

# Impacts of Color Space Transformations on Dysplastic Nuclei Segmentation Using CNN

Dalí F. D. dos Santos\*, Adriano B. Silva, Paulo R. de Faria, Bruno A. N. Travençolo, Marcelo Z. do Nascimento  
Federal University of Uberlândia, Uberlândia, Minas Gerais, Brazil  
\*dalifreire@gmail.com

**Abstract**—Oral epithelial dysplasia is a common precancerous lesion type that can be graded as mild, moderate and severe. Although not all oral epithelial dysplasia become cancer over time, this premalignant condition has a significant rate of progressing to cancer and the early treatment has been shown to be considerably more successful. The diagnosis and distinctions between mild, moderate, and severe grades are made by pathologists through a complex and time-consuming process where some cytological features, including nuclear shape, are analysed. The use of computer-aided diagnosis can be applied as a tool to aid and enhance the pathologist decisions. Recently, deep learning based methods are earning more and more attention and have been successfully applied to nuclei segmentation problems in several scenarios. In this paper, we evaluated the impact of different color spaces transformations for automated nuclei segmentation on histological images of oral dysplastic tissues using fully convolutional neural networks (CNN). The CNN were trained using different color spaces from a dataset of tongue images from mice diagnosed with oral epithelial dysplasia. The CIE L\*a\*b\* color space transformation achieved the best averaged accuracy over all analyzed color space configurations (88.2%). The results show that the chrominance information, or the color values, does not play the most significant role for nuclei segmentation purpose on a mice tongue histopathological images dataset.

**Index Terms**—CNN, deep learning, dysplasia, nuclei segmentation, color spaces

## I. INTRODUCTION

Cancer can be defined as a group of non-communicable diseases (NCD) that can start almost anywhere in the body. The disease is caused when cells begin to divide uncontrollably with potential to invade other parts of the body and/or spread to other organs and surrounding tissues. According to the World Health Organization (WHO), cancer is the second leading cause of death globally, accounting for 18.1 million cases and 9.6 million deaths worldwide in 2018 [1]. Projections from the Instituto Nacional do Cancer (INCA) show that in 2020, 686 thousand new cancer cases will be registered in Brazil and, from those cases, 15 thousand will be oral cavity-derived cancer [2].

Tumors of the oral cavity can be grouped into two broad categories: benign and malignant tumor (cancer). Benign tumors are not considered cancer, as they do not invade other tissues and do not spread to other parts of the body. Precancerous

conditions, on the other hand, are still harmless lesions, but some of those precancerous conditions can give rise to cancer over time.

Dysplasia is an important precancerous condition characterized by the presence of abnormal cells in the oral mucosa. The histological evaluation, i.e. the study of tissue samples of affected region under the microscope, remains the most reliable way for diagnosing and grade oral epithelial dysplasia [3], [4]. Despite being referred as gold-standard in cancer diagnosis, the histological evaluation is plagued by inter- and intra-observer variability problem. This difficulty requires experienced pathologists and is an expensive, highly skilled and very time-consuming process.

Fast scanners can be used to capture tissue samples into digital images to obtain the so-called digital histological images, allowing to view tissue samples on computer rather than through a microscope. Digitized histological samples can be analyzed by computational image processing techniques to aid and enhance the pathologist decision making, minimizing human interventions, discovering measurable and traceable clinical information, providing reliable prognostic factors and eliminating the inter- and intra-observer variability.

There are several potential applications of digital pathology, such as nuclei detection and segmentation that are critical prerequisite steps for diagnosing and grading dysplasia in image-based computer-aided diagnosis (CAD). Nuclei extracted features are critical for evaluating the existence of diseases and its severity. Furthermore, nuclei commonly appear in overlapping clusters, have heterogeneous aspects and remain a challenging problem, which keeps nuclei segmentation methods under investigation [5]. Particularly, deep learning techniques, such as convolutional neural networks (CNN), has been successfully applied in medical and biological researches [6].

As the color is one of the most dominant and visually distinguishable visual properties, color variations could play a high influence on automated analysis of histological images. This paper evaluates the impact of different color space transformations applied to our previously proposed method for automated nuclei segmentation on dysplastic oral tissue histological images using fully convolutional neural networks [7]. For this, CNN models were trained using different color spaces from a dataset of tongue images from epithelial dysplasia-harboring mice.

The rest of the paper is organized as follows. The next

This work was supported by the National Council for Scientific and Technological Development - CNPq (Grant 304848/2018-2) and the State of Minas Gerais Research Foundation - FAPEMIG (Grant APQ-00578-18).

section describes some important background concepts. Section III presents the experimental evaluation and the Section IV presents the main results achieved and its discussion. Finally, the Section V concludes the paper and presents further work directions.

## II. BACKGROUND REVIEW

### A. Color Spaces

The purpose of a color space is to facilitate the specification of colors in some standard providing a coordinate system and a subspace in which each color is represented by a single point. There have been numerous different color spaces in use today. Some of these color spaces are ideally suited for hardware implementations and others for the way that humans describe and interpret colors [8]–[10].

In this paper, we focused on using RGB, HSV and CIE  $L^*a^*b^*$  color spaces in hematoxylin-eosin (H&E) stained histological images for nuclei segmentation purposes:

- RGB (red, green, blue) color space defines each color as a combination of the three primary spectral components: red, green, and blue. The RGB color space is the hardware-oriented color space most widely used for a broad class of video cameras and color monitors.
- HSV (hue, saturation, value/brightness) color space is a nonlinear transformation of the RGB that describes the pure color (hue) in terms of gray presented in each color (saturation) and how bright the color is (value). The HSV color space corresponds closely with the way humans describe and interpret color.
- CIE  $L^*a^*b^*$  was defined by the International Commission on Illumination (CIE) and is also based on human perception. The  $L^*$  channel indicates lightness and  $a^*$  and  $b^*$  channels indicate chromaticity directions:  $a^*$  indicates the color value between green and red and  $b^*$  indicates the color value between blue and yellow.

### B. Histological Images Dataset

The histological images dataset was built using H&E-stained tongue slides extracted from 30 mice previously diagnosed with oral epithelial dysplasia. The images were digitized using a Leica DM500 light microscope with original magnification of  $400\times$ . A total of 66 images were scanned and saved in TIFF format using the RGB color space and resolution of  $2048\times 1536$  pixels. A experienced pathologist used the criteria described by Lumerman et al. [11] to classify each image into four predominant classes: healthy tissue, mild, moderate or severe dysplasia.

The digitized images were then cropped into regions of interest (ROI) with size of  $448\times 256$  pixels, totalling 120 ROI images – 30 ROI images for each class. Examples of the produced histological images are shown in Fig. 1 and examples of the extracted ROI images are shown in the first column in Fig. 3 (3a, 3e, 3i, 3m).

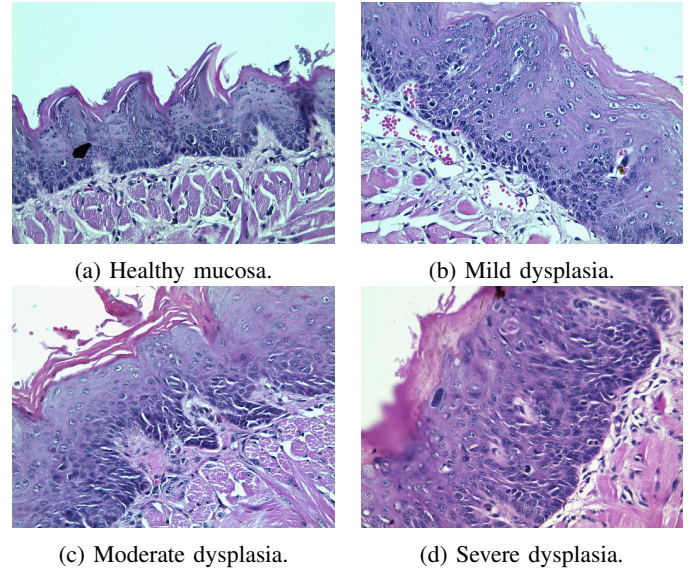


Fig. 1: H&E stained histological images of mice oral epithelial tissues.

### C. Automated Nuclei Segmentation Using CNN

The method used for nuclei segmentation in oral tissue histological images was originally proposed by dos Santos et al. [7]. The proposed CNN architecture, which was slightly modified to support images with one or three color channels as input, is depicted in Fig. 2.

Deep learning segmentation models requires a large number of samples and their corresponding segmentation masks to be properly trained. Since producing these image samples is a very hard and time consuming-task, data augmentation becomes an essential technique to overcome this general problem of scarcity of available training samples [13]. Six different image transformation techniques were used together as data augmentation: horizontal/vertical flip, rotation, elastic transformation, grid distortion and optical distortion. Some examples of data augmentation applied to the ROI images and their corresponding targets are illustrated in the most right two columns in Fig. 3 (3c, 3d, 3g, 3h, 3k, 3l, 3o, 3p).

In the training step, we directly apply the histological ROI images (e.g., Figs. 3a, 3c, 3e, 3g, 3i, 3k, 3m and 3o) and their corresponding masks (e.g., Figs. 3b, 3d, 3f, 3h, 3j, 3l, 3n and 3p) to the deep neural network to train the model. After the last convolution step be performed, the Otsu [14] threshold is applied to binarize the resulting mask image that predicts the nuclei for the input ROI image.

In order to investigate the influence of color space transformations on the evaluated nuclei segmentation method, we transformed the 120 ROI images from the original training and test sets into three different color spaces (RGB, HSV, and CIE  $L^*a^*b^*$ ). We retrained the nuclei segmentation model using the grayscale image and each color space and their respective individual channels separately, resulting in 13 new training datasets (thirteen sets of 96 images) and performed evaluations

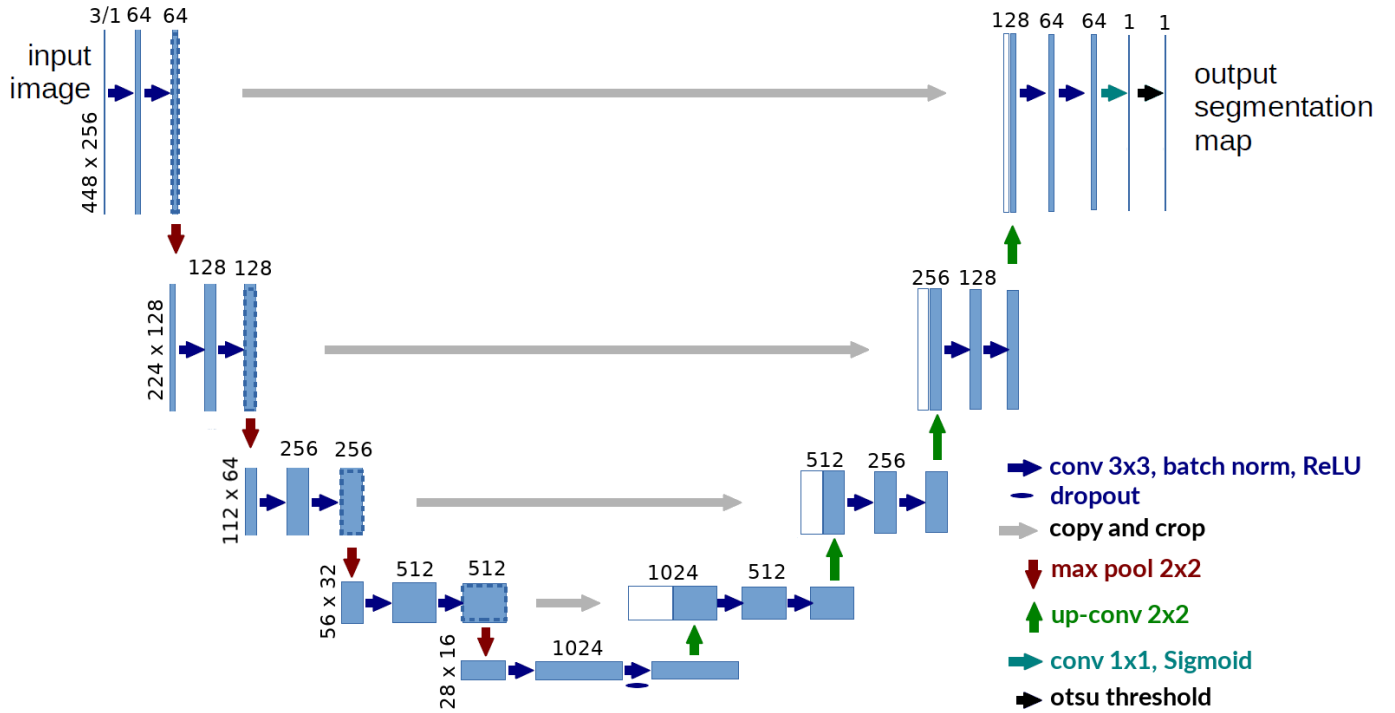


Fig. 2: Adaptation of the CNN architecture proposed in [7], which is based on U-Net [12] model. The number of convolutional feature channels and the height×width of images for each layer are denoted on their corresponding boxes.

using the test set (24 images transformed thirteen times for each corresponding color space configuration).

The nuclei segmentation model was implemented using the PyTorch framework [15]. The models were trained using a desktop computer (Intel Core i7 3.4GHz×8 processor, 32 GB memory, 1TB SSD) equipped with GeForce GTX 1050 Ti graphic card and Ubuntu 20.04 operational system. The elapsed time to train the models with 500 epochs for each color space transformation was about 400 minutes. After training, the elapsed time to process an input image was about 0.3 seconds. To provide better understanding and make this work as reproducible as possible, the source code is publicly available at: [https://github.com/dalifreire/dysplastic\\_oral\\_tissues\\_segmentation](https://github.com/dalifreire/dysplastic_oral_tissues_segmentation).

### III. EXPERIMENTAL EVALUATION

We performed experimental evaluations using thirteen configurations for the 24 ROI images from the test subset: Grayscale image, RGB, only the R channel from RGB, only the G channel from RGB, only the B channel from RGB, HSV, only the H channel from HSV, only the S channel from HSV, only the V channel from HSV, CIE L\*a\*b\*, only the L\* channel from CIE L\*a\*b\*, only the a\* channel from CIE L\*a\*b\* and only the b\* channel from CIE L\*a\*b\*.

The influence of different color space transformations on the automated nuclei segmentation using fully convolutional neural network [7] was investigated. The automated nuclei segmentation results were compared quantitatively and qualitatively (visually) with the nuclei manually segmented by

the specialist. The segmentation performance was measured calculating the overlapping regions of the resulting automated segmented image and the regions of a reference image segmented by the specialist. The average performance was measured by six of the most commonly used quantitative criteria with respect to pixel classification (nuclei or non-nuclei). Accuracy is defined as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (1)$$

Precision indicates if the segmentation results bring only nuclei areas and does not bring any non-nuclei areas:

$$Precision = \frac{TP}{TP + FP}. \quad (2)$$

Sensitivity (or Recall) indicates if the segmentation results bring all the nuclei areas and is defined as:

$$Sensitivity = \frac{TP}{TP + FN}. \quad (3)$$

F1 score (or Dice Coefficient) is the harmonic mean of the precision and recall and can be defined as:

$$F1 = \frac{2 \cdot TP}{2 \cdot TP + FP + FN}. \quad (4)$$

Jaccard index emphasizes similarity between gold-standard and segmentation results for both nuclei and non-nuclei areas and is defined as:

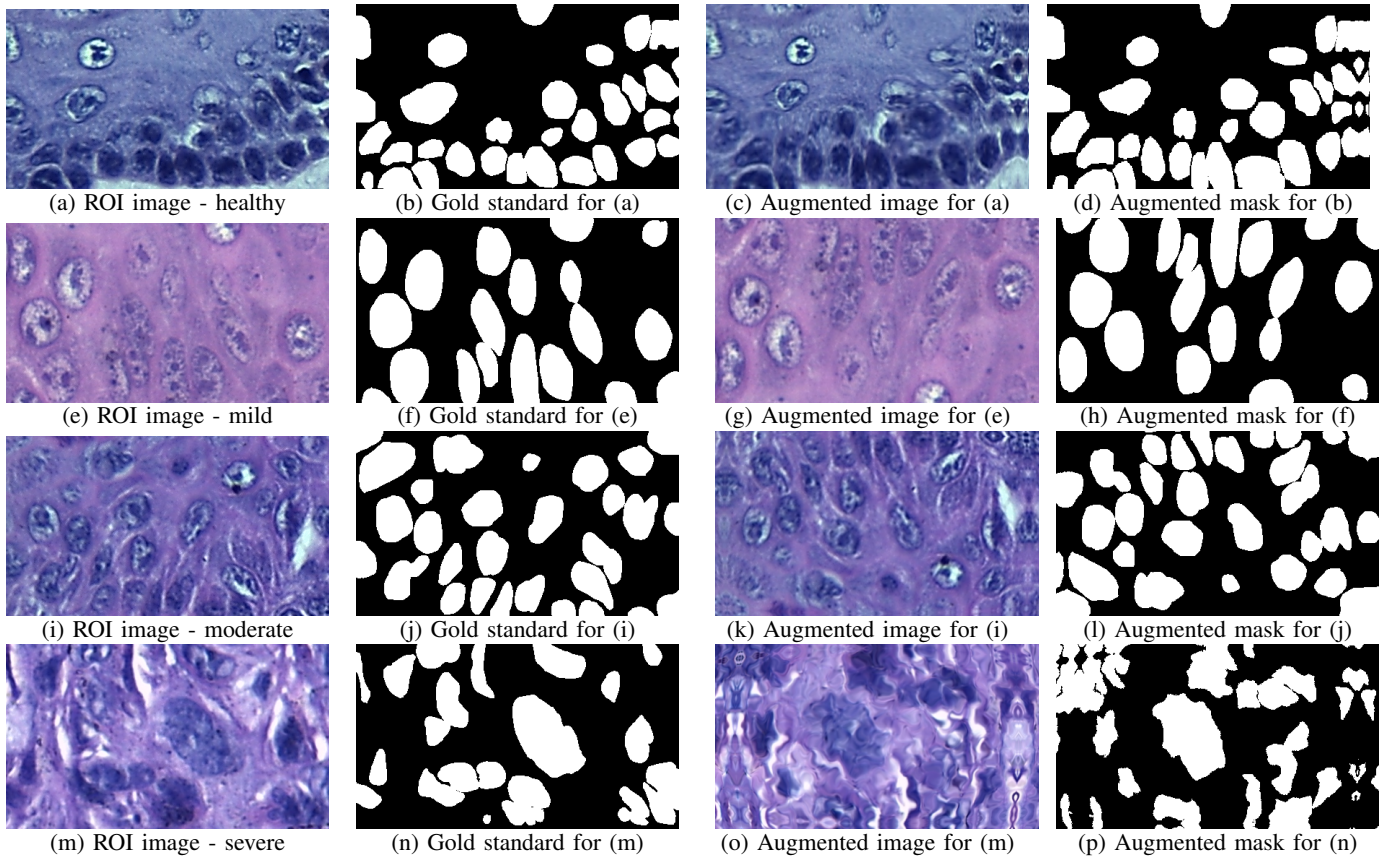


Fig. 3: ROI images from healthy and epithelial dysplasia-harboring tongue mice: 1st row shows healthy mucosa class; 2nd row shows mild dysplasia class; 3rd row shows moderate dysplasia class and 4th row shows severe dysplasia class.

$$Jaccard = \frac{TP}{TP + FP + FN}. \quad (5)$$

Specificity measures the proportion of non-nuclei areas correctly identified and is defined as:

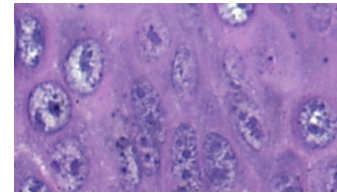
$$Specificity = \frac{TN}{TN + FP}, \quad (6)$$

where  $TP$  means true positives (the amount of correctly detected pixels),  $TN$  means true negatives (the amount of correctly undetected pixels),  $FP$  means false positives (number of incorrectly detected pixels) and  $FN$  means false negatives (the number of incorrectly undetected pixels).

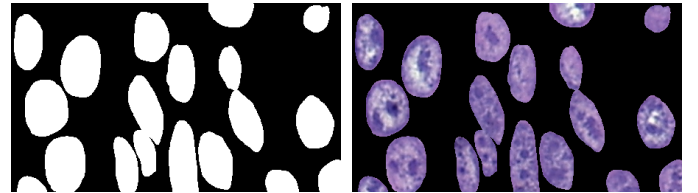
#### IV. RESULTS

In this section qualitative and quantitative results are shown. Fig. 4 shows one selected image from the dataset to serve as a reference for the qualitative analysis. The original image is shown in Fig. 4a, its corresponding nuclei mask manually delimited by the specialist is shown in Fig. 4b, and the segmentation result using the masks manually delimited by the specialist – referenced as gold-standard – is shown in Fig. 4c. Fig. 5 reveals the results obtained by the automated segmentation method for each configuration on RGB color space and the Grayscale image. Fig. 6 depicts the results for HSV color

space transformation and its respective individual channels. Fig. 7 shows the results obtained by CIE L\*a\*b\* color space transformation and its respective individual channels.



(a) Original input ROI image.



(b) Manually segmented nuclei. (c) Nuclei segmentation result.

Fig. 4: Gold standard segmentation.

Regarding quantitative analysis, the Table I summarizes the average results for each quantitative measure in the test subset. The same 24 ROI images – test subset transformed thirteen

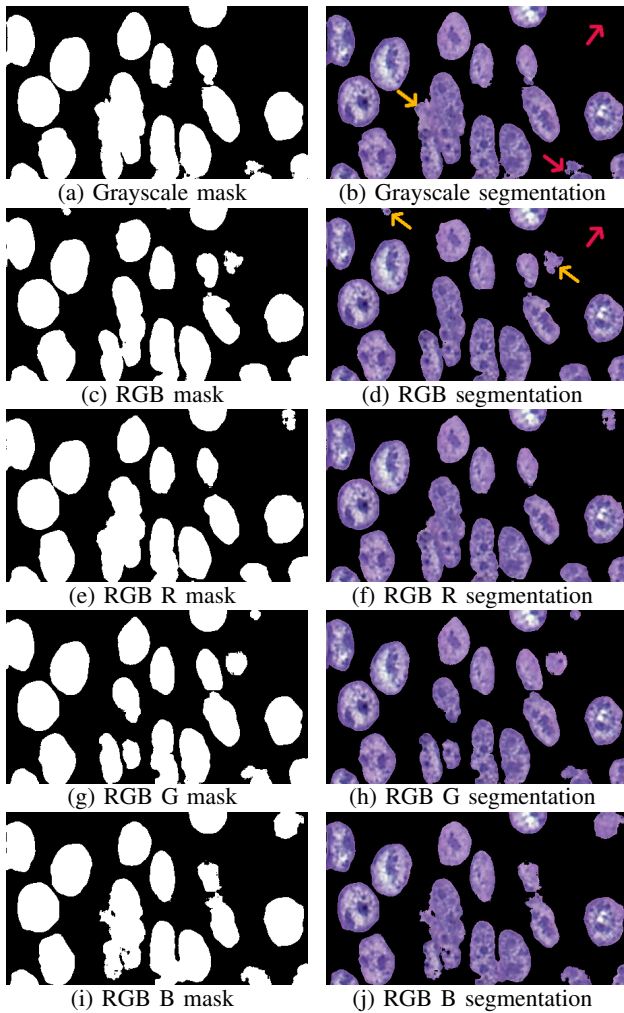


Fig. 5: Qualitative analysis for RGB color space segmentation: first column (a, c, e, g, i) shows nuclei masks obtained; second column (b, d, f, h, j) shows the final segmentation results – red arrows indicate some false negative areas; yellow arrows indicate some false positive areas.

times for each corresponding color space configuration – were employed to evaluate all configurations.

As we can see in Figs. 5, 6 and 7 and Table I, the original RGB and the CIE  $L^*a^*b^*$  color spaces present the best results, outperforming all the configurations. It is important to note that despite visually presenting good results and contours close to the gold-standard, as indicated in the images by the red and yellow arrows, all configurations present some false negative and false positive regions.

The worst results were presented by the CIE  $a^*$  and HSV H color space configurations. It is worth noting that these results indicate that the chrominance information alone, or the color values, does not play the most significant role for nuclei segmentation purpose on oral epithelial dysplasia-harboring tongue mice image datasets. The color space channels that indicate lightness information (Grayscale, CIE  $L^*$ , HSV S/V) performed very close to the best configurations, revealing that

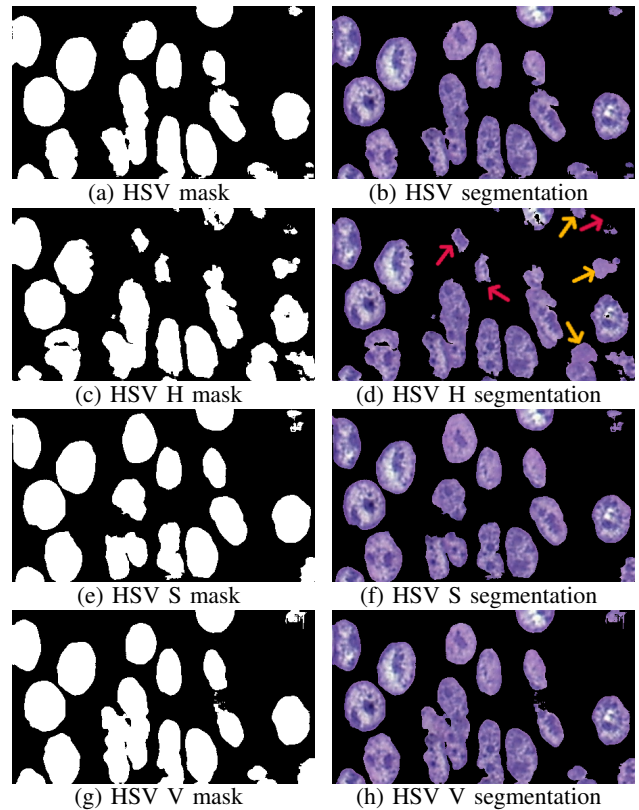


Fig. 6: Qualitative analysis for HSV color space segmentation: first column (a, c, e, g) shows nuclei masks obtained; second column (b, d, f, h) shows the final segmentation results – red arrows indicate some false negative areas; yellow arrows indicate some false positive areas.

lightness information plays a pivotal role for nuclei segmentation purposes (note also that Grayscale and  $L^*$  channel are computed from a very similar equation from the RGB color space).

## V. CONCLUSIONS

In this paper we evaluated the impact of different color space transformations applied to H&E-stained histological images for nuclei segmentation purposes. A fully convolutional neural networks model for automated nuclei segmentation was trained and run using thirteen different color space configurations of tongue mice-derived epithelial dysplasia image datasets.

Experimental results revealed that the chrominance information does not play the most significant role for nuclei segmentation purposes in H&E-stained histological images. Furthermore, the results indicates that most significant role for nuclei segmentation purposes may be played by the lightness information contained in the color spaces.

The dataset employed in this study has a reduced number of images and, in future works, the number of images will be expanded and images of human oral tissues will also be employed.

Color Space	Accuracy	Precision	F1 / Dice	Jaccard	Sensitivity / Recall	Specificity
Grayscale	0.862	0.769	0.795	0.667	0.842	0.877
RGB	0.879	0.793	<b>0.820</b>	<b>0.699</b>	<b>0.860</b>	0.891
R channel from RGB	0.874	0.788	0.811	0.689	0.848	0.893
G channel from RGB	0.860	0.794	0.780	0.647	0.782	0.908
B channel from RGB	0.862	0.775	0.794	0.666	0.832	0.885
HSV	0.881	0.807	0.819	0.698	0.844	0.904
H channel from HSV	0.784	0.691	0.659	0.503	0.648	0.858
S channel from HSV	0.842	0.752	0.765	0.628	0.801	0.871
V channel from HSV	0.863	0.783	0.792	0.663	0.817	0.895
Lab	<b>0.882</b>	<b>0.853</b>	0.814	0.689	0.788	<b>0.937</b>
L channel from Lab	0.865	0.775	0.797	0.670	0.835	0.886
a channel from Lab	0.758	0.637	0.648	0.490	0.679	0.803
b channel from Lab	0.819	0.711	0.738	0.590	0.790	0.841

TABLE I: The average quantitative results by each color space transformation applied on the test set (24 images).

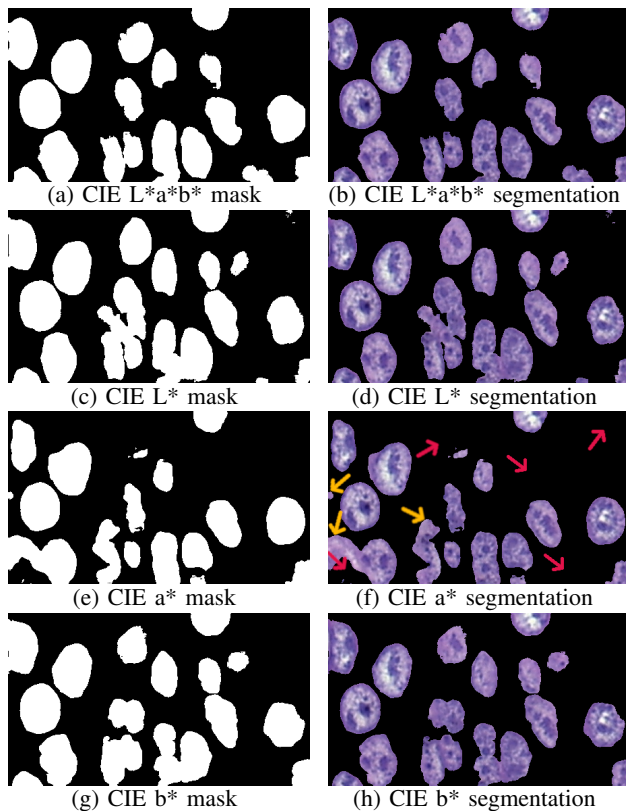


Fig. 7: Qualitative analysis for CIE L\*a\*b\* color space segmentation: first column (a, c, e, g) shows nuclei masks obtained; second column (b, d, f, h) shows the final segmentation results – red arrows indicate some false negative areas; yellow arrows indicate some false positive areas.

## REFERENCES

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. Siegel, L. Torre, and A. Jemal, "Global cancer statistics 2018: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries: Global cancer statistics 2018," *CA: A Cancer Journal for Clinicians*, vol. 68, 09 2018. 1
- [2] INCA. (2020) Instituto nacional de câncer josé alencar gomes da silva (inca). [Online]. Available: <https://www.inca.gov.br/tipos-de-cancer/cancer-de-boca> 1
- [3] J. Baik, Q. Ye, L. Zhang, C. Poh, M. Rosin, C. Macaulay, and M. Guillaud, "Automated classification of oral premalignant lesions using image cytometry and random forests-based algorithms," *Cellular Oncology*, vol. 37, p. 193 – 202, 2014. [Online]. Available: <https://doi.org/10.1007/s13402-014-0172-x> 1
- [4] D. K. Das, S. Bose, A. K. Maiti, B. Mitra, G. Mukherjee, and P. K. Dutta, "Automatic identification of clinically relevant regions from oral tissue histological images for oral squamous cell carcinoma diagnosis," *Tissue and Cell*, vol. 53, pp. 111 – 119, 2018. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S0040816618301137> 1
- [5] H. Irshad, A. Veillard, L. Roux, and D. Racoceanu, "Methods for nuclei detection, segmentation and classification in digital histopathology: A review current status and future potential," *IEEE reviews in biomedical engineering*, vol. 7, pp. 97–114, 05 2014. 1
- [6] F. Xing, Y. Xie, H. Su, F. Liu, and L. Yang, "Deep learning in microscopy image analysis: A survey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 29, no. 10, pp. 4550–4568, 2018. 1
- [7] D. F. D. dos Santos, T. A. A. Tosta, A. B. Silva, P. R. de Faria, B. A. N. Travençolo, and M. Z. do Nascimento, "Automated nuclei segmentation on dysplastic oral tissues using CNN," in *Proceedings of the 2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*, July 2020, pp. 45–50. 1, 2, 3
- [8] A. R. Robertson, "The CIE 1976 color-difference formulae," in *Color Research & Application*, vol. 2, 1977, pp. 7–11. 2
- [9] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2006. 2
- [10] M. M. Fernández-Carrobles, G. Bueno, O. Déniz, J. Salido, M. García-Rojo, and L. González-López, "Influence of texture and colour in breast tma classification," *PloS one*, vol. 10, 2015. 2
- [11] H. Lumerman, P. Freedman, and S. Kerpel, "Oral epithelial dysplasia and the development of invasive squamous cell carcinoma," *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology, and Endodontology*, vol. 79, no. 3, pp. 321 – 329, 1995. [Online]. Available: <http://www.sciencedirect.com/science/article/pii/S1079210405802264> 2
- [12] T. Falk, D. Mai, R. Bensch, Özgün Çiçek, A. Abdulkadir, Y. Marrakchi, A. Böhm, J. Deubner, Z. Jäckel, K. Seiwald, A. Dovzhenko, O. Tietz, C. D. Bosco, S. Walsh, D. Saltukoglu, T. L. Tay, M. Prinz, K. Palme, M. Simons, I. Diester, T. Brox, and O. Ronneberger, "U-net – deep learning for cell counting, detection, and morphometry," *Nat. Methods*, vol. 16, pp. 67–70, 2019. 3
- [13] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, p. 60, Jul 2019. [Online]. Available: <https://doi.org/10.1186/s40537-019-0197-0> 2
- [14] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 9, no. 1, pp. 62–66, Jan 1979. 2
- [15] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer, "Automatic differentiation in PyTorch," in *NIPS Autodiff Workshop*, 2017. 3