while the testing dataset includes 24 WSIs.

Each of the slides has been manually annotated by expert pathologists to generate the ground truth data for the two-class classifier: i) areas with damaged tissue ii) areas with non-damaged tissue (normal tissue, meninges, empty areas). Training data cohort includes significantly more tiles with damages than non-damaged tiles in order to train network to detect these types of areas.

The annotations include image tiles of 1024 × 1024 pixels and the corresponding binary mask obtained by segmenting the two classes: 1-damaged tissue regions and 0-non-damaged tissue regions. Figure 2 shows an example of manual annotated tiles.

Since the annotated samples may seem barely enough to be fed into a Convolutional Neural Network, a data augmentation technique was agreed upon to be applied. A combination of rotations in 0°, 90°, 180° and 270° and vertical flip were performed. The final number of samples used in the training process, each one containing Damaged and non-Damaged tissue, was established in 4,115 tiles.

3. Method

As mentioned before, we propose a detection and segmentation method for damaged tissue regions in Ki-67 brain tumor sections. The method is based on CNN using U-net model trained by using manual annotated database. The workflow of the proposed system consists of the following three steps:

- CNN model training
- Tissue area segmentation
- Damaged tissue segmentation by U-net.

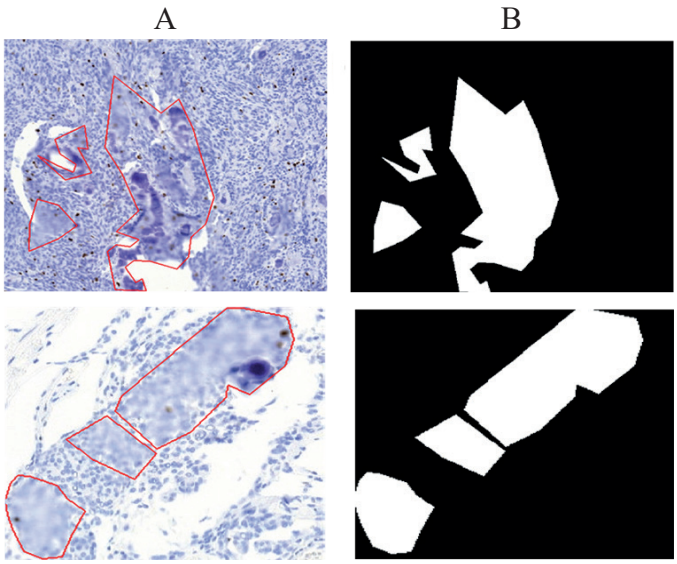

Fig. 2. Examples of annotated tiles. A – original tiles with damaged regions (marked by red). B – corresponding masks

3.1. CNN model training. In this study, we used the U-net deep learning network [16] to configure the CNN classifier. This convolutional network architecture achieves a fast and precise image segmentation suitable for semantic segmentation tasks. The U-net architecture was presented by Ronneberger et al. in [16]. It is based on a contracting path to capture the context and a symmetric expanding path that enables a precise localization (see Fig. 3):

- The contracting path is based on the repeated application of two 3x3 convolutions, each followed by a rectified linear unit (ReLU) and a 2×2 max pooling operation with stride 2 for downsampling. The number of feature channels is doubling in each of downsampling step.
- The expansive path consists of an upsampling of the feature map followed by a 2×2 convolution, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3×3 convolutions, each followed by a ReLU. Besides, in the final layer, a 1×1 convolution is used to map each 64-component feature vector to the desired number of classes. Jointly, the network has 23 convolutional layers.

It should be noticed that the upsampling part has many feature channels, which allow the network to propagate context information to higher resolution layers. As a result, the expansive path is symmetric to the contracting path and gives a final u-shaped architecture (see Fig. 3). In this network, we do not have any fully connected layers, only the valid part of each convolution is used. This approach allows the seamless pixel-wise segmentation of arbitrarily large images by an overlap-tile

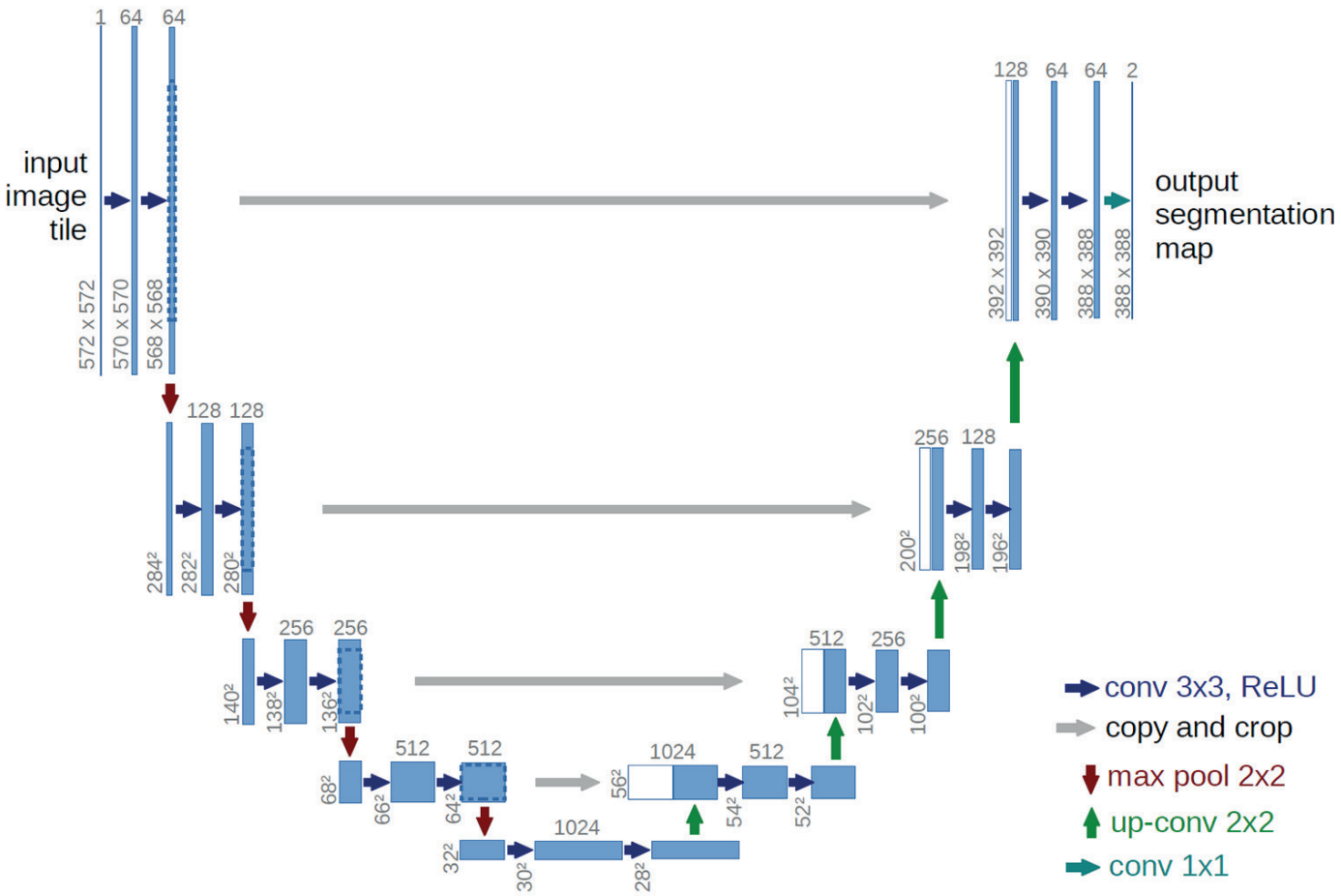

Fig. 3. Graphical representation of U-net model [16]

strategy. Two of the main advantages of the U-net are both the possibility to be trained end-to-end with few images and the fast performance of this network in classification.

In order to prevent neural network from the overfitting [7], we propose to extend the U-net architecture by adding dropout layers. The dropout layers are used to randomly drop units from the neural network during the training, thus avoiding units from co-adapting too much. This approach is useful in the case of limited number of training data. We applied the dropout factor equal to 0.25, which means that 25% of neurons are randomly dropped out.

The model was trained with stochastic Gradient Descent optimizer (SGD), learning rate parameter equalto 0.05, binay-crossentropy loss function and it was monitored by the accuracy metric. The training set parameters were adjusted based on the available literature, our own previous experience [21, 13] and the conducted experiments. It should be noticed that both values and initialization of the network parameters are central to the network training. Incorrect adjusted parameters can lead to an incorrect network training. We train the CNN with the training dataset (described in Section 2) by downsampling 4 times the data resolution to finally work with tiles of 256×256 pixels. Our experiments show that this resizing yields correct classification results while reducing the computational time.

All calculations have been performed on a computer with graphics card Nvidia Quadro 400 and Cuda version 8.0. The network has been implemented in Python 3.5 with Keras [19] and Tensorflow [20] libraries.

3.2. Tissue area segmentation. WSIs include large empty areas without tissue. In order to avoid the classification of these image regions, we apply a tissue segmentation to detect the tissue localization in the WSI. It can significantly limit the amount of data to classify, thus reducing the time of calculations. The tissue mask is obtained by applying the Otsu image thresholding [22] to a 8 times downsampled version of the WSI. Figure 4 shows an example of tissue area segmentation.

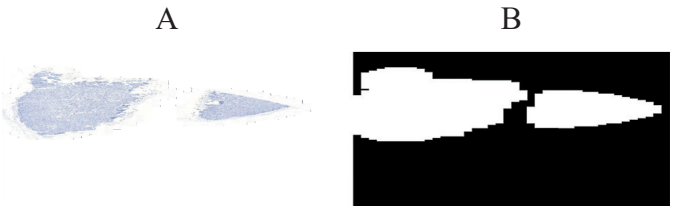

Fig. 4. Examples of a whole slide image (A) and mask for whole slide image (B)

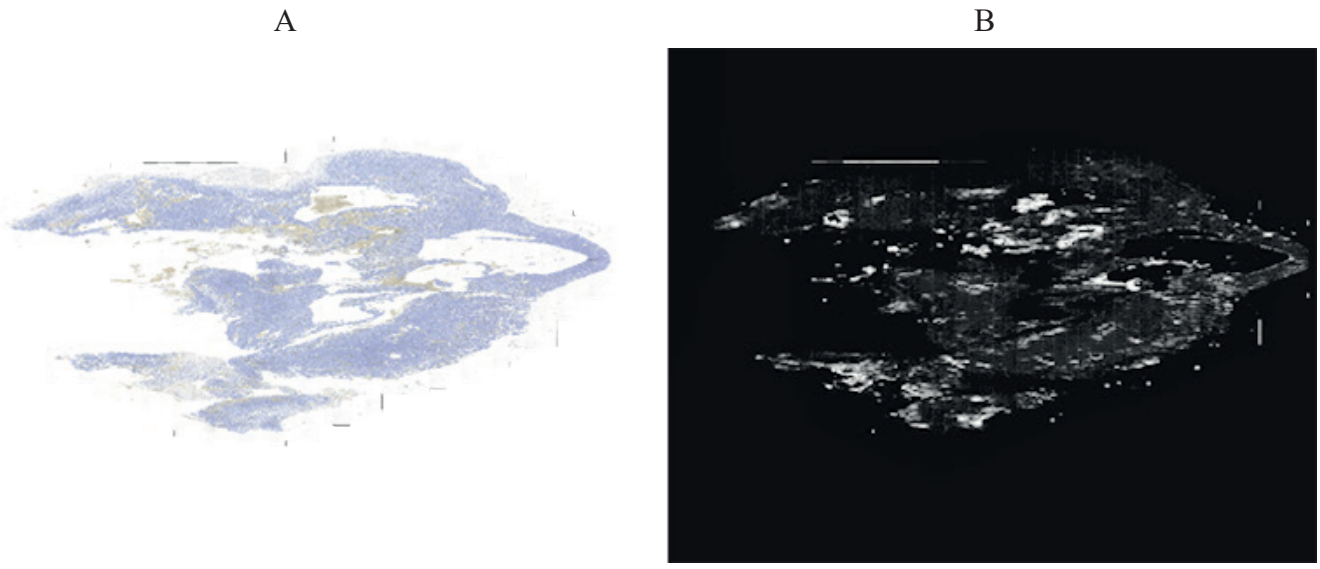

Fig. 5. Examples of original WSI (A) and the resultant Damaged tissue probability map (B)

3.3. WSI-classification. The original WSI under classification is divided into 1024×1024 pixels tiles. Only tissue tiles that present more than 30% of tissue, according to the computed tissue mask, are analyzed. We apply a 4 times downsampling to achieve 256×256 pixel tiles to be classified into both Damaged and non-Damaged classes by the U-net. One of the features of stained specimens pertains to the significant differences in intensity. Hence, a color standardization pre-processing step is applied in order to standardize color distribution in the WSI.

The classification for each tile results in a probability map of 256×256 pixels, where each pixel ranges from 0 to 1, according to the probability belonging to the Damaged tissue class. All the probability maps obtained for each tile create the whole probability map of the 4 times downsampled WSI. Figure 5 shows an example of a resultant Damaged tissue probability map for a WSI.

Finally, in order to obtain the damaged regions segmentation from the probability map, some post-processing operations have been applied: first, we apply a Gaussian blurring filter with parameter 5, to avoid strong changes in the probability map; next, region growing operations are used to obtain the segmentation of Damaged regions (pixels with probability higher than 0.6 are selected as starting points) and finally, we apply an area filtering to remove spurious detections. We use an area threshold of 50 pixels. All the parameters have been selected based on experimental tests. Figure 6 presents an example of the final segmentation results.

3.4. Performance evaluation. The ground truth created by expert pathologists is used to perform the evaluation of our classifier. The manual annotations have been compared with the classification results in a pixel-wise level. The parameters used in this evaluation are:

- Sensitivity (True positive rate) = TP/(TP + FP),
- Specificity (True negative rate) = TN/(TN + FP),

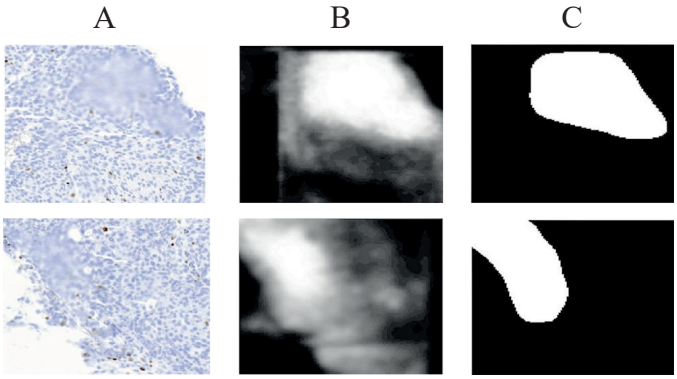

Fig. 6. Examples of damaged regions classification, where In A – original tiles, B – probability maps and C – final classification

- Accuracy: ACC = (TP + TN)/(TP + FP + FN + TN),
- Precision: P = TP/(TP + FP),
- Intersection over union: IoU = OA/UA,

where TP is the number of true positive pixels, TN is its counterpart for true negative pixels, FP are the false positive pixels, FN are the false negative pixels, OA is the overlapped area between annotated masks and the resultant classification and UA stands for the area of union between both regions.

Intersection over union (IoU) is a popular metric applied in the literature for assessing the results performance in semantic segmentation tasks. Since pixel-wise evaluation is very sensitive to contour errors, IoU helps to better evaluate the classification results.

4. Results

In order to evaluate our proposal we have performed quantitative and qualitative evaluations for the proposed classification

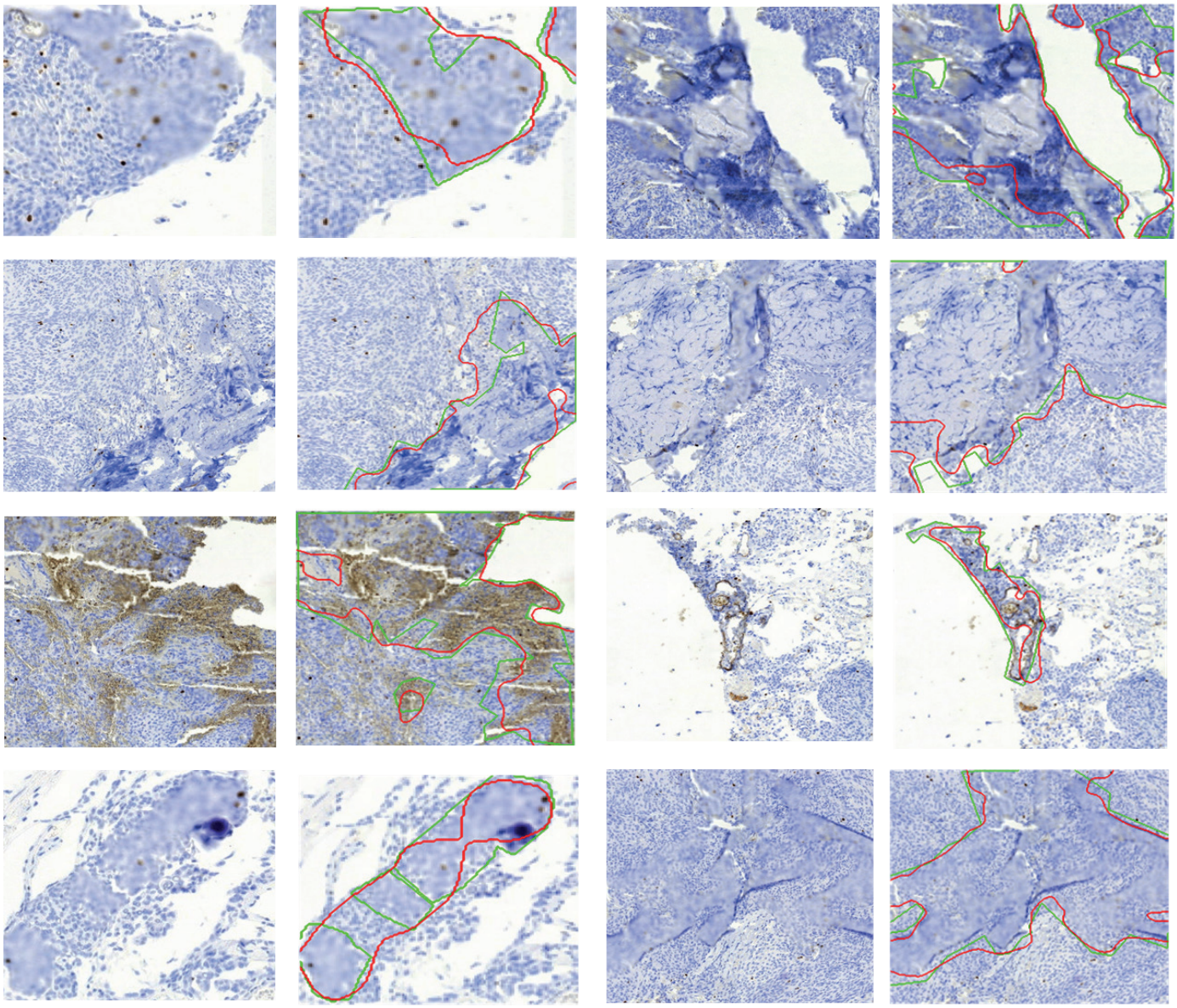

Fig. 7. Example of qualitative results in damaged regions segmentation. Manual annotations are marked in green color. Results obtained using our proposed segmentation are painted in red color

method. We have used 1,893 images from 24 brain tumor slides for algorithm evaluation. As mentioned before, images have been annotated and selected by an expert from WSI specimens where damaged regions are present.

Qualitative results are shown in Fig. 7. As can be observed, the proposed segmentation (red color) is very similar to the ground truth annotations carried out by expert pathologists (green color). Although some regions can present slightly over or under segmentation in the edges, the areas considered as damaged tissue regions by the pathologists and our proposed detection are similar and allow the automation of this segmentation.

Quantitative results have been obtained by using the testing dataset. We have evaluated the performance of the proposed system by means of the parameters described in Section 3.4, and the achieved results are presented in Table 1. As can be observed, the proposed classification and segmentation method achieves 90% accuracy with a precision of 80% and IoU of 69%, which highlight the correct performance of the algorithm in terms of regions detection. Meanwhile, their segmentation partially matches the manual annotated one mainly due to the

Table 1
Statistical evaluation of the achieved results

Sensitivity	Specificity	Precision	Accuracy	IoU
0.83	0.92	0.80	0.90	0.69

differences between manual and detected contours localization. Since damaged regions present an unclear definition of their contours, ensuring their exact positions is a complicated task, which can weaken the performance quality in segmentation purposes.

5. Conclusions

In this work, we have proposed a method to detect and segment tissue damaged areas based on U-net model. Since the standard benchmark data for this challenge are not available, we developed and tested our methodology on brain tumor specimens acquired at our institution.

The damaged tissue areas can have a significant impact on an automatic specimen analysis. They can lead to uncorrected structure detection and classification, and finally, to incorrect specimen evaluation. The detection of tissue regions with the damaged structure is very important in the quantitative analysis of immunohistochemical stains.

Achieved results show that the proposed approach allows for a correct detection of damaged regions with 90% accuracy. Regarding the segmentation, our method achieves an IoU = 69%, which means the existence of differences between expert annotations and our result in the contours of the regions.

In the literature, only few methods related to artifacts and tissue damage detection have been proposed and evaluated. Most of them are devoted to artifacts detection in Hematoxylin and Eosin (H&E) stained specimens [23–25]. The H&E staining is a gold standard for many types of disease. S. Kothari et al. [23] proposed a method to eliminate tissue-fold artifacts in histopathological H&E whole-slide images. The method was used to enhance the image-based cancer detection, achieving a sensitivity of 47% and a specificity of 98%. In our research, we propose a DL-based method to detect several types of artifacts and damages, not only tissue-folds, in Ki67 specimens. Our method achieved sensitivity of 83% and a specificity of 92%.

The results obtained show that the proposed approach is suitable for WSI pre-processing in order to remove damaged areas, which can have a negative impact on automatic specimen analysis. Therefore, our future research will focus on improving the proposed method by using a larger multi-center dataset to work with several stainings and tissue sections.

Acknowledgements. This research was approved by the Research Ethics Board of the Military Institute of Medicine (Number: 30/WIM/2016). This work was in part funded by National Science Center (Poland) by the grant UMO-2016/23/N/ST6/02076.

References

[1] M. Bojarski, P. Yeres, A. Choromanska, K. Choromanski, B. Firner, L. Jackel, and U. Muller, "Explaining How a Deep Neural Network Trained with End-to-End Learning Steers a Car", *arXiv preprint arXiv:1704.07911*, (2017).

[2] X. Liu, W. Xue, L. Xiao, and B. Zhang, "PBODL: Parallel Bayesian Online Deep Learning for Click-Through Rate Prediction in Tencent Advertising System", *arXiv preprint arXiv:1707.00802*, (2017).

[3] A. Madabhushi and G. Lee, "Image analysis and machine learning in digital pathology: Challenges and opportunities", *Medical image analysis*, 33(170–175), (2016).

[4] G. Litjens, C.I. Sanchez, N. Timofeeva, M. Hermsen, I. Nagtegaal, I. Kovacs, C. Hulsbergen-van de Kaa, P. Bult, B. van Ginneken, and J. van der Laak, "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis", *Scientific Reports*, 6, (2016).

[5] G. Bueno, M. Fernandez-Carrobles, O. Deniz, and M. Garcia Rojo, "New trends of emerging technologies in digital pathology", *Pathobiology*, (2016).

[6] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, and C.I. Sanchez, "A survey on deep learning in medical image analysis", *arXiv preprint arXiv:1702.05747*, (2017).

[7] J. Isaksson, I. Arvidsson, K. Aastrom, and A. Heyden, "Semantic segmentation of microscopic images of H&E stained prostatic tissue using CNN", *Neural Networks (IJCNN)*, 1252–1256, (2017).

[8] L. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. Yuille, "Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs", *arXiv preprint arXiv:1606.00915*, (2015).

[9] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks", *arXiv preprint arXiv:1705.03820*, (2017).

[10] D.X. Xue, R. Zhang, Y.Y. Zhao, J.M. Xu, and Y.L.Wang, "Fully convolutional networks with double-label for esophageal cancer image segmentation by self-transfer learning", *Ninth International Conference on Digital Image Processing (ICDIP 2017)*, (2017).

[11] P. Naylor, M. Lae, F. Reyal, and T.Walter, "Nuclei segmentation in histopathology images using deep neural networks", *Biomedical Imaging (ISBI 2017)*, 933–936, (2017).

[12] A. BenTaieb, J. Kawahara, and G. Hamarneh, "Multi-loss convolutional networks for gland analysis in microscopy", *Biomedical Imaging (ISBI)*, 642–645, (2016).

[13] J. Gallego, A. Pedraza, S. Lopez, G. Steiner, L. Gonzalez, A. Laurinavicius, and G. Bueno, "Glomerulus Classification and Detection Based on Convolutional Neural Networks", *Journal of Imaging*, 4(1), 20, (2018).

[14] A. Krizhevsky, I. Sutskever, and G. Hinton, "Imagenet classification with deep convolutional neural networks", *Proceedings of the Advances in Neural Information Processing Systems*, pp. 1097–1105, (2012).

[15] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, and A. Rabinovich, "Going deeper with convolutions", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9, (2015).

[16] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation", *In International Conference on Medical Image Computing and Computer- Assisted Intervention*, 234–241, (2015).

[17] N. Smith and C. Womack, "A matrix approach to guide IHCbased tissue biomarker development in oncology drug discovery", *J. Pathol.*, 232(2), pp. 190–198, (2014).

[18] S. Kothari, J.H. Phan, T.H. Stokes, and M.D. Wang, "Pathology imaging informatics for quantitative analysis of whole-slide images", *JAMIA (J. Am. Med. Inf. Assoc.)*, 20(6), pp. 1099–1108, (2013).

[19] F. Chollet and others, “Keras”, GitHub, https://github.com/keras-team/keras, (2015).

[20] M. Abadi et al., “TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems”, *Software available from tensorflow. org*, (2015).

[21] Z. Swiderska-Chadaj, T. Markiewicz, B. Grala, M. Lorent, and A. Gertych, “A Deep Learning Pipeline to Delineate Proliferative Areas of Intracranial Tumors in Digital Slides”, *Annual Conference on Medical Image Understanding and Analysis (pp. 448–458)*. Springer, Cham, 2017.

[22] N. Otsu, “A threshold selection method from gray-level histograms”, *IEEE Systems, Man, and Cybernetics Society*, 9(1), pp. 62–66, (1979).

[23] S. Kothari, J.H Phan, and M.D. Wang, “Eliminating tissue-fold artifacts in histopathological whole-slide images for improved image-based prediction of cancer grade”, *Journal of pathology informatics*, 4, (2013).

[24] J.P. Johnson, E.A. Krupinski, M. Yan, H. Roehrig, A.R. Graham, and R.S. Weinstein, “Using a visual discrimination model for the detection of compression artifacts in virtual pathology images”, *IEEE transactions on medical imaging*, 30(2), 306–314,(2011).

[25] H. Wu, J.H. Phan, A.K. Bhatia, C.A. Cundiff, B.M. Shehata, and M.D. Wang, “Detection of blur artifacts in histopathological whole-slide images of endomyocardial biopsies”, *37th Annual International Conference of the IEEE*, pp. 727–730, (2015).