



Original Research Article

Automatic carcinoma identification from breast epithelial tissue WSI through U-Net deep learning network

Rajasekaran Subramanian^{1,*}, R. Devika Rubi², Sai Sandeep Mutyala¹

¹Dept. of Research & Development, Keshav Memorial Institute of Technology, Hyderabad, Telangana, India

²Keshav Memorial Institute of Technology, Hyderabad, Telangana, India



ARTICLE INFO

Article history:

Received 06-04-2020

Accepted 08-05-2020

Available online 04-06-2020

Keywords:

Carcinoma

Breast Cancer

Epithelial Tissue

U-Net

Epithelium Segmentation

CNN

Diagnostic Pathology

Digital Pathology

WSI Images

ABSTRACT

Epithelium tissue covers and lines all internal organs of the body. Breast cancer carcinomas arise from the epithelial cells of the breast. The epithelial component lines tubules, ducts, etc. and segmenting out this region from other tissues is important to detect breast cancer. This paper proposes application of deep learning technique U-Net, to segment epithelium tissue automatically from a whole slide image (WSI) image. It also implements U-Net, an image segmentation algorithm to automatically learn the features of epithelium components from the experts-annotated WSI image dataset and tested the results. Various image processing techniques such as thresholding are applied to improve the quality of dataset before and after the training of images by U-Net. The performance of the U-Net is measured by statistic parameters Sørensen–Dice coefficient and F1 score. The automatic system generated an epithelium segmentation of accuracy of 0.932.

Context: Automatic, faster and accurate computational technique implementation for breast cancer diagnostic pathology.

Aims: Automatic identification of carcinoma in WSI of Breast Epithelium Tissue by using deep learning network, U-Net.

Settings and Design: Epithelium tissue covers and lines all internal organs of the body. Breast cancer carcinomas arise from the epithelial cells of the breast. The epithelial component lines tubules, ducts, etc. and segmenting out this region from other tissues is important to detect breast cancer.

Methods and Material: This paper proposes application of deep learning technique U-Net, to segment epithelium tissue automatically from a whole slide image (WSI) image. It also implements U-Net, an image segmentation algorithm to automatically learn the features of epithelium components from the experts-annotated WSI image dataset and tested the results. Various image processing techniques such as thresholding are applied to improve the quality of dataset before and after the training of images by U-Net.

Statistical analysis used: The performance of the U-Net is measured by statistic parameters Sørensen–Dice coefficient and F1 score.

Results: The automatic system achieved an accuracy of 0.932 for the given dataset.

Conclusions: The automatic system achieved an accuracy of 0.932 for 40 images. System should be trained for more images to realistically achieve more accuracy for higher complex digital pathology images.

Key Messages: [S1] Automatic identification of carcinoma in WSI of Breast Epithelium Tissue by using deep learning network, U-Net.

© 2020 Published by Innovative Publication. This is an open access article under the CC BY-NC license (<https://creativecommons.org/licenses/by-nc/4.0/>)

1. Introduction

In 2018, breast cancer is the most commonly diagnosed cancer for women and the second leading cause of cancer deaths with 2 million new cases and 600,000 deaths.¹ Most

* Corresponding author.

E-mail address: rajasekarans@kmit.in (R. Subramanian).

breast cancers are carcinomas, which are tumors generally tend to originate from epithelial cells. Most cancers develop in epithelial tissue which lines body cavities. Epithelial tissues work as boundaries of the body to the outside world. Epithelial organization makes them to act as a barrier and do molecules' vectorial transport between various compartments of body. Epithelia comprises much of the tissue in glands. Carcinomas develop from epithelial tissue and they account for as many as 90 percent of all human cancers.

Epithelial tissue component of breast consists of "ductal" or "luminal" cells. Epithelial cells make milk under normal condition. In breast tissue, cancer arises predominantly from the luminal epithelial cells that line both the ducts and milk-producing lobules, and less frequently from the outer layer of basal cells. Cancer follows from a series of somatic mutations. Somatic mutations are genetic alterations acquired by the cell, which is passed to the progeny of the mutated cell in the course of cell division. Renewal of epithelial tissues showing large number of cell replications lead to an increase of somatic mutations, tending to carcinoma incidence. Carcinoma, a cancer starts in the skin or the tissues that line other organs, especially Epithelial tissue component of breast.

Breast cancer can be identified and diagnosed with the help of ultrasound, X-rays, but the conclusive evidence of cancerous cells is given by WSI images. WSI images are gigapixels in size and it is manually difficult and time consuming to analyze them. Deep Learning has shown super human performance in classification, object detection, segmentation tasks and can save a lot of human effort and help to improve the diagnostic accuracy, if the analysis can be automated.

Segmenting epithelium regions is crucial as cancer carcinomas typically arise in the epithelium.² This will allow pathologists to separate out relevant cancer tissue. Since it filters out relevant regions of an image, this will also reduce the amount of computation used. The dataset contains breast epithelium images along with their respective mask images (ground truth) which are annotated by experts.³ However, this is a difficult task as often non-relevant tissue finds its way into the masks. So, the quality of segmentation heavily relies on the quality of the dataset.

This paper uses an image segmentation model U-Net to segment out the epithelium tissue of an image.⁴ Image preprocessing techniques such as thresholding, noise reducing methods to remove small connected cells components, etc, are performed to obtain finer mask images.⁵ Similarly, image denoising techniques like connected component removal, dilation, closing, etc. are performed in the postprocessing to improve the accuracy of results. Binary cross-entropy loss function is used to calculate loss energy weight in the images. U-Net uses Adam optimizer algorithm to learn the features of

the dataset. Statistical parameters such as dice, Jaccard coefficients, accuracy and F-score are used as performance parameters to evaluate U-Net. Data augmentation is used to increase the size of the dataset. Learning rate would be high if more data are fed into the U-Net network.

Each WSI image is gigapixels in size and analyzing the entire image manually is extremely difficult which consumes lot of pathologist's time. Due to the time-consuming and tedious nature of the diagnosis of breast cancer, there is a growing need for it to be automated. This paper envisions an Epithelium segmentation model which will help the experts to perform fast and accurate diagnosis task.

Anant Madabhushi et al introduces the usage of Deep Learning Techniques for Digital Pathology images processing, with various onco-pathology diagnostic use cases such as Carcinoma Localization, Nuclei Segmentation, Epithelium Segmentation, Tubule Segmentation, Lymphocyte detection, Mitosis Detection and Lymphoma classification.⁶ Authors have used 34 digital pathology (WSI-Whole Slide Images of 1000 x 1000 pixels) training images and 8 validation images to perform the Epithelium Segmentation use case. Each single WSI image is divided into multiple patches of size 32 x 32 pixels. To remove uninterested stroma regions from the mask of grayscale image, a user-defined thresholding is applied. Patches containing the edges of epithelium regions are taken so that the network can learn crisp boundaries. AlexNet, the popular Convolutional Neural Network (CNN) algorithm is applied for image classification on the generated patches. While generating the output, white regions are removed again by applying user-defined thresholding. And positive regions of size less than 300 pixels are removed, which aren't clinically relevant. This method has generated an average F1 score of 0.84.

Segmentation is the same as classification of each pixel in the image. Ciresan et al used a CNN to classify each pixel using its local patch.⁷ As the number of patches generated through this technique were really large and the CNN has to run over each one of them which leads to high runtime.

With the introduction of U-Net, which has better speed and accuracy than the other existing models of the time, image segmentation algorithms started using U-Net model. U-Net has processes called down-convolutions and up-convolutions. Down-convolutions facilitate data localization, whereas up-convolutions generate an image of matching size. Its concatenation layers help to prevent context loss and generate a clean image. All these processes and concatenation layers of U-Net facilitate faster runtime and better accuracy than other CNNs.

Wouter Bulten et al proposes the usage of U-Net for prostate epithelium segmentation on Hematoxylin and Eosin (H&E) stained prostatectomy slides which uses immunohistochemistry (IHC) as reference standard.⁸

Nearly 102 tissue sections were stained with H&E and subsequently re-stained with P63 and CK8/18 IHC markers to highlight epithelial structures. The H&E and IHC images are co-registered using a non-linear image registration method. Due to multi-staining of the image, the co-registration pipeline follows specialized methodology to account for the multi-modality registration problem. The co-registration pipeline initially converts image to grayscale, followed by parametric (affine) registration of the same, non-parametric registration and finally a patch-based registration with Normalized Gradient Fields (NGF) and Curvature. Non-parametric registration uses NGF distance which measures the alignment of image gradients, to account for the multi-modality registration problem. By applying color deconvolution on IHC images P63/CK8-18 components are separated. Subsequently, using thresholding each channel was then converted to binary mask. The U-Net is trained with 50 training images and 12 validation images. The network's output is mapped to H&E version of the specimens using registration which were fed to our final model as inputs.

Philipp Kainz et al proposed classic LeNet-5 Network architecture to semantically segment glands in the Warwick-QU dataset, presented at GlaS@MICCAI 2015 challenge.^{9,10} The dataset contains 161 H&E stained images taken at 20x magnification of benign and malignant colorectal adenocarcinomas. Authors proposed 3-steps pipeline, comprising preprocessing, with 2 Convolutional Neural Networks namely Object-Net and Separator-Net. Object-Net network is used to predict glands and Separator-Net network is used to segment gland-separating structures. The proposed methodology has color deconvolution and CLAHE equalization as preprocessing steps. CLAHE equalization helps to equalize training images contrasts. Object-Net network provides the probability of a pixel belonging to a gland or background, whereas Separator-Net network identifies structures which separate the gland from the background objects. The model generated an F1 score of 0.78 without Separator-Net and 0.87 with Separator-Net.

CNNs were used for brain tumor classification and segmentation.¹¹ Inputs of size 224 x 224 pixel are passed into the network. Initially, convolutional layers are applied to down-sample the image. Subsequently, the outputs are flattened and linear layers are applied on them. Finally, a 4096 dimensional feature extractor was extracted. The model attained classification accuracy of 97.5% and got 0.84 dice coefficient for segmentation.

1.1. Subjects and Methods

1.2. Model Description

Convolutional Neural Network (CNN) is the class of deep neural networks, used to analyze images. U-Net is a Deep Learning Segmentation model developed by

Olaf Ronneberger et al for fast and accurate image segmentation. It was originally developed for biomedical image segmentation, but has become a standard image segmentation model for all domains. CNNs are typically used for classification tasks. Segmentation can also be seen as a classification task, as each pixel needs to be classified into one of several types. Prior to U-Net a sliding window technique was used along with a CNN to classify each pixel, by looking at the region around it.¹² The run-time of sliding window technique was high because the CNN need to run huge number of times before the entire image gets classified. U-net outperforms this task in terms of both speed and accuracy. U-Net model was also used for object detection and object counting tasks, as proposed by Hao Dong et al and Falk, T et al.^{13,14} For detection, critical objects can be accurately outlined and segmented out using U-Net.

U-Net consists of two primary parts namely down-sampling (shown at left side in Figure) and up-sampling (Figure 1). The down-sampling comprises of 5 convolutional blocks, with each block consist of two 3x3 convolutional layers and a max pooling layer. Similarly, up-sampling consists of 5 reverse-convolutional blocks. Each convolutional block consists of a 2x2 convolution transpose layer, a concatenate layer which concatenates the current activations with their respective down-sampled activations and two 3x3 convolutional layers. Finally, the activations are passed through a 1x1 convolution layer and a soft max activation layer. Each convolution block is followed by a ReLU. The concatenate layer prevents loss of context and detail. The 1x1 convolutional layer maps the feature vector for each pixel into the desired number of classes. The soft max is used to calculate class probabilities of each pixel, and can be replaced with a sigmoid layer for binary classification. The entire architecture is shown is Figure 1.

Performance of a CNN model can be assessed using a Loss function. Binary-cross-entropy loss function, which classifies each image pixel into two classes viz Presence (positive/White color) and Absence (Negative/Black color) of epithelium. Critical regions or interested regions are very limited on most of the biomedical images and critical regions occupy a small portion on the image map, which leads to an unbalanced optimization problem.

The experimental use case dataset is also susceptible to unbalanced optimization problem due to the existence of sparse regions within the images. As negative regions (non-epithelium regions) within the images are common, which may lead to data overfitting, while training dataset becomes unbalanced (Figure 2). By applying dice loss which focus on positive regions (regions containing epithelium), avoids data overfitting.¹⁵ U-Net uses Adam optimizer, a modified Stochastic Gradient Descent optimizer for smoother and finer learning curves.¹⁶

Binary cross-entropy and dice loss functions are added and the model is trained on the new function. The optimizer

is used with $b_1 = 0.9$, $b_2 = 0.999$, learning rate = 0.001, and batch size = 4. The learning rate is exponentially decayed by 0.5 every 400 training steps.

1.3. Model Evaluation Metrics

The model is evaluated by using the metrics, dice coefficient, Jaccard coefficient, accuracy, and F1-score.^{17,18} Accuracy is usually the most common way to evaluate models. However, a model evaluated by accuracy may fail to identify small, positive, and critical regions. In imbalanced datasets, a model may have high accuracy, but may fail to identify such regions. To properly assess the model in such situations, dice, jaccard coefficients, and F1-score are used. A detailed comparison of the metrics was done by Vikas Thada et al.¹⁹

1.4. Data

Epithelium is a tissue which lines the outer surfaces of organs in the body. The digitized biopsy image (WSI, whole slide image) generally contains regions like stroma, tubules etc, along with epithelium. To perform epithelium segmentation task on WSI image using CNN requires a sample image along with respective expert-annotated mask showing the region containing epithelium tissue on WSI. Epithelium region is shown as the dark purplish region in the Figure 3.

The experimental dataset consists of 42 epithelium tissue images of size 1000x1000 pixels with their respective mask images. To fit into CNN model, each image is appropriately resized and divided into patches of size 320x320 leading to 378 pairs (image and mask).

Data augmentation is the process of performing transformations over existing data to produce new, additional data samples. Data augmentation techniques are used to increase the size of training data and obtain a better fit on the model. Few sample images, masks and respective data augmentations are shown in Figure 4. These images are augmented using random horizontal, vertical flips, random crops and rotations before being passed into the model. Perez et al demonstrates the effectiveness of data augmentation in improving deep learning models.²⁰ A SmallNet was used on the Cats vs Dogs dataset. With data Augmentation validation accuracy improved from 0.705 to 0.775. On the dogs vs goldfish dataset, validation accuracy improved from 0.855 to 0.890 using traditional data augmentation, and 0.915 using neural data augmentation.

1.5. Pre and Post processing Methods

Generally, there would be few white stroma regions within the epithelium, which are considered to be epithelium by mask-annotations as shown in Figure 5. Similarly, there also small purplish epithelium regions outside the epithelium chunks, those were not marked as epithelium

region as shown in Figure 6. This confuses the training network and predict them as noisy. Pre and post processing techniques which are discussed as given below would override these issues.

The annotated masks in the given dataset are slightly inconsistent, marking even white stroma regions as epithelium at few places. Binary thresholding is applied to remove these mis-markings. Initially, the original image is converted to grayscale, where the value of each pixel corresponds to the darkness at that point. Over these images, apply binary thresholding value of 0.5, wherein all positive pixels with a value less than a threshold value of are marked as negative in the mask as shown in Figure 7. Thus, this creates negative spaces in the mask wherever there are white regions in the original image and reduces the number of wrongly marked stroma regions.

1.6. Post-processing Methods

Connected components are groups of connected pixels belonging to the same class in the segmented image. Removing small connected components within an image is an effective way to de-noise a binary image. Coupier et al and Chan et al describe how removing connected components can assist in denoising an image.^{21,22} From the generated mask, white generated components which contain fewer than 300 pixels are removed. This removes most white noise in the image, and reduces the issue of falsely marked non-epithelial regions. The image is slightly dilated to close tiny gaps between epithelial regions. This process is illustrated in Figure 8.

2. Results

The automatic system generated an accuracy of 0.932 for 40 images. The system generated Sørensen–Dice coefficient, a statistic used to gauge the similarity of two samples as 0.7 as shown in the Table 1. The system achieved a precision of 0.94 and an F1 score (also F-score or F-measure), a measure of a test's accuracy in a binary classification problem as 0.86 which is considerably better than the previous results as shown in the Table 2.

3. Discussion

This paper discussed an automatic image segmentation method to segment out epithelium tissue from WSI images of the breast cancer using a CNN network called U-Net. This method used the combination of binary cross-entropy and dice loss functions to counter imbalanced training. Adam optimizer was used for training. Data augmentation techniques like rotation, cropping, horizontal and vertical flips, etc. are used to artificially increase the dataset size. This paper also discussed about information loss issues in the dataset, like the presence of false positives and false negatives and discussed the pre-processing techniques like

Table 1: Dice Coefficient and Jaccard Coefficient

Metric Value	Test-set size (Images/Patches)	Dice Coefficient	Accuracy	Jaccard
	7/63	0.7	0.932	0.753

Table 2: F-Score and Precision

Metric Value	F1-Score	Precision	Recall
	0.86	0.94	0.792

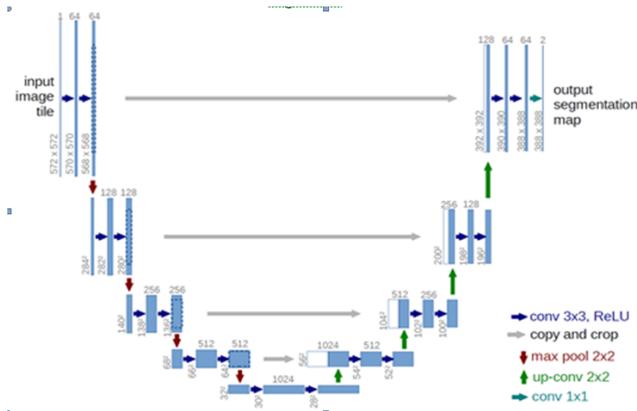


Fig. 1: U-Net model architecture

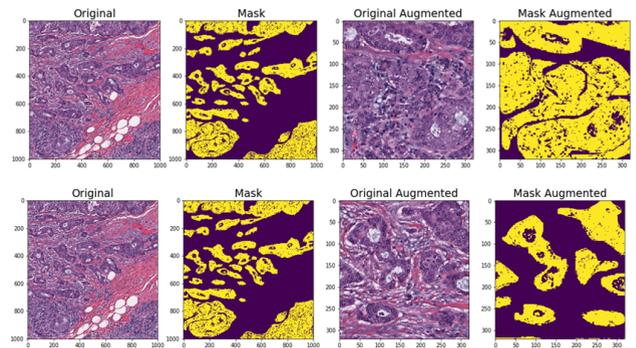


Fig. 4: Image Augmentation Samples

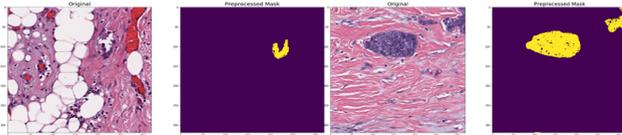


Fig. 2: Examples of unbalanced training samples

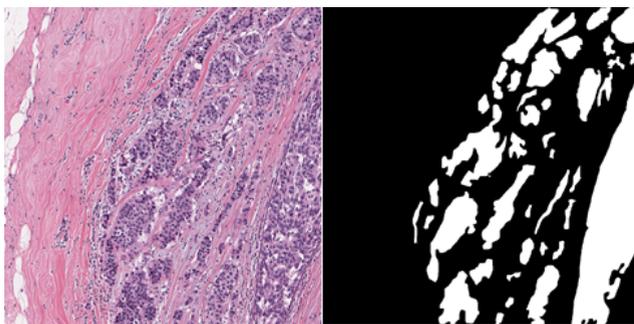


Fig. 3: Image– Mask Pair

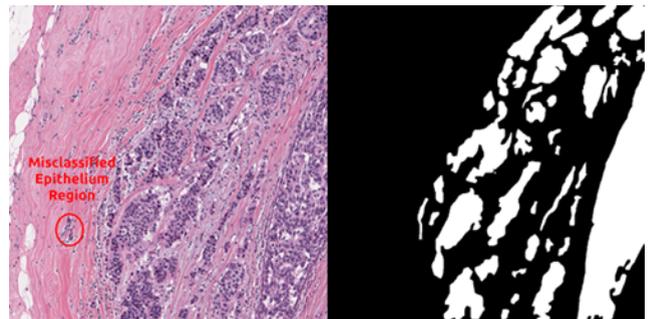


Fig. 5: Example of misclassified epithelium

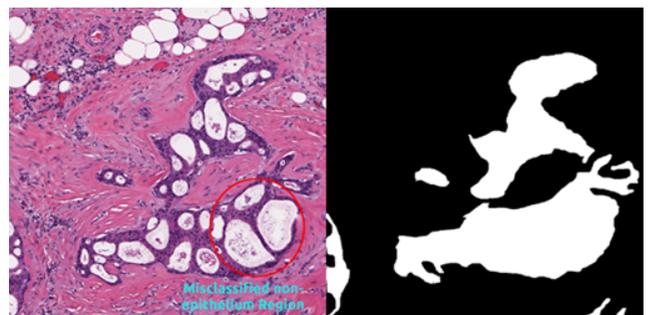


Fig. 6: Example of misclassified non-epithelium

binary thresholding to avoid the information loss. It also discussed about postprocessing techniques like blurring, removing small connected components, etc. to generate cleaner and more interpretable images.

The main disadvantage over this method is noise persistence, despite applying pre and post processing techniques. Due to the existence of few false-positives and false negatives in the data and generalized boundaries, noise

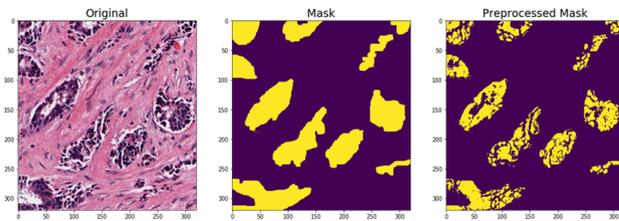


Fig. 7: Preprocessing mask

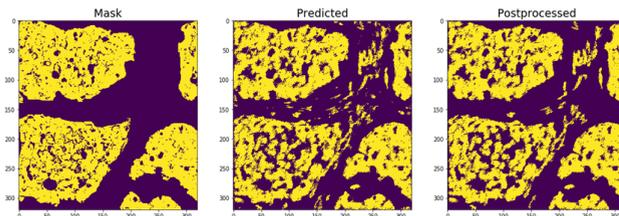


Fig. 8: Post processing generated mask

persists in the output. The postprocessing and dilating the output image can help to clean up the output generated and make it accessible to a human viewer, but still some unnecessary regions do get in which can be avoided by applying denoising techniques over the output image.

4. Acknowledgement

None.

5. Source of Funding

None.

6. Conflict of Interest

None.

References

1. American Cancer Society. American Cancer Society; 2018. Available from: <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2018/cancer-facts-and-figures-2018.pdf>.
2. Andrew B, Ankur S, Samuel L, Robert M, Torsten N, Marc V, et al. Systematic Analysis of Breast Cancer Morphology Uncovers Stromal Features Associated with Survival. *Sci Transl Med*. 2011;3.
3. Janowczyk A. Deep Learning Datasets ; 2018. Available from: <http://www.andrewjanowczyk.com/deep-learning/>.
4. Ronneberger O, Fischer P, Brox T, U-Net. Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Springer. 2015;p. 9351.
5. Jeyavathana R, Balasubramanian R, Anbarasa P, Survey. A Survey: Analysis on Pre-processing and Segmentation Techniques for Medical Images. *Int J Res Scientific Innov*. 2016;.
6. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. *J Pathol Inform*. 2016;7(1):29.
7. Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Deep Neural Networks Segment Neuronal Membranes in Electron Microscopy

Images. . *NIPS*. 2012;p. 2852–60.

8. Bulten W, Bándi P, Hoven J, van de Loo R, Lotz J, Weiss N, et al. Epithelium segmentation using deep learning in H&E-stained prostate specimens with immunohistochemistry as reference standard. *Scientific Rep*. 2019;9(1):864.
9. Philipp K, Michael P, Martin U. Semantic Segmentation of Colon Glands with Deep Convolutional Neural Networks and Total Variation Segmentation. ArXiv e-prints; 2015.
10. Glas@MICCAI'2015: Gland Segmentation Challenge Contest; 2015. Available from: <https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest>.
11. Xu Y, Jia Z, Ai Y, Zhang F, Lai M, Chang EI. Deep convolutional activation features for large scale Brain Tumor histopathology image classification and segmentation. In: Conference on Acoustics, Speech and Signal Processing (ICASSP), Brisbane, QLD; 2015. p. 947–951.
12. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer J Clin*. 2018;68(6):394–424.
13. Dong H, Yang G, Liu F, Mo T, Guo Y. Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks. Berlin Heidelberg: Springer-Verlag; 2011.
14. Falk T, Mai D, Bensch R. U-Net: deep learning for cell counting, detection, and morphometry. *Nat Methods*. 2019;(16):67–70.
15. Carole S, Wenqi L, Tom V, Sebastien O, and CMJ. Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. Lecture Notes in Computer Science Book Series (LNCS). Springer; 2017. p. 240–8.
16. Diederik K, Jimmy B. Adam: A Method for Stochastic Optimization. In: International Conference on Learning Representations. 3rd International Conference for Learning Representations. San Diego; 2015.
17. Reuben S, Yuval D, Jinyoung K, Guillermo S, Noam H. Continuous Dice Coefficient: a Method for Evaluating Probabilistic Segmentations. bioRxiv preprint; 2018.
18. Shi R, Ngan KN, Li S. Jaccard index compensation for object segmentation evaluation. In: 2014 IEEE International Conference on Image Processing (ICIP), Paris; 2014. p. 4457–61.
19. Thada V, Jaglan V. Comparison of Jaccard, Dice, Cosine Similarity Coefficient To Find Best Fitness Value for Web Retrieved Documents Using Genetic Algorithm. *Int J Innov Eng Technol (IJJET)*. 2013;2(4).
20. Perez L, Wang J. The Effectiveness of Data Augmentation in Image Classification using Deep Learning; 2017. Available from: <https://arxiv.org/abs/1712.04621>.
21. Coupier D, Desolneux A, Ycart B. Image Denoising by Statistical Area Thresholding. *J Math Imaging Vision*. 2005;22(2-3):183–97.
22. Chan TF, Esedoglu S, Nikolova M. Algorithms for Finding Global Minimizers of Image Segmentation and Denoising Models. *SIAM Journal on Applied Mathematics*. 2006;66(5):1632–1648. Available from: <https://dx.doi.org/10.1137/040615286>. doi:10.1137/040615286.

Author biography

Rajasekaran Subramanian Professor

R. Devika Rubi Professor

Sai Sandeep Mutyala Student

Cite this article: Subramanian R, Rubi RD, Mutyala SS. Automatic carcinoma identification from breast epithelial tissue WSI through U-Net deep learning network. *IP J Diagn Pathol Oncol* 2020;5(2):163-168.