Expert Systems with Applications (2020)

Automatic Segmentation of Whole-Slide H&E Stained Breast Histopathology Images using a Deep Convolutional Neural Network Architecture

Blanca Maria Priego-Torres[a,b], Daniel Sanchez-Morillo[a,b,*], Miguel Angel Fernandez-Granero[a,b], and Marcial Garcia-Rojo[a,c]

[a] *Biomedical Research and Innovation Institute of Cadiz (INiBICA), Hospital Universitario Puerta del Mar, Avda Ana de Viya 21, Cadiz, Spain.*
[b] *Biomedical Engineering and Telemedicine Research Group, Department of Automation Engineering, Electronics and Computer Architecture and Networks, School of Engineering, University of Cadiz, Avda. Universidad de Cádiz 10, Puerto Real, Cadiz, Spain*
[c] *Department of Pathology, Puerta del Mar University Hospital, Avda. Ana de Viya 21, Cadiz, Spain*

Abstract

The segmentation of malignant breast tissue from histological images represents a crucial task for the diagnosis of breast cancer (BC). This is a time-consuming process that could be alleviated with the help of computerized segmentation methods, leading to elevated precision and reproducibility results. However, this automated segmentation poses a challenge due to the large size of histological whole-slide images and the significant variability, heterogeneity and complexity of features in them.

In this research, we propose a processing pipeline for the automatic segmentation of stained BC images presenting different types of histopathological patterns. To deal with the gigantic size of whole-slide images, the digital preparations were processed in a tile-wise manner: a large part of the image is split into patches. Then, the segmentation of each tile was accomplished by applying a deep convolutional neural network (DCNN) along with an encoder-decoder with separable atrous convolution architecture, which, once successfully validated, has revealed to be a promising method to segment pathological image patches. Next, in order to combine the local segmentation results (segmented tiles), while avoiding discontinuities and inconsistencies, an improved merging strategy based on an efficient fully connected Conditional Random Field (CRF) was applied.

Experimental results on a collection of patches of breast cancer images demonstrate how the designed processing pipeline performs properly regardless the size, texture or any other colour-shape features typical of the malignant carcinomas considered in this study. The estimated segmentation accuracy and frequency weighted intersection over union ($FWIoU$) were 95.62 %, 92.52 %, respectively. Additionally, in order to facilitate the collaboration between pathologists and researchers to extract the specialist knowledge in form of training datasets that allows the training of new algorithms, a web-based platform which includes a slide-viewer and an annotation tool was developed. The automatic segmentation method proposed in this work was integrated into this platform and currently, it is being used as a decision support tool by pathologists.

Keywords: Breast cancer; Segmentation; Deep Learning; H&E staining; Whole-Slide Imaging

[*] **Corresponding author**

Name: Daniel Sánchez Morillo

Full Postal Address – Escuela Superior de Ingeniería, Avda. Universidad de Cádiz 10, Puerto Real, Cadiz, Spain

e-mail: daniel.morillo@uca.es

Telephone: 0034-956015709

E-mail addresses

blanca.priego@inibica.es (Blanca Maria Priego-Torres)

daniel.morillo@uca.es (Daniel Sanchez-Morillo)

ma.fernandez@uca.es (Miguel Angel Fernandez-Granero)

marcial.garcia.sspa@juntadeandalucia.es (Marcial Garcia-Rojo)

1. Introduction

Breast cancer is the most common malignancy in women worldwide. Estimates of the incidence from cancer in 2018 edition of GLOBOSCAN has shown that breast cancer was the most frequently diagnosed cancer in women around the world with an estimated age-standardized incidence rate of 46.3 per 100.000. Breast cancer also presents the highest mortality rate (13 per 100.000) from cancer among women worldwide (Ferlay et al., 2019). These incidence and mortality rates are rising worldwide, especially across countries with resource-poor settings.

Early detection approaches have been shown critical for reducing the breast cancer burden, including morbidity and mortality (Oeffinger et al., 2015). Detection and diagnosis of breast cancer are performed through clinical exams and imaging modalities. Mammography and ultrasound have become the gold standard technique for breast cancer screening, showing high sensitivity and specificity, cost-efficiency and well tolerance (Crystal, Strano, Shcharynski, & Koretz, 2003; Wellings, Vassiliades, & Abdalla, 2016). Magnetic resonance imaging (MRI) can be used efficiently in the case of high-risk women (Sung et al., 2016).

None of the above mentioned diagnostic screening techniques can deliver 100% accuracy. As a result, once there are suspicions or the tumour is detected by one or more of the above techniques, a biopsy is performed to confirm the diagnosis. A biopsy consists in the extraction by a pathologist of a sample of tissue for testing. Tissue preparation with formalin fixation and paraffin embedding are part of the histopathology techniques, as well as the subsequent staining on glass slides to highlight structures of interest like nuclei and cytoplasm. Haematoxylin and eosin (H&E), complemented with Immunohistochemistry (IHC), are among the most common staining techniques.

Pathology experts will apply visual inspection of histological samples (i.e. size and unproportioned number of cells and tissue abnormalities) to extract quantitative data and to determine the grade, type and distribution of tumour cells, tumour stage, as well as the evaluation of hormone receptors and the HER2 gene and proliferation index, among other biomarkers. These biological characteristics provide great information on the type of breast cancer and consequently the most appropriate treatment for each case (Sharma, Dave, Sanadya, Sharma, & Sharma, 2010).

In recent years, tremendous advancement in image acquisition devices (digital slide scanners) has enabled the scanning of conventional glass slides to produce high-quality digital slides, also called whole slide images (WSI). Scanning magnification factors of 20x and 40x, with an effective pixel size of around 0.50 μm and 0.20 μm respectively, are commonly used (Spanhol, Oliveira, Petitjean, & Heutte, 2016; Veta, Pluim, van Diest, & Viergever, 2014). This development of hardware systems capable of creating a WSI in a time-efficient manner and with reasonable operating costs has caused WSI to be widely used by pathology departments worldwide promoting the field of digital pathology (Madabhushi & Lee, 2016). As a matter of fact, diagnostic, educational, and research purposes of WSI continue to expand nowadays (Pantanowitz, Farahani, & Parwani, 2015).

In digital pathology, histopathological analysis in breast cancer deals with the study of breast tissue WSIs. Notwithstanding, the complex information and wide variations of the content in the histopathological images turn the process into a highly specialized task, which requires experienced and well-trained pathologists to perform it at least as efficiently as in conventional microscopic examination (Gu & Yang, 2019). The whole process is highly time-consuming, expensive, and prone to factors such as a decrease of attention and fatigue, with the significant

consequences for patient outcomes (Chan & Tuszynski, 2016). With the aim to address these limiting factors, computer-aided diagnosis (CAD) systems are being developed to support the pathologist along the process. In this regard, as a large number of WSIs are being generated in pathology departments, efforts are focusing on the development of faster and more precise software support tools through computerized analysis of these histopathological images (Robertson, Azizpour, Smith, & Hartman, 2018).

In addition to classical CAD support systems, recent years have seen a rise in the research into the application of digital pathological image analysis using machine learning algorithms (Bhargava & Madabhushi, 2016; Komura & Ishikawa, 2018; Litjens et al., 2017; Shen, Wu, & Suk, 2017; Xing & Yang, 2016). Among the factors that underlay behind these efforts are an urgent need to develop systems that support the pathologists in their routine tasks, alleviate their workload and address the inter-observer agreement issues among pathologists. Generally adopted workflows in CAD image tools for breast cancer diagnosis have focused on quantitative image analysis (Madabhushi & Lee, 2016), including pre-processing, features extraction and classification stages. In pre-processing, noise reduction (Gil, Wu, & Wang, 2002), image segmentation for detecting regions of interest (ROI) in images (Demir & Yener, 2005), as well as colour normalization (Bejnordi et al., 2016) have been ordinarily used. Likewise, image segmentation to extract morphological (Kowal, Filipczuk, Obuchowicz, Korbicz, & Monczak, 2013; Wang, Hu, Li, Liu, & Zhu, 2016), texture (Peikari, Gangeh, Zubovits, Clarke, & Martel, 2016; Sanchez-Morillo, González, García-Rojo, & Ortega, 2018; Spanhol et al., 2016), fractal (Chan & Tuszynski, 2016; Maipas, Nonni, Politi, Sarlanis, & Kavantzas, 2018), and intensity-based features (Basavanhally & Madabhushi, 2013; Bejnordi et al., 2016) have been common approaches for obtaining discriminative information. Finally, classification has been carried out using well-known learning algorithms such as neural networks (Azar & El-Said, 2013; Dihge, Ohlsson, Edén, Bendahl, & Rydén, 2019), support vector machines (Sumbaly, Vishnusri, & Jeyalatha, 2014), decision trees (Nguyen, Wang, & Nguyen, 2013; Sumbaly et al., 2014), or linear or quadratic discriminant analysis (Raghavendra et al., 2016; Spanhol et al., 2016) among others.

Nevertheless, traditional ruled-based CAD systems and machine learning techniques have been shown to be not sufficient to face the challenge of dealing with the complex problem of supporting breast cancer diagnosis (Razzak, Naz, & Zaib, 2018). Factors of these limitations include the need for crafted extraction of discriminative features by experts, poor convergence and stability issues, insufficient labelled images, the high intra-patient variability, the complexity of the clinical feature representation, and the awkward handling of large WSIs that may consist of as many as tens of billions of pixels (Komura & Ishikawa, 2018; Xu et al., 2017).

Very recently, the availability of computational resources that can satisfy the needs derived from the application of new deep learning techniques, has boosted the research on these methods, displacing in many domains the traditional machine learning methods and converting deep learning algorithms into the dominant technique for supervised learning from data. Deep learning networks are fed with image raw data and can overcome the manual feature extraction by automatically learning multiple levels of abstraction, representation and information. Deep learning approaches have demonstrated promising results in a plethora of domains (Hinton, 2018; Zhao et al., 2019), and current trends in digital pathology point out to the assessment of basic H&E-stained tissue sections and to the development of effective and reliable CAD algorithms for classification and segmentation in WSIs (Robertson et al., 2018). However, the

success of deep learning-based strategies is currently limited by the scarce availability of large manually annotated datasets.

For all the above-mentioned reasons, automatic segmentation of WSIs may contribute to alleviating the rising need to unburden the workload on pathologists, by screening benign areas, so that the expert can focus on the suspicious cases that can be more difficult to diagnose (Gurcan et al., 2009). In automatic segmentation, the algorithm takes part of the WSI as input, and segments the regions of the image that present certain differentiated characteristics. In this regard, automated segmentation of clinically meaningful tissue structures or ROIs in breast histopathological images is a crucial initial step in the development of automated CAD systems. In addition, accurate segmentation of these ROIs may allow performing an efficient automatic analysis of large breast histopathological images by operating in specific relevant areas within the WSI.

However, the automatic ROIs segmentation in breast cancer is challenging due to the high computational costs of pixel-wise segmentation, and to the large regional variations (Su et al., 2015) that these images present. Deep Convolutional Neural Networks (DCNN) have demonstrated a great potential for image classification tasks (Greenspan, van Ginneken, & Summers, 2016; Ting, Tan & Sim, 2019), and more specifically for automatic segmentation of ROIs in WSIs (Bejnordi et al., 2017; de Bel et al., 2018; Su et al., 2015; Vu et al., 2019). As a step forward from segmentation, semantic segmentation aims at assigning a class to every pixel in a given image. Nevertheless, there are few and very recent studies published in the semantic segmentation of histopathological images of breast cancer and most of them describe time-consuming methods that provide a coarse prediction of lesion areas. Among them, Guo et al. (Guo et al., 2019) proposed a fast and refined cancer regions segmentation framework to preselect metastatic regions in H&E-stained WSI of lymph node sections combining a coarse segmentation using a classification model Inception-v3, and a refined segmentation using DCNN. A heat-map of WSI and polygons of lesion regions were extracted to support the pathologists, achieving a free- response receiver operating characteristic (FROC) score of 83.5% on a dataset of metastases in haematoxylin and eosin (H&E) stained whole-slide images of lymph node sections.

The limited availability of studies in this field suggests that may exist some aspects that hinder the development of automatic methods based on deep learning for the processing of WSI. The first concern is related to the lack of publicly available labelled histopathological datasets. When they exist, they are specific to a type of cancer or body location (Bejnordi et al., 2017), or they are restricted to small fixed-sized images (Spanhol et al., 2016), which prevent algorithms from being properly trained. Secondly, methodological problems arise when processing arbitrarily sized ROIs or complete WSIs. Standard DCNNs are designed to receive fixed-size images, which entails the development of new strategies that make these methods flexible to variations in the input size. Finally, the application of deep learning-based methods for the segmentation of tumoral and non-tumoral regions is intrinsically challenging due to the high heterogeneity found in histological images, which will depend on the type of cancer, malignancy, staging, location, etc. The suitability of deep learning algorithms is still poorly explored in this context, so a greater effort is required in the design and validation of these algorithms for the automation of segmentation tasks in breast cancer.

As far as we know, this research is the first to address all the steps involved in the development of a framework for the automatic segmentation of histopathological images in breast cancer, being capable of fulfilling the following aspects:

- Creation of training and test image datasets. We have developed a web-based platform, including a WSI viewer and annotation tool that allows pathology specialists to create ground truth images of ROIs, establishing in this way a methodology to extract expert knowledge in the form of training data sets. Using this platform, sets of RGB (red, green, blue) images and their corresponding ground-truth segmentation maps for the different diagnostic classes (carcinoma types) can be extracted and subsequently used to train a DCNN-based segmentation algorithm.
- Processing of large pathological slides. The proposed framework tackles the segmentation of arbitrary-sized images following a splitting image procedure that allows the individual processing of sub-images, also named tiles or patches. Once the tiles are processed independently, a new merging procedure based on an efficient fully connected Conditional Random Field (CRF) is applied to combine the resulting segmentation maps, ensuring the continuity and consistency on the borders between patches.
- The design of an expert system for the automatic segmentation of stained BC tissue samples. In this research, a DCNN along with an encoder-decoder with separable atrous convolution architecture is posited and successfully validated for the segmentation of H&E-stained BC images when the most common types of malignant breast carcinomas are present on the image dataset. This method reveals to be robust considering the heterogeneity in terms of size, texture and colour that comes with the morphology of tumoral and non-tumoral breast tissues.

The remainder of this article is organized as follows. Firstly, section 2 introduces a description of the training and testing dataset used in this work, along with the annotation and decision support tool developed to create the mentioned dataset. Then, a DCNN-based architecture to segment single H&E stained breast histopathology images is presented, detailing the mechanisms followed to improve the performance of the learning process of these models. This section concludes with the two different strategies proposed to deal with large-scale pathology images, based on a tile-wise processing approach. Section 3 addresses the experimental results of applying the segmentation methods to H&E stained images of breast cancer containing different types of histopathological patterns. Finally, limitations of the study, some concluding remarks, and future research lines are summarized in Section 4.

2. Material and methods

2.1 Dataset description

In this study, training and testing was performed using 12 breast cancer WSI stained with H&E and scanned at $40\times$ magnification level with a 0.2524 μm per pixel resolution in both X and Y dimensions. On each of the digital slides, pathologists and histo-technicians followed an agreed protocol and used a web-based WSI viewer, developed as part of this work, for the accomplishment of the study (see section 2.2). They made annotations through free-hand drawings on approximately 14 areas of 1.000×1.000 pixels, corresponding to areas containing tumour tissue and healthy or non-tumour tissue. After that, these 1.000×1000 pixel areas were split into 4 squared patches, leading to a total of 640 patches of 500×500 pixels.

Table 1 shows the number of patches included in this study considering the type of breast carcinoma present on them.

Figure 1 depicts some samples of the RGB image patches and their corresponding ground-truth segmentation maps for the different diagnostic classes in the image dataset. The ground-truth segmentation maps were defined as the reference maps outlined manually by the pathology experts.

Table 1. Distribution of patches in the image dataset per breast carcinoma type. DCIS: Ductal Carcinoma in Situ; IDC: Invasive Ductal Carcinoma; LCIS: Carcinoma Lobular in Situ; ILC: Invasive Lobular Carcinoma.

Non-tumour tissue	Ductal Carcinoma In Situ (DCIS)	Invasive Ductal Carcinoma (IDC)	Carcinoma Lobular In Situ (LCIS)	Invasive Lobular Carcinoma (ILC)	Other Tumours
304	56	196	28	28	28

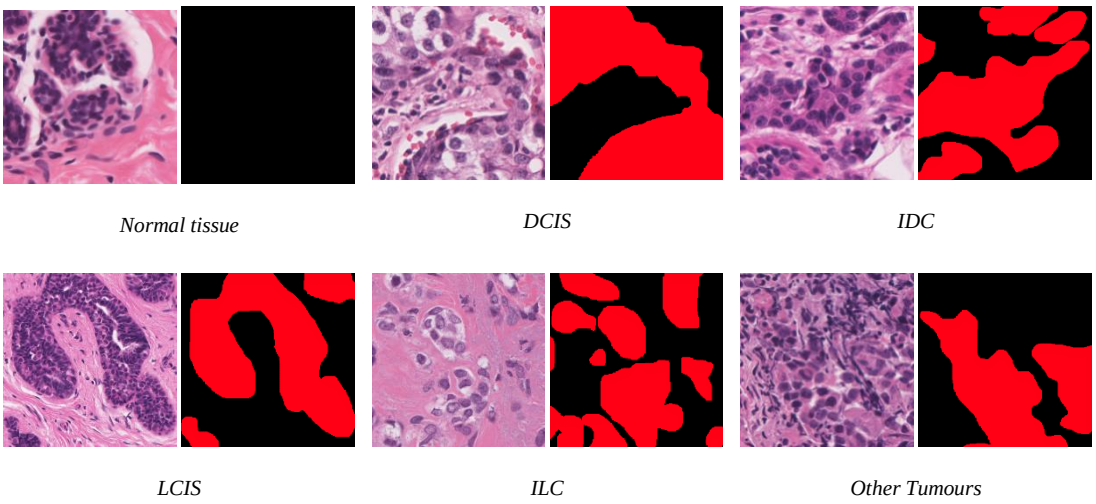


Figure 1. Samples of haematoxylin and eosin (H&E) stained images and their corresponding ground-truth segmentation maps for the different types of carcinoma distributed in the image dataset.

2.2 Annotation and decision support tool

The development of new CAD methods for digital pathology based on deep learning techniques requires to provide the learning algorithms with medical expert knowledge in the form of training and testing datasets. For this and other future works, a web-based vendor-neutral digital slide viewer was developed. It included an easy and intuitive interface for the annotation of WSIs. This viewer allowed to establish a procedure for the generation of datasets, enabling a fast transition from annotation to development of new techniques for image processing and to the validation and visualization of the results.

The web-based platform was built using Python Flask framework. Patient-related information, as well as the paths pointing to where WSIs and annotation files were hosted, were stored using an SQLite database. The OpenSlide image library was used for accessing WSI images. These WSI images were visualized through OpenSeadragon technology. Finally, annotations were vector objects drawn using Paper.js on top of OpenSeadragon and then exported into JSON files.

2.3 Deep-neural-network based segmentation architecture

This study aims at the semantic segmentation of pathology images. As previously detailed, semantic segmentation is based on the assignation of a label from a class/label space to each pixel from the image. Semantic segmentation may be conceived as the next step to image classification and object detection tasks, in terms of complexity, time consumption and detail level. While the image classification problem has been widely addressed in the scientific literature over the last ten years, the use of DCNN-based methods to solve the semantic segmentation problem is very recent, dating back to 2015 (Long, Shelhamer, & Darrell, 2015). Only a few works have been published since then (Badrinarayanan, Kendall, & Cipolla, 2017; L.-C. Chen et al., 2018; L. C. Chen, Zhu, Papandreou, Schroff, & Adam, 2018; Lin et al., 2017), all of them sharing a similar high-level structure.

Recent segmentation works take a convolutional neural network architecture trained for classification, from now on named *backbone architecture*, such as Inception-v3 (Szegedy, Vanhoucke, Ioffe, Shlens, & Wojna, 2016), Resnet (He, Zhang, Ren, & Sun, 2016) or Xception (Chollet, 2017), and remove the last fully connected layers in order to produce downsampled feature maps containing dense representations of the input image. This step, along with some further processing applied to the feature maps, is commonly termed as *encoder*.

Afterwards, the low-resolution image representations are reconverted into prediction or segmentation maps of the same size as the input image, what implies the gradual recovering of spatial information keeping the boundaries between objects or classes as well-defined as possible. This part is named *decoder* and normally entails the main difference among the segmentation methods proposed so far.

From the available state-of-the-art of semantic segmentation methods and motivated by the high scores provided in the segmentation of objects, *DeepLabv3+* method was selected in order to assess its suitability when applied to pathology images (L. C. Chen et al., 2018).

This segmentation method takes modified versions of the Mobilenetv2 (Sandler, Howard, Zhu, Zhmoginov, & Chen, 2018), Xception or ResNet architectures as a backbone network. Then, it substitutes the max pooling and batch normalization of the last blocks for a convolutional variant named atrous or dilated convolution (Lin et al., 2017). This convolution block allows the explicit control of the resolution at which feature responses are computed. The application of atrous convolutions with different rates enables the adaptive modification of the filter's field-of-view without increasing the amount of computation, working as a mechanism to balance the accurate localization of objects (low rate values, implying reduced field-of-view) and context assimilation (larger rate values). Additionally, these atrous convolutions are applied in a separable manner in order to speed up the training and segmentation prediction. For this, atrous convolution is decomposed into a depth-wise convolution (spatial convolution applied to each channel of the input image independently) and pointwise convolution (1×1 convolution applied to the output of the depth-wise convolution).

The Deeplabv3 encoder structure is based on an atrous spatial pyramid pooling (ASPP). It condenses the features extracted by the backbone network into multi-scale contextual information by the application of parallel atrous convolutions at multiple sampling rates (different fields-of-views) and pooling operations (Figure 2). In this way, ASPP helps to improve the segmentation of objects belonging to the same class that may appear at different scales in the input image. The feature maps obtained from the atrous convolutional branches are concatenated and then fused together with 256 filters through with 1×1 convolutions.

In the decoder part, the encoder feature maps are upsampled by a factor of 4 and then concatenated with those features from the backbone network that have the same spatial resolution after applying a 1×1 convolution to them. Finally, the feature maps feed a few 3×3 convolutional layers for features refining and the outputs are upsampled by a factor of 4, resulting in the final segmented image.

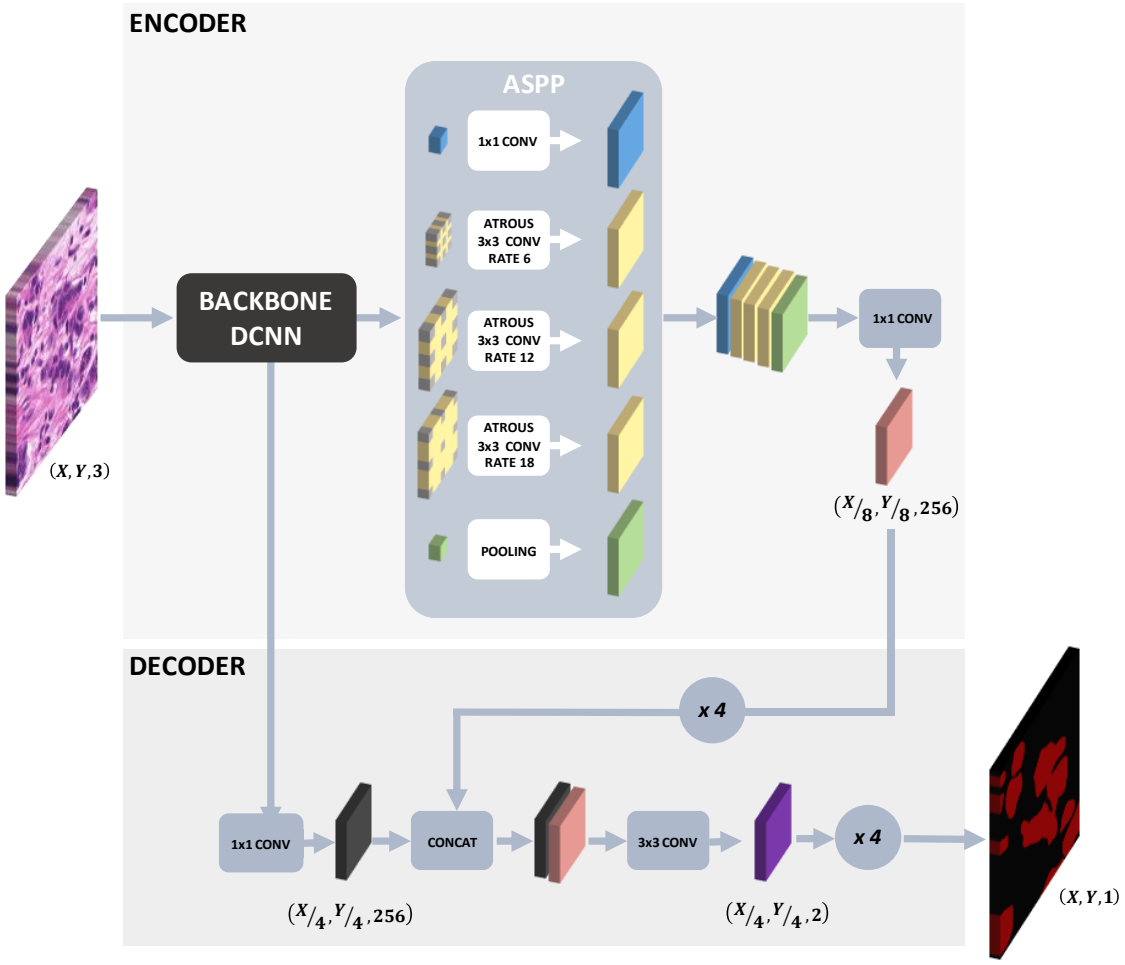


Figure 2: Illustration of the Deep-Neural-Network-based pipeline proposed for the segmentation of haematoxylin and eosin (H&E) stained breast histopathology images.

2.4 Transfer learning and data augmentation

Transfer learning has revealed to be an effective method to overcome a typical problem that appears when artificial intelligence algorithms deal with real-world problems: learning or adjusting a large number of parameters that make up the deep architectures with limited labelled datasets. Collecting massive hand-labelled training sets is often highly time-consuming and expensive, and on scenarios like digital pathology requires expert knowledge. Through inductive transfer learning (Pan & Yang, 2010), a model is pre-trained using large labelled datasets from an unrelated problem and then adapted to the problem at hand just with minimal retraining, thus avoiding much expensive data labelling efforts. In the particular application of Deeplabv3+ to segment H&E stained breast histopathology images, only the parameters of the

backbone network were re-used, while the rest of the parameters in the encoder and decoder parts were not initially transferred.

Data augmentation is a standard technique to tackle data scarcity in deep learning without collecting new data. This strategy aims at expanding the training dataset in order to improve the performance and ability of the model to generalize. Data augmentation consists in applying image transformations to existing labelled images creating new transformed versions of images in the training dataset that belong to the same class as the original image. Data augmentation was accomplished in this work by applying three different operations: random zooming, cropping and flipping.

Zooming operation allows either zooming in or zooming out: the maximum zoom-in percentage applied to images was 120%, followed by a random cropping of the image; and the minimum zoom-out percentage considered was 80%, which required to add padding to the borders of the image using the mean RGB value of the image as constant.

Flipping allows mirroring the image. Both horizontal and vertical flips were applied with a probability of 0.5 respectively.

Finally, in this work, augmentation was applied on-line, that is, augmenting every data sample stochastically each time the training loop was being fed.

2.5 Tile-wise segmentation strategy

WSIs in digital pathology are extremely large in size compared to other types of medical imaging techniques, such as MRI or X-ray images. For this reason, and due to the limitation of the graphics processing unit (GPU) memory and performance, deep neural network-based segmentation methods are not able to tackle the direct processing of complete images.

To handle this problem, a tile-wise segmentation strategy was applied in this work. It enabled the segmentation of regions of any size by splitting the image into as many tiles as needed. Since context information is very relevant for DCNNs to segment tumour tissues, overlapped tiles were extracted in order to alleviate potential segmentation issues in boundaries.

The merging of tile-wise segmentation results was addressed following two different approaches. In the first strategy (Figure 3a), the image was split into tiles of 500×500 pixels, which were independently segmented by the trained Deeplabv3+ algorithm, so that the tile-wise segmentation results were directly assigned to the whole segmentation map.

Since some overlapping was introduced in the splitting process, the rule followed for the label assignation in the overlapping areas was $l_i = \cup_{s=1}^{N} l_{i,s}$ where l_i is the final label assigned to the pixel i laying in the overlapping area, taking $\{0, 1\}$ values, $l_{i,s}$ is the output label of the pixel i in tile s, N is the number of tiles that overlap on pixel i, and $\cup$ refers to OR operation.

Given that no post-processing is generally applied to the Deeplabv3+ segmentation results, the adopted strategy presented two major drawbacks. First, the whole segmentation maps were likely to be noisy. And second, discontinuities might appear in between the boundaries of overlapping regions. In order to overcome these issues, a novel second strategy was proposed (Figure 3b). Again, tiles of 500×500 pixels were considered in the splitting process, with a 50% overlapping rate. Then, the tile-wise logit maps of the last block of Deeplabv3+ architecture were normalized by applying a softmax function. As a result, tile-wise probability maps were obtained, which were later averaged in their overlapping areas. Next, the whole

probability maps were combined with an efficient fully connected Conditional Random Field (CRF) (Krähenbühl et al 2011), in which the pairwise edge potentials are defined by linear combinations of Gaussian kernels, in order to improve the performance of tumour tissue localization by defining the fine edge details and capturing long-range dependencies.

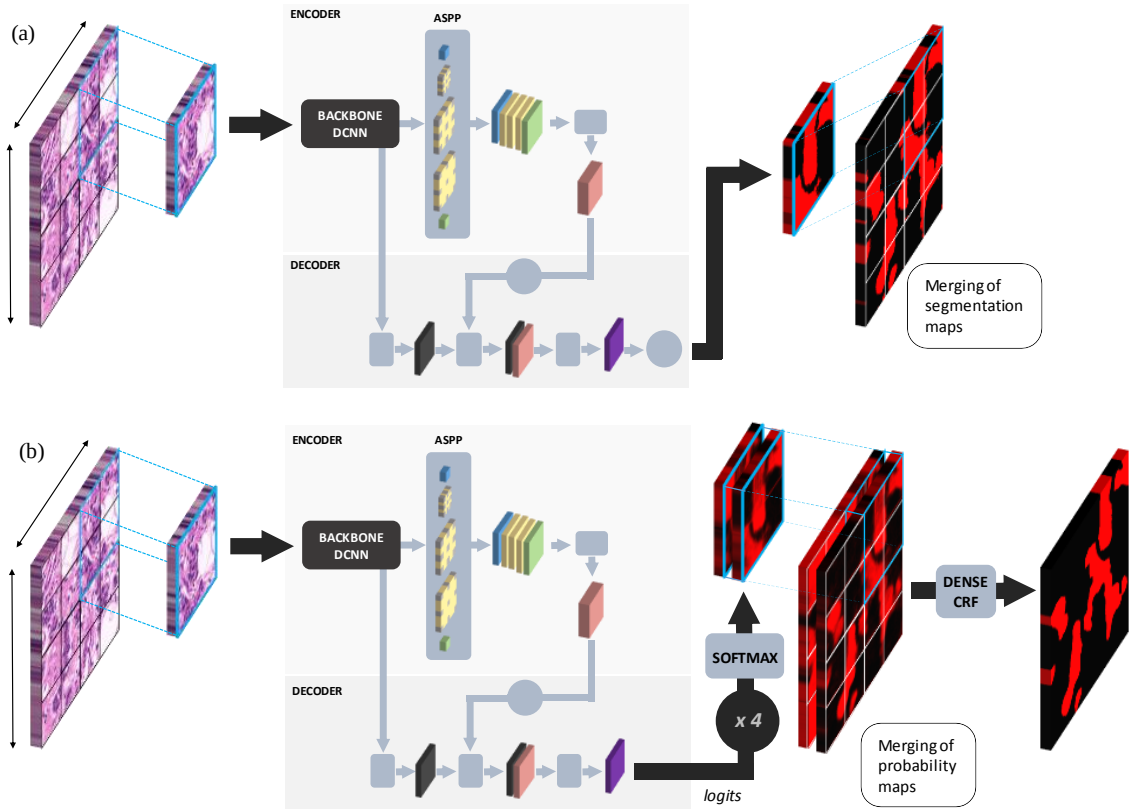


Figure 3. Tile-wise segmentation strategies addressed in this study: a) direct tile-wise merging; b) tile-wise merging based on a Conditional Random Field (CRF).

2.6 Validation and evaluation metrics

To assess the generalization ability of the proposed DCNN-based architecture, the dataset of 640 H&E stained breast histopathology images was divided into five parts according to fivefold cross-validation principle. For each fold, 512 (80%) patches were selected from the 640 images and used to generate a training set. The remaining 128 (20%) images not selected for training were used to build a testing set. The process of assigning images to the training and validation sets was random, although the percentage of samples from each class was preserved.

To evaluate the performance of the segmentation algorithm, two success rates and three similarity metrics commonly applied in semantic segmentation studies were used in this work to compare the consistency between the automated segmentation and the ground truth.

Success rates included sensitivity (Se), and specificity (Sp). To detail these evaluation metrics, some definitions were considered. Let Z be the set of n pixels in an image. Let the ground truth segmentation (performed by pathologists manually) be represented by the set of pixels G. Let S

be the set of pixels labelled as segmented region by the proposed automated segmentation strategy, and being evaluated. Let the following disjoint sets be defined as follows:

- True positive (TP) as set of pixels for which the label assigned by the proposed method matches the ground truth label ($TP = G \cap S$).
- True negative (TN) as set of the non-tumour region labelled pixels between these two sets ($TN = \bar{G} \cap \bar{S}$).
- False positive (FP) set as $FP = \bar{G} \cap S$
- False negative (FN) set as $FN = G \cap \bar{S}$

According to the above definitions, Se and Sp were calculated following the Eqs. (1) and (2):

$$Se = \frac{TP}{TP+FN} \tag{1}$$

$$Sp = \frac{TN}{TN+FP} \tag{2}$$

Sensitivity reports the amount of manually detected pixels that were also correctly segmented by the proposed algorithm whereas specificity estimates the percentage of true negatives correctly labelled by the proposed segmentation method.

Similarity metrics included accuracy (Acc), mean intersection over union ($MIoU$), and frequency weighted intersection over union ($FWIoU$).

Let k denote the number of different segmentation classes ($k = 2$ in this study) and n_{ij} denote the number of pixels of class i predicted to belong to class j. Acc represents the proportion of correctly segmented pixels to total pixels and can be expressed as follows:

$$Acc = \frac{\sum_{i=0}^{k-1} n_{ii}}{\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} n_{ij}} \tag{3}$$

The intersection over union (IoU) metric quantifies the overlap between the ground truth and the predicted output. $MIoU$ refers to the IoU computed on a per-class basis and then averaged. In $FWIoU$, the importance of each class score is weighted according to its frequency of occurrence. $MIoU$ and $FWIoU$ can be formulated according to Eqs. (4) and (5).

$$MIoU = \frac{1}{k}\sum_{i=0}^{k-1} \frac{n_{ii}}{\sum_{j=0}^{k-1} n_{ij} + \sum_{j=0}^{k} n_{ji} - n_{ii}} \tag{4}$$

$$FWIoU = \frac{1}{\sum_{i=0}^{k-1}\sum_{j=0}^{k-1} n_{ij}} \sum_{i=0}^{k-1} \frac{\sum_{j=0}^{k-1} n_{ij} n_{ii}}{\sum_{j=0}^{k-1} n_{ij} + \sum_{j=0}^{k} n_{ji} - n_{ii}} \tag{5}$$

Se, Sp and Acc were calculated globally, considering each pixel independent from the patch it belonged, whereas $MIoU$ and $FWIoU$ were calculated image by image and afterwards averaged.

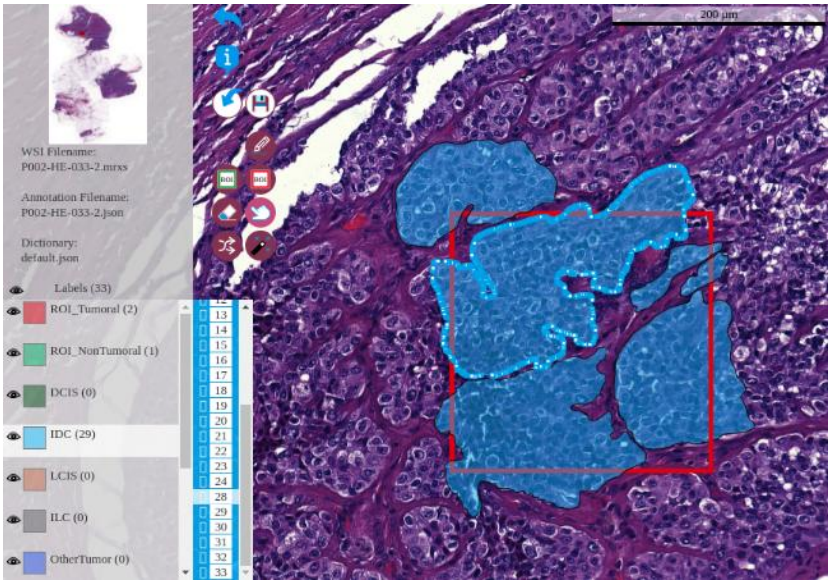
2.7 System Configuration

Training and tests of the automatic segmentation algorithm applied to H&E stained breast histopathology images using deep learning was accelerated by using an NVIDIA DGX station. DGX consists of four NVIDIA® Tesla® V100 Tensor Core GPUs, integrated with a fully-connected four-way NVLink™ architecture. The segmentation method was based on the

publicly available model Deeplabv3+ pre-trained on PASCAL VOC 2012 dataset. Resnet v1, Mobilenet v2 and Xception 65 were used as backbone networks.

For the training of Deeplabv3+, a mini-batch of 8 patches and an initial learning rate of 0.0001 are set. The learning rate is multiplied by 0.1 after 2000 iterations, using weight decay of 0.00004 and a momentum of 0.9 and the number of steps used for training in all the experiments was 150000. Each experimental learning process took 21 hours approximately and the inference time for each patch in the test dataset is about 0.23 seconds.

(a)



(b)

Figure 4. Screenshots of the whole slide image viewer and annotation tools.

3. Results and Discussion

3.1 Annotation and decision support tool

The web-based annotation and decision support tool was designed and implemented gradually, trying to satisfy several usability suggestions made by the pathologists involved in this work. Some relevant features of this tool are gradual zooming, free-hand drawing of annotations on the slide, an intuitive procedure for editing already-made annotations, and easy navigation among them (Figure 4). The tool also allows importing previously made annotations and incorporates a decision support button for automatically segment a region of interest. The video included as supplementary material illustrates the use of the developed tool.

3.2 Segmentation assessment of single tiles

First experiments were devoted to assessing the suitability of the DCNN-based segmentation architecture to segment H&E stained breast cancer images following the criteria from pathologists when annotating histopathological patterns in the form of ground truth or reference maps. To this purpose, Deeplabv3+ algorithm relying on Xception 65 as backbone network was retrained.

Once the segmentation model was fine-tuned, it was validated over the test set, reaching an accuracy, specificity, sensitivity, $MIoU$ and $FWIoU$ of 95.62%, 97.39%, 88.58%, 64.52% and 92.52%, respectively. Figure 5 displays selected image patches revealing different types of histopathological patterns and their corresponding segmentation results after applying the proposed method. It can be observed that the segmentation maps follow closely the segmentation criteria from expert pathologists, although the resulting images tend to present slight noise around the boundaries between tumour and non- tumour areas.

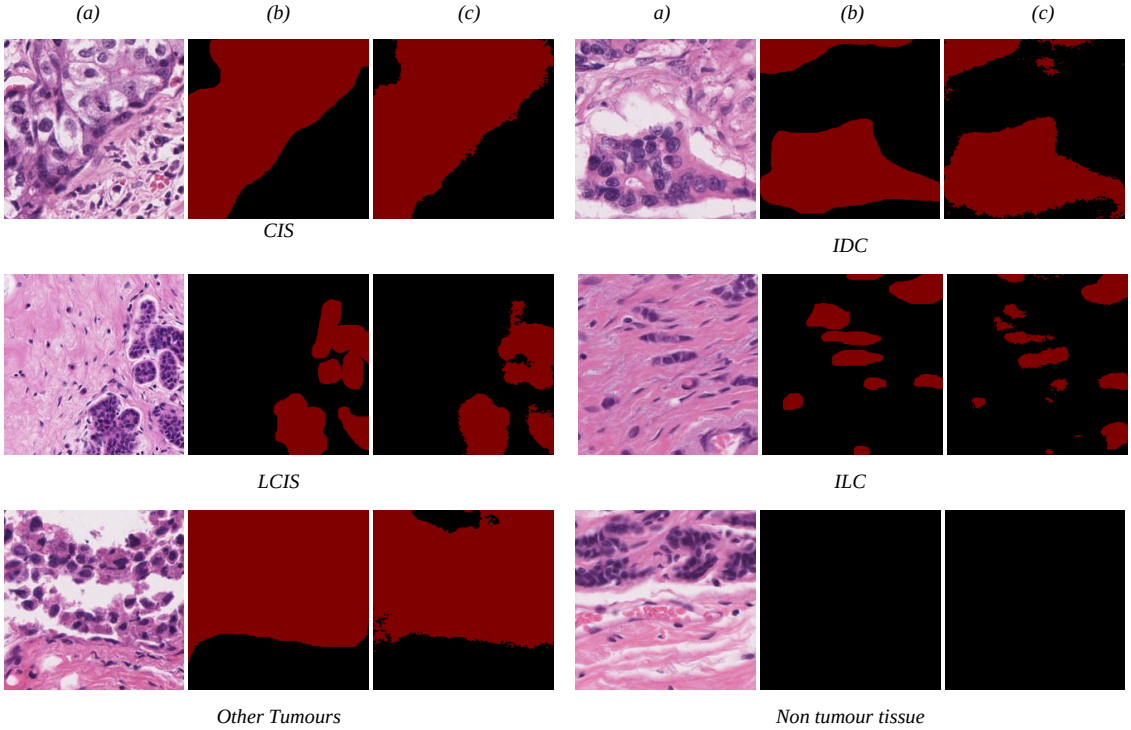


Figure 5. Segmentation results for different types of histopathological pattern (a) RGB images of

the haematoxylin and eosin (H&E) stained images from the test set (b) ground-truth segmentation maps and (c) predicted segmentation.

Table 2 shows a more detailed analysis of pixel-wise classification metrics considering the histopathological patterns included in the test set. For the extraction of these metrics, only patches presenting tumour areas were considered. The results of the evaluation indicate that patches containing in situ (DCIS and LCIS) carcinomas were more reliably segmented than those with invasive histopathological patterns, such as ILC, which ended up with the lowest performance values.

Table 2. Pixel-wise assessment metrics according to the type of breast carcinoma. DCIS: Ductal Carcinoma in Situ; IDC: Invasive Ductal Carcinoma; LCIS: Carcinoma Lobular in Situ; ILC: Invasive Lobular Carcinoma.

	DCIS	*IDC*	*LCIS*	*ILC*	*Other Tumours*
Accuracy (%)	96.91	92.16	97.75	88.72	93.81
Specificity (%)	98.30	92.64	98.56	95.65	97.57
Sensitivity (%)	94.84	91.37	94.40	55.06	89.55

Once Deeplabv3+ model was correctly validated, subsequent experiments analysed the performance of the DCNN-based segmentation method according to the backbone network used to extract the feature maps in Deeplabv3+ algorithm. Xception 65, Mobilenet v2 and Resnet v1 networks were used in this task. Again, the segmentation algorithms were retrained and validated using a 5-fold cross-validation strategy, maintaining the same five training and test folds/sets for all the experiments.

Table 3 illustrates how Xception 65 performed better as backbone network than the others, providing the highest scores for the majority of considered metrics, except for specificity, obtaining a close value to the maximum provided by Resnet v1. Figure 6 plots the learning curves in the training process when the three backbone networks were used in the encoder module of the segmentation algorithm. The loss curves represent the cross-entropy calculated on 8-image batches as a function of training steps. As it can be observed, even though Xception 65 provides the highest loss values at first, the loss falls below the other two curves after approximately 10.000 steps, stabilizing much faster.

Table 3. Segmentation assessment metrics using different DCNN as backbone networks in Deeplabv3+. Acc: accuracy; Sp: specificity; Se: sensitivity; $MIoU$: mean Intersection over Union; $FWIoU$: frequency-weighted Intersection over Union.

	Acc (%)	Sp (%)	Se (%)	$MIoU$(%)	$FWIoU$(%)
Xception 65	95.62	97.39	88.58	64.52	92.52
Mobilenet v2	92.90	96.49	78.45	59.72	88.59
Resnet v1	91.16	98.06	62.55	57.25	86.19

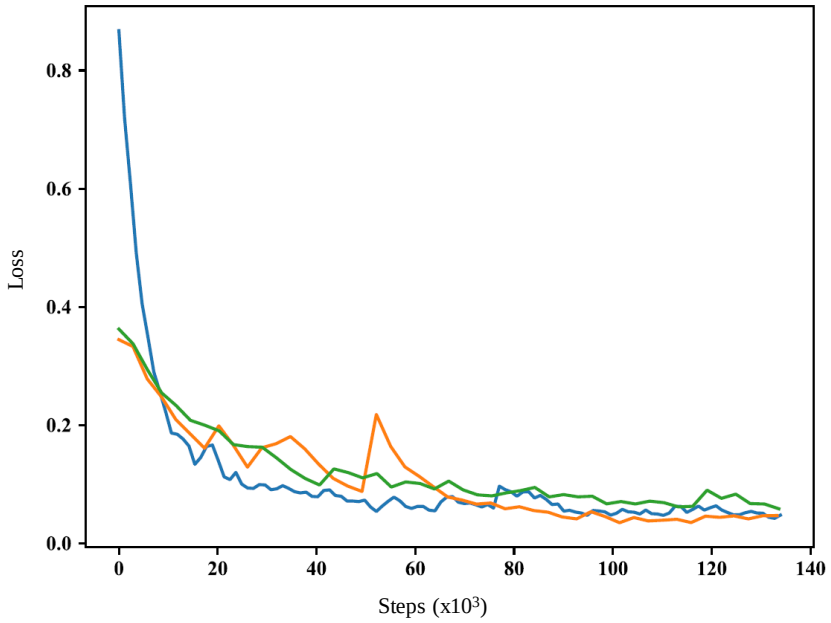


Figure 6: Learning curves in training processes when different backbone networks are employed in the encoder module of the segmentation algorithm. Blue, orange, and green lines refer to Xception 65, Resnet V1 and mobilenet V2 backbone networks respectively.

Finally, in order to evaluate the effect of the ASPP module on the segmentation performance of Deeplabv3+ method applied to H&E stained breast histopathology images, two different training setups were configured by varying the atrous rates of the dilated convolution blocks. In the first setup, the atrous rates were set to 6, 12 and 18 for the three convolutional networks, whereas in the second one, these rates were modified to 12, 24 and 36. For these models, Xception 65 was used as a backbone network.

As mentioned previously, atrous rates enable to enlarge the filter's field-of-view in convolutional operations, incorporating in this way larger spatial context.

Table 4 presents the assessment results of both models after applying a 5-fold cross-validation method. The model that performs better is the one with smaller atrous rates, which leads us to think that, for the resolution of images that make up the dataset in this work, larger filter's field-of-view may miss spatial information relevant for the segmentation process.

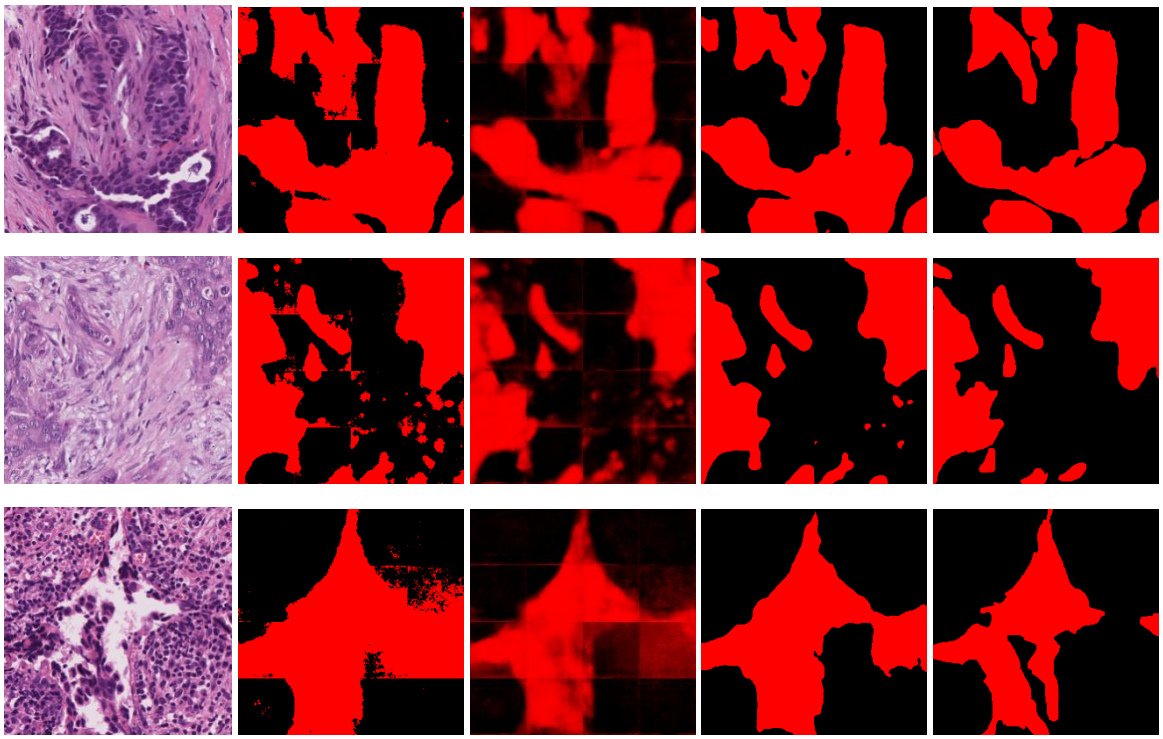
Table 4. Segmentation assessment metrics using different atrous rates in the ASPP module when Xception 65 is used as a backbone network. Acc: accuracy; Sp: specificity; Se: sensitivity; $MIoU$: mean Intersection over Union; $FWIoU$: frequency-weighted Intersection over Union.

Atrous rates	Acc (%)	Sp (%)	Se (%)	$MIoU$(%)	$FWIoU$(%)
[6,12, 18]	95.62	97.39	88.58	64.52	92.52
[12, 24, 36]	93.98	98.14	77.23	61.73	90.13

3.3 Segmentation of large histopathology images: tile-wise merging strategy results

Experiments in this section aimed to compare the effectiveness of the two different tile-wise merging strategies addressed in this work. As stated in section 2.5, tiles of 500 x 500 pixels with an overlapping rate of 50 % (250 pixels in both spatial dimensions) were considered in the splitting process in order to deal with the segmentation of the large WSIs. Figure 7 shows some examples of the segmentation results by applying both tile-wise merging strategies: column (b) in Figure 7 represents the segmentation maps obtained through the direct merging of the local segmentation results when tiles were processed independently; column (c) depicts the composed tumour tissue probability maps, which feed the input of the dense CRF; column (d) shows the CRF-based tile-wise merging results; and column (e) displays the ground-truth maps.

It can be observed that the CRF-based merging strategy reduces the impact of noise, resulting in softer boundaries and cleaner segmentation maps. Furthermore, this merging strategy preserves continuity in between the boundaries of overlapping regions. Finally, performing a qualitative evaluation by visually comparing (b), (d) and (e) columns in Figure 7, it can be noticed that the CRF-based tile-wise merging strategy provides for all cases segmentation results that are closer to the reference segmentations (ground truth maps).



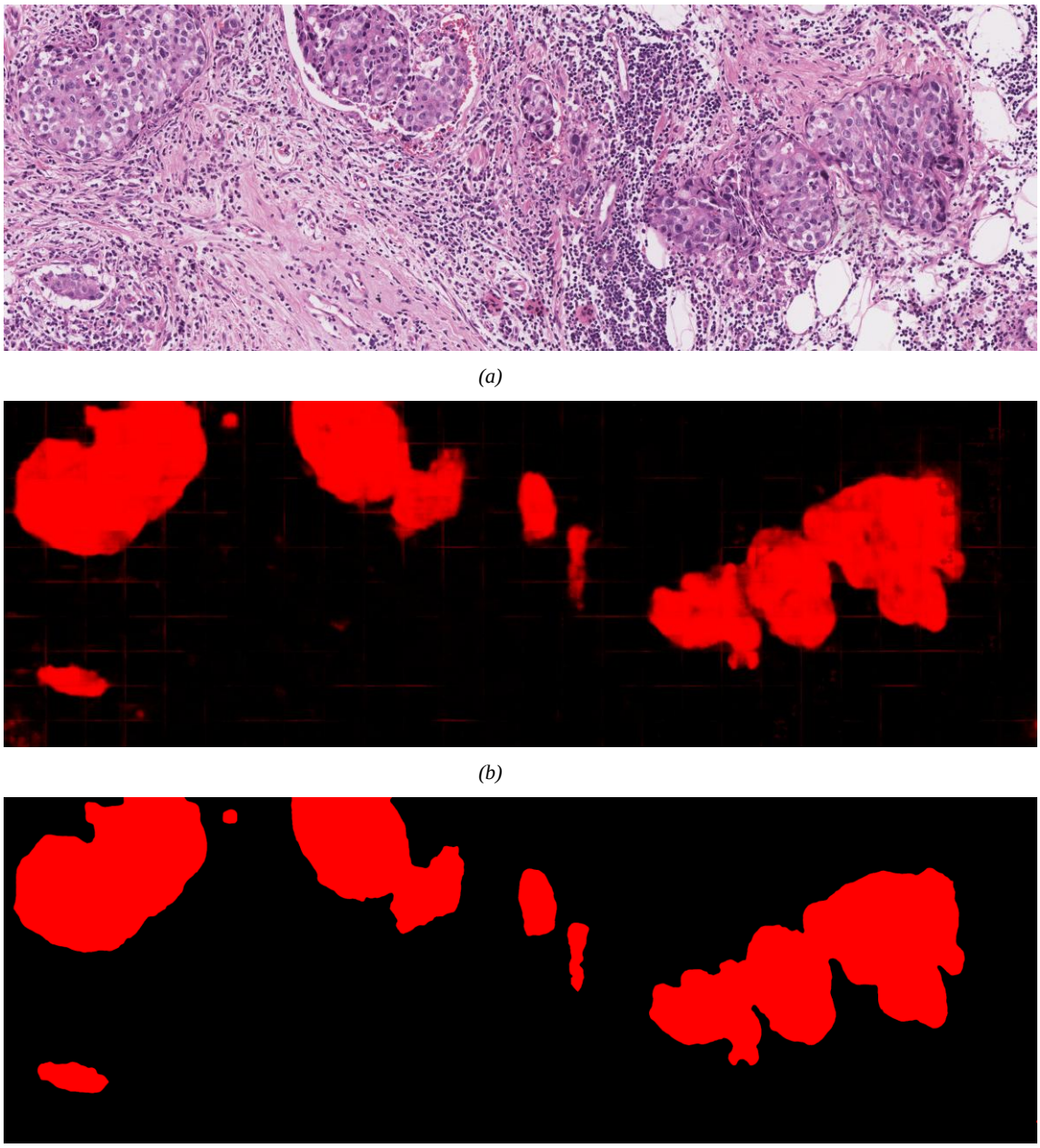
(a) *(b)* *(c)* *(d)* *(e)*

Figure 7. Tile-wise merging results. (a) RGB images of the haematoxylin and eosin (H&E) stained images, (b) direct tile-wise merging results, (c) composed tumour tissue probability maps, (d) Proposed CRF-based tile-wise merging results and (e) ground-truth maps.

Finally, Figures 8 and 9 display the segmentation maps after applying the CRF-based merging strategy to two larger histopathology images containing different types of histopathological patterns. These results also demonstrate how the proposed method performs properly regardless of the size, texture or any other colour/shape features typical of the malignant carcinomas considered in this study.

4. Conclusions and future work

In this article, a DCNN-based method to segment H&E stained breast histopathology images was discussed and evaluated. To this end, a web-based viewer and annotation tool was developed. It allowed establishing a methodology for creating datasets in collaboration with pathologists and technicians, boosting in this way the learning process of new deep-learning-based algorithms. Using this web platform, expert pathologists made segmentation annotations on image patches which composed the image datasets for the training and validation of a segmentation algorithm based on an encoder-decoder with separable atrous convolution.



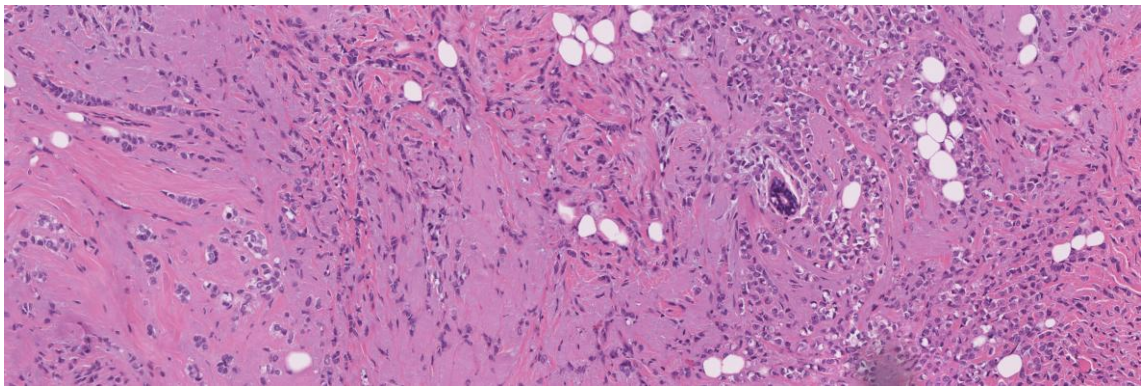
(a)

(b)

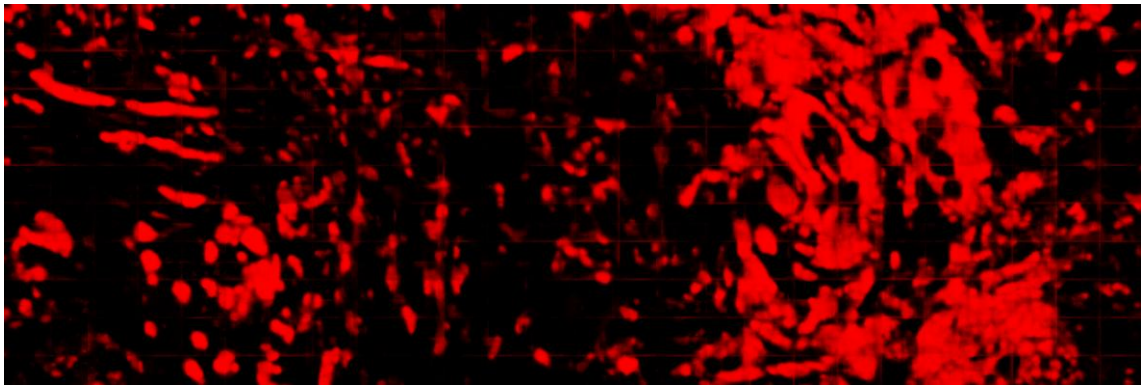
(c)

Figure 8. (a) Haematoxylin and eosin (H&E) stained RGB image, (b) composed tumour tissue probability map, (c) CRF-based tile-wise merging segmentation.

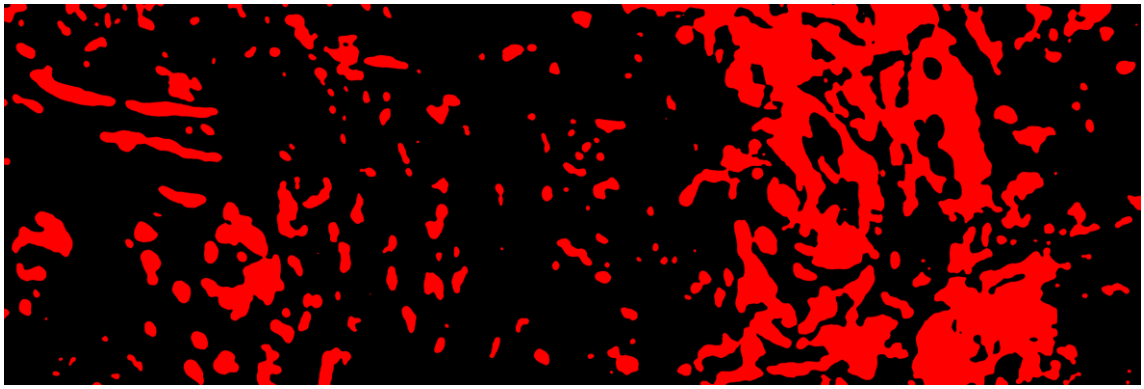
The trained segmentation models have shown to perform well in terms of standard success rates and similarity segmentation metrics, especially considering that the datasets include WSI images with high tumour variability. In this connection, no pre-processing steps were applied to RGB images prior to the segmentation process, such us colour normalization, meaning that the trained deep-learning-based segmentation method reveals to be robust to non-uniformities that may appear among WSIs, such as staining colour variations.



(a)



(b)



(c)

Figure 9. (a) Haematoxylin and eosin (H&E) stained RGB image, (b) composed tumour tissue probability map, (c) CRF-based tile-wise merging segmentation.

Finally, a new strategy to deal with complete WSI images has been proposed in this work. It is driven by the splitting of image tiles and a posterior and improved merging of local segmentation maps through dense CRF.

This study has some limitations. The processing of artefacts present in the images and arising during the process of sample handling, staining and scanning (i.e. stains, air bubbles, or unfocused areas) is not addressed in this work. The addition of a pre-processing stage for filtering these artefacts, or the addition to the training set of a significant number of images that include these elements, could lead to an improvement in the robustness of the system. Moreover, all the images used in this study were stained in the same laboratory and digitized using the same scanner. Increasing the heterogeneity of the training set with images from other sources could provide a better generalization of the results.

Meanwhile, certain future works are derived from this study. Although DCNNs are the state-of-the-art tool for semantic segmentation, they have reached a certain degree of maturity, and further improvements of segmentation quality will probably require a problem-specific approach. Nevertheless, in the specific scenario of DCNNs applied to the automated segmentation of breast tissue WSIs, there are several promising directions for future research. First, addressing the complete WSI processing has become an issue. An interesting approach would be to develop methods for the detection of hotspots, defined as regions with a high probability of containing carcinomas, and the subsequent fine segmentation of ROIs.

The fair comparison between the proposed method and other state-of-the-art segmentation algorithms applied to H&E stained breast histopathology images should be addressed as new techniques arise. The validation of the segmentation algorithm with lower resolution images, which would accelerate the processing time, will be also part of future work. Finally, another interesting avenue of future research is applying the network to other types of carcinomas. This task will require to increase the dataset.

Summing up, the results of the proposed processing pipeline have shown to be promising, revealing that this strategy performs properly regardless the size, texture or any other colour-shape features typical of the malignant carcinomas considered in this study. The algorithm has been included in a web-based platform, which includes a slide-viewer and an annotation tool, that facilitates the collaboration between pathologists and researchers to extract the specialist knowledge in form of training datasets. Currently, this tool can be used as a decision support instrument by pathologists.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

Funding: Project PI-0032-2017. Subvención para la financiación de la investigación y la innovación biomédica y en Ciencias de la Salud en el marco de la iniciativa territorial integrada 2014–2020 para la provincia de Cádiz. Consejería de Salud. Junta de Andalucía. Unión Europea, financed by the Fondo de Desarrollo Regional (FEDER).

References

Azar, A. T., & El-Said, S. A. (2013). Probabilistic neural network for breast cancer classification. *Neural Computing and Applications*, *23*(6), 1737–1751. https://doi.org/10.1007/s00521-012-1134-8

Badrinarayanan, V., Kendall, A., & Cipolla, R. (2017). SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(12), 2481–2495. https://doi.org/10.1109/TPAMI.2016.2644615

Basavanhally, A., & Madabhushi, A. (2013). EM-based segmentation-driven color standardization of digitized histopathology. *Medical Imaging 2013: Digital Pathology*, *8676*, 86760G. https://doi.org/10.1117/12.2007173

Bejnordi, B. E., Litjens, G., Timofeeva, N., Otte-Höller, I., Homeyer, A., Karssemeijer, N., & Van Der Laak, J. A. W. M. (2016). Stain specific standardization of whole-slide histopathological images. *IEEE Transactions on Medical Imaging*, *35*(2), 404–415. https://doi.org/10.1109/TMI.2015.2476509

Bejnordi, B. E., Veta, M., Van Diest, P. J., Van Ginneken, B., Karssemeijer, N., Litjens, G., … Venâncio, R. (2017). Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA - Journal of the American Medical Association*, *318*(22), 2199–2210. https://doi.org/10.1001/jama.2017.14585

Bhargava, R., & Madabhushi, A. (2016). Emerging Themes in Image Informatics and Molecular Analysis for Digital Pathology. *Annual Review of Biomedical Engineering*, *18*(1), 387–412. https://doi.org/10.1146/annurev-bioeng-112415-114722

Chan, A., & Tuszynski, J. A. (2016). Automatic prediction of tumour malignancy in breast cancer with fractal dimension. *Royal Society Open Science*, *3*(12), 160558. https://doi.org/10.1098/rsos.160558

Chen, L.-C., Papandreou, G., Kokkinos, I., Murphy, K., & Yuille, A. L. (2018). DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. In *IEEE transactions on pattern analysis and machine intelligence* (Vol. 40). https://doi.org/10.1109/TPAMI.2017.2699184

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder-decoder with atrous separable convolution for semantic image segmentation. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11211 LNCS*, 833–851. https://doi.org/10.1007/978-3-030-01234-2_49

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-Janua*, 1800–1807. https://doi.org/10.1109/CVPR.2017.195

Crystal, P., Strano, S. D., Shcharynski, S., & Koretz, M. J. (2003). Using Sonography to Screen Women with Mammographically Dense Breasts. *American Journal of Roentgenology*, *181*(1), 177–182. https://doi.org/10.2214/ajr.181.1.1810177

de Bel, T., Hermsen, M., van der Laak, J., Litjens, G. J. S., Smeets, B., & Hilbrands, L. (2018). Automatic segmentation of histopathological slides of renal tissue using deep learning. In M. N. Gurcan & J. E. Tomaszewski (Eds.), *Medical Imaging 2018: Digital Pathology* (Vol. 10581, p. 37). https://doi.org/10.1117/12.2293717

Demir, C., & Yener, B. (2005). Automated cancer diagnosis based on histopathological images: a systematic survey. In *Computer*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.61.1199&rep=rep1&type=pdf

Dihge, L., Ohlsson, M., Edén, P., Bendahl, P.-O., & Rydén, L. (2019). Artificial neural network models to predict nodal status in clinically node-negative breast cancer. *BMC Cancer*, *19*(1), 610. https://doi.org/10.1186/s12885-019-5827-6

Ferlay, J., Colombet, M., Soerjomataram, I., Mathers, C., Parkin, D. M., Piñeros, M., … Bray, F. (2019). Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods. *International Journal of Cancer*, *144*(8), 1941–1953. https://doi.org/10.1002/ijc.31937

Gil, J., Wu, H., & Wang, B. Y. (2002). Image analysis and morphometry in the diagnosis of breast cancer. *Microscopy Research and Technique*, *59*(2), 109–118. https://doi.org/10.1002/jemt.10182

Greenspan, H., van Ginneken, B., & Summers, R. M. (2016). Guest Editorial Deep Learning in Medical Imaging: Overview and Future Promise of an Exciting New Technique. *IEEE Transactions on Medical Imaging*, *35*(5), 1153–1159. https://doi.org/10.1109/TMI.2016.2553401

Gu, Y., & Yang, J. (2019). Multi-level magnification correlation hashing for scalable histopathological image retrieval. *Neurocomputing*, *351*, 134–145. https://doi.org/10.1016/j.neucom.2019.03.050

Guo, Z., Liu, H., Ni, H., Wang, X., Su, M., Guo, W., … Qian, Y. (2019). A Fast and Refined Cancer Regions Segmentation Framework in Whole-slide Breast Pathological Images. *Scientific Reports*, *9*(1), 882. https://doi.org/10.1038/s41598-018-37492-9

Gurcan, M. N., Boucheron, L. E., Can, A., Madabhushi, A., Rajpoot, N. M., & Yener, B. (2009). Histopathological image analysis: a review. *IEEE Reviews in Biomedical Engineering*, *2*, 147–171. https://doi.org/10.1109/RBME.2009.2034865

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Hinton, G. (2018). Deep learning-a technology with the potential to transform health care. *JAMA - Journal of the American Medical Association*. https://doi.org/10.1001/jama.2018.11100

Komura, D., & Ishikawa, S. (2018). Machine Learning Methods for Histopathological Image Analysis. *Computational and Structural Biotechnology Journal*, Vol. 16, pp. 34–42. https://doi.org/10.1016/j.csbj.2018.01.001

Kowal, M., Filipczuk, P., Obuchowicz, A., Korbicz, J., & Monczak, R. (2013). Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. *Computers in Biology and Medicine*, *43*(10), 1563–1572. https://doi.org/10.1016/j.compbiomed.2013.08.003

Lin, T. Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). Feature pyramid networks for object detection. *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, *2017-Janua*, 936–944. https://doi.org/10.1109/CVPR.2017.106

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., … Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, *42*, 60–88. https://doi.org/10.1016/J.MEDIA.2017.07.005

Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3431–3440. https://doi.org/10.1109/CVPR.2015.7298965

Madabhushi, A., & Lee, G. (2016). Image analysis and machine learning in digital pathology: Challenges and opportunities. *Medical Image Analysis*, *33*, 170–175. https://doi.org/10.1016/j.media.2016.06.037

Maipas, S., Nonni, A., Politi, E., Sarlanis, H., & Kavantzas, N. G. (2018). The Goodness-of-fit of the Fractal Dimension as a Diagnostic Factor in Breast Cancer. *Cureus*. https://doi.org/10.7759/cureus.3630

Nguyen, C., Wang, Y., & Nguyen, H. N. (2013). Random forest classifier combined with feature selection for breast cancer diagnosis and prognostic. *Journal of Biomedical Science and Engineering*, *06*(05), 551–560. https://doi.org/10.4236/jbise.2013.65070

Oeffinger, K. C., Fontham, E. T. H., Etzioni, R., Herzig, A., Michaelson, J. S., Shih, Y.-C. T., … Wender, R. (2015). Breast Cancer Screening for Women at Average Risk. *JAMA*, *314*(15), 1599. https://doi.org/10.1001/jama.2015.12783

Pan, S. J., & Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on Knowledge and Data Engineering*. https://doi.org/10.1109/TKDE.2009.191

Pantanowitz, L., Farahani, N., & Parwani, A. (2015). Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathology and Laboratory Medicine International*, *7*, 23.

https://doi.org/10.2147/PLMI.S59826

Peikari, M., Gangeh, M. J., Zubovits, J., Clarke, G., & Martel, A. L. (2016). Triaging diagnostically relevant regions from pathology whole slides of breast cancer: A texture based approach. *IEEE Transactions on Medical Imaging*, *35*(1), 307–315. https://doi.org/10.1109/TMI.2015.2470529

Raghavendra, U., Rajendra Acharya, U., Fujita, H., Gudigar, A., Tan, J. H., & Chokkadi, S. (2016). Application of Gabor wavelet and Locality Sensitive Discriminant Analysis for automated identification of breast cancer using digitized mammogram images. *Applied Soft Computing Journal*. https://doi.org/10.1016/j.asoc.2016.04.036

Razzak, M. I., Naz, S., & Zaib, A. (2018). Deep learning for medical image processing: Overview, challenges and the future. *Lecture Notes in Computational Vision and Biomechanics*, *26*, 323–350. https://doi.org/10.1007/978-3-319-65981-7_12

Robertson, S., Azizpour, H., Smith, K., & Hartman, J. (2018). Digital image analysis in breast pathology—from image processing techniques to artificial intelligence. *Translational Research*, *194*, 19–35. https://doi.org/10.1016/j.trsl.2017.10.010

Sanchez-Morillo, D., González, J., García-Rojo, M., & Ortega, J. (2018). Classification of Breast Cancer Histopathological Images Using KAZE Features. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *10814 LNBI*, 276–286. https://doi.org/10.1007/978-3-319-78759-6_26

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4510–4520. https://doi.org/10.1109/CVPR.2018.00474

Sharma, G. N., Dave, R., Sanadya, J., Sharma, P., & Sharma, K. K. (2010). Various types and management of breast cancer: an overview. *Journal of Advanced Pharmaceutical Technology & Research*, *1*(2), 109–126. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/22247839

Shen, D., Wu, G., & Suk, H.-I. (2017). Deep Learning in Medical Image Analysis. *Annual Review of Biomedical Engineering*, *19*(1), 221–248. https://doi.org/10.1146/annurev-bioeng-071516-044442

Spanhol, F. A., Oliveira, L. S., Petitjean, C., & Heutte, L. (2016). A Dataset for Breast Cancer Histopathological Image Classification. *IEEE Transactions on Biomedical Engineering*, *63*(7), 1455–1462. https://doi.org/10.1109/TBME.2015.2496264

Su, H., Liu, F., Xie, Y., Xing, F., Meyyappan, S., & Yang, L. (2015). Region segmentation in histopathological breast cancer images using deep convolutional neural network. *Proceedings - International Symposium on Biomedical Imaging*, *2015-July*, 55–58. https://doi.org/10.1109/ISBI.2015.7163815

Sumbaly, R., Vishnusri, N., & Jeyalatha, S. (2014). Diagnosis of Breast Cancer using Decision Tree Data Mining Technique. *International Journal of Computer Applications*. https://doi.org/10.5120/17219-7456

Sung, J. S., Stamler, S., Brooks, J., Kaplan, J., Huang, T., David, D., … Comstock, C. E. (2016). Breast cancers Detected at screening Mr imaging and Mammography in Patients at high risk: Method of Detection Reflects Tumor Histopathologic Results 1. *Radiology.Rsna.Org n Radiology*, *280*. https://doi.org/10.1148/radiol.2016151419

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-Decem*, 2818–2826. https://doi.org/10.1109/CVPR.2016.308

Ting, F. F., Tan, Y. J., & Sim, K. S. (2019). Convolutional neural network improvement for breast cancer classification. *Expert Systems with Applications*, *120*, *103-115*. https://doi.org/10.1016/j.eswa.2018.11.008

Veta, M., Pluim, J. P. W., van Diest, P. J., & Viergever, M. (2014). Breast cancer histopathology image analysis: A review. *Biomedical Engineering, IEEE Transactions On*, *61*(5), 1400–1411.

Vu, Q. D., Graham, S., Kurc, T., To, M. N. N., Shaban, M., Qaiser, T., … Farahani, K. (2019). Methods for Segmentation and Classification of Digital Microscopy Tissue Images. *Frontiers in Bioengineering and Biotechnology*, *7*, 53. https://doi.org/10.3389/fbioe.2019.00053

Wang, P., Hu, X., Li, Y., Liu, Q., & Zhu, X. (2016). Automatic cell nuclei segmentation and classification of breast cancer histopathology images. *Signal Processing*, *122*, 1–13. https://doi.org/10.1016/j.sigpro.2015.11.011

Wellings, E., Vassiliades, L., & Abdalla, R. (2016). Breast Cancer Screening for High-Risk Patients of Different Ages and Risk - Which Modality Is Most Effective? *Cureus*, *8*(12), e945. https://doi.org/10.7759/cureus.945

Xing, F., & Yang, L. (2016). Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review. *IEEE Reviews in Biomedical Engineering*, *9*, 234–263. https://doi.org/10.1109/RBME.2016.2515127

Xu, Y., Jia, Z., Wang, L. B., Ai, Y., Zhang, F., Lai, M., & Chang, E. I. C. (2017). Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics*. https://doi.org/10.1186/s12859-017-1685-x

Zhao, R., Yan, R., Chen, Z., Mao, K., Wang, P., & Gao, R. X. (2019). Deep learning and its applications to machine health monitoring. *Mechanical Systems and Signal Processing*. https://doi.org/10.1016/j.ymssp.2018.05.050