Intelligent Health Systems – From Technology to Data and Knowledge
E. Andrikopoulou et al. (Eds.)

doi:10.3233/SHTI250272

Understanding Stain Separation Improves Cross-Scanner Adenocarcinoma Segmentation with Joint Multi-Task Learning

Ho Heon KIM[a,1], Won Chan JEONG[a,1], Youngjin PARK[a,2], and Young Sin KO[a,2]
[a] *AI Research Center, Seegene Medical Foundation, Seoul, South Korea*
ORCiD ID:
Ho Heon Kim https://orcid.org/0000-0001-7260-7504
Won Chan Jeong https://orcid.org/0009-0008-1931-5957
Youngjin Park https://orcid.org/0000-0002-0754-8103
Young Sin Ko https://orcid.org/0000-0003-1319-4847

Abstract. Digital pathology has made significant advances in tumor diagnosis and segmentation; however, image variability resulting from tissue preparation and digitization - referred to as domain shift - remains a significant challenge. Variations caused by heterogeneous scanners introduce color inconsistencies that negatively affect the performance of segmentation algorithms. To address this issue, we have developed a joint multitask U-net architecture trained for both segmentation and stain separation. This model isolates the stain matrix and stain density, allowing it to handle color variations and improve generalization across different scanners. On 180 stain images from three different scanners, our model achieved a Dice score of 0.898 and an Intersection Over Union (IoU) score of 0.816, outperforming conventional supervised learning methods by +1.5% and +2.5%, respectively. On external datasets containing images from six different scanners, our model averaged a Dice score and IoU of 0.792. By leveraging our novel approach to stain separation, we improved segmentation accuracy and generalization across diverse histopathological samples. These advances may pave the way for more reliable and consistent diagnostic tools for breast adenocarcinoma.

Keywords. Image Segmentation, Stain separation, Multi-task learning

1. Introduction

The field of digital pathology has made significant progress, particularly in tumor diagnosis and segmentation [1]. However, the effectiveness of current algorithms is often limited by the variability in digital pathology images caused by differences in organs, tissue preparation, and image acquisition - a challenge known as *domain shift* [2]. Addressing *domain shift* is critical to ensure consistent performance of segmentation algorithms across different domains.

[1] These authors contributed equally to this work.
[2]Corresponding authors: Youngjin PARK and Young Sin KO are corresponding authors. Contact: E-mails: youngjpark@mf.seegene.com and noteasy@mf.seegene.com.

Adenocarcinoma, a common form of cancer that arises from epithelial cells, occurs in various organs, including the breast. Despite consistent staining techniques such as hematoxylin and eosin (H&E), variability still occurs due to differences in scanner types, leading to domain shift which can significantly impact the performance of machine learning models for adenocarcinoma. Although structural information of tissue or cells (e.g., contrasts) may be preserved when scanned with different devices, scanner-induced variability can adversely affect AI model performance. Understanding how stains interact with tissues - particularly in terms of color contrast and intensity - is critical to improving image analysis. Despite the importance of this knowledge, the integration of stain separation into artificial intelligence models remains underexplored.

Stain separation, which isolates structural information from pure stain style, has the potential to improve segmentation results by making algorithms more robust to scanner induced variability. The use of multi-task learning frameworks, in which a model learns multiple goals simultaneously, offers a promising approach to improving segmentation quality. The objective of this study is to investigate whether multi-task learning using stain separation can enhance segmentation performance for breast adenocarcinoma in heterogeneous scanner H&E images.

2. Methods

2.1. Multi-task learning based on stain separation

Motivated by unsupervised learning based on stain separation, our approach for joint multi-task learning leverages stain separation to isolate the stain matrix W (stain color appearance) and stain map density H [3]. Given Beer-Lambert transformed image $I \in \mathbb{R}^{n\times m}$, where m is the number of channels, and n is the number of pixels, the image, I, can be decomposed into HW. In this decomposition, the stain matrix $W \in \mathbb{R}^{r\times m}$ represents the basis colors for each stain with r being the number of stains [4]. The stain density $H \in \mathbb{R}^{n\times r}$ represents the concentration of stains at each pixel.

By learning to separate these stain components, our model was designed to effectively capture histological structures from stain density and manage variations from the stain matrix, despite differences in scanner color bases. We implemented a multi-decoder AutoEncoder within a multi-task learning framework, where each decoder serves a specific function: the stain matrix decoder (f_m) predicts the stain matrix ($\widehat{W}$) to understand various colour styles at a pixel-wise level ($\widehat{W} \in \mathbb{R}^{n\times r\times m}$), addressing inter-stain variance among pixels; the stain density decoder estimates the predicted stain density ($\widehat{H} \in \mathbb{R}^{n\times r}$). A classification header (f_c) is then attached to the model, leveraging two feature maps from both decoders for segmentation (Figure 1). We hypothesized that if the stain matrix header (f_h) and stain density header (f_d) effectively learn histological information, their corresponding convolutional neural network (CNN) kernels would be well-trained for segmentation in domain shifted dataset. To utilize these stain related features, we concatenated the predicted stain density and stain matrix ($\widehat{H}^T$, $\widehat{W}^T$) along the number of pixel axis and fed them into the classification header, yielding pixel-wise logits ($\hat{y}$) for segmentation:

$$\hat{y} = f_c(\widehat{H}^T \oplus \widehat{W}^T) \tag{1}$$

For joint multi-task learning, we formulated the objective function ($\mathcal{L}_{total}$) as a weighted average of two loss functions: reconstruction error ($\mathcal{L}_{recon}$), calculated from elementwise multiplication of $\widehat{H}$ and $\widehat{W}$ along the number of pixel axis, and the pixel-wise classification error (L_{seg}) for segmentation:

$$\mathcal{L}_{total} = \alpha\mathcal{L}_{recon} + \mathcal{L}_{seg} \tag{2}$$

Here, α denotes the coefficient for the reconstruction error, balancing the contributions of the two loss functions. Code is available at: https://github.com/4pygmalion/cosas

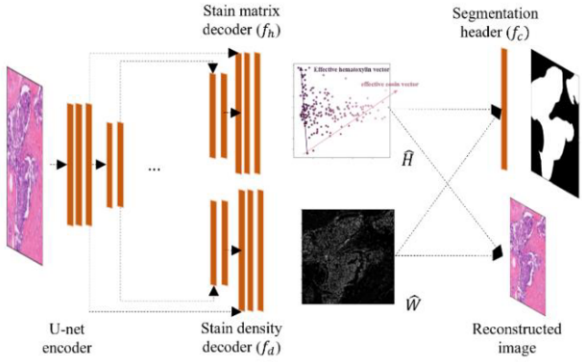

Figure 1. Multi-decoder Unet architecture for joint multi-task learning

2.2. Mixture of stain augmentation

To train a domain-generalized deep learning model, we created the realistic histopathologic image by mixing two stain augmentation methods: 1) RandStainNA [5] and 2) our proposed stain separation-based stain augmentation [6]. We mixed the augmentation probabilities by giving 0.25 to RandStainNA, and Stain separation based stain augmentation respectively. The stain separation-based augmentation technique uses SPCN (Structure-Preserving Color Normalization) to extract stain vectors and stain densities, modifying the stain vector's color basis through random distribution. This approach produces realistic H&E stained images while preserving essential histological information, as SPCN only alters the stain matrix in terms of *H*.

2.3. Training and test configuration

We used 180 patch images from 3 types of scanners including KF-BIO Pro400, TEKSQRAY SQS-600P, and 3DHISTECH 1000 Pannoramic as interval validation set from COSAS(Cross-Organ and Cross-Scanner Adenocarcinoma Segmentation). In addition, we evaluated 90 images from 6 scanners as external validation dataset [7].

We employed a multi-task U-Net architecture built on a pretrained EfficientNet-B7 backbone derived from ImageNet1K. The model training involved four-fold stratified cross-validation with an interval validation dataset from three different scanners. In addition to stain augmentation, we applied random vertical and horizontal flips to enhance generalization. Furthermore, test-time augmentation was conducted by

randomly rotating input images in 90° increments, resulting in four possible orientations (90°, 180°, 270°, and 360°).

3. Results

3.1. Model performance

The proposed model demonstrated superior performance compared to the conventional supervised learning approach, as evidenced by a Dice score of 0.887 and an IoU score of 0.805 in a 4-fold stratified cross-validation setting. These results represent a significant improvement of +1.5%, +2.5% over the baseline method (Table 1). Furthermore, joint multi-task learning has been demonstrated to enhance segmentation performance by +1.4%, +2.3%, respectively, over an identical architecture without the reconstruction objective function. In the external validation set, the model achieved a summed score of 0.7924, which is the average of the Dice and IoU scores.

Table 1. 4 Fold-stratified segmentation performance in interval validation dataset

Architecture	Backbone	Training strategy	Average Dice	Average IoU
U-net	EfficientNet-b7	Supervised	0.883	0.791
U-net	EfficientNet-b7	Multi-task,$\alpha = 0.0$	0.884	0.793
U-net	EfficientNet-b7	Multi-task,$\alpha = 0.3$	0.894	0.816

3.2. Performance sensitivity

In order to ascertain the optimal coefficient for joint multi-task learning, a sensitivity analysis was conducted. This involved a grid search, in which the value of α was modified from 0 to 1 with a step size of 0.1. The weight coefficients of the loss function were found to be optimal when α was set at 0.3.

3.3. Case review

In the internal test dataset, the U-net model with EfficientNet-b7, utilizing conventional supervised learning, yielded a Dice coefficient of 0.7575 and an Intersection over Union (IoU) of 0.609 for the adenocarcinoma patch image (Figure 2). This performance was affected by the positive prediction of lymphatic nodes, which contributed to the lower performance. In contrast, the multi-task architecture without reconstruction loss performed significantly better (Dice = 0.81, IoU = 0.68). Furthermore, when optimized, the algorithm exhibited an even higher level of accuracy (Dice = 0.85, IoU = 0.74).

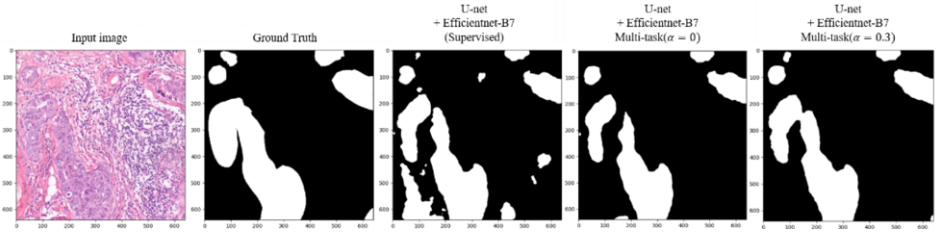

Figure 2. Example of predictions for each method: supervised, multi-task learning with α = 0 and α = 0.3

4. Discussion and Conclusions

We have developed a joint multi-task architecture which enables the learning of stain matrix and stain density through an unsupervised approach. This enables the model to effectively separate histological structures from colour variations. Conventional methodologies frequently encounter challenges when confronted with the inherent variability of scanners and stains. However, by simply integrating stain separation into our approach, we are able to achieve more consistent interpretations of histopathology images. A multihead autoencoder was devised which is able to separately predict both the stain matrix and stain density. This allows the model to distinguish between true histologic features and scanner-induced artefacts. This enhances the model's capacity to generalize across domain shifts by facilitating the acquisition of structural information. The approach was validated using three scanners for the interval validation dataset, in addition to six scanners as external dataset. This limitation arises primarily from the inherent difficulty of obtaining histopathology images scanned on multiple devices, as such datasets are rare and often restricted by scanner availability. Expanding the dataset to include larger and more diverse samples could further enhance the generalizability of the findings. In summary, we provide multi-task learning framework in computational pathology which jointly learns segmentation and stain separation task to improves conventional supervised learning on H&E stained breast cancer images from heterogeneous scanner.

References

[1] Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. Lancet Oncol. 2019;20:e253–e261. doi:10.1016/S1470-2045(19)30154-8.
[2] Stacke K, Eilertsen G, Unger J, Lundström C. Measuring domain shift for deep learning in histopathology. IEEE J Biomed Health Inform. 2021;25:325–36. doi:10.1109/JBHI.2020.3032060.
[3] Moyes A, Zhang K, Ji M, Zhou H, Crookes D. Unsupervised deep learning for stain separation and artifact detection in histopathology images. In: Medical Image Understanding and Analysis – MIUA 2020. 2020. p. 221–34. doi:10.1007/978-3-030-52791-4_18.
[4] Vahadane A, Peng T, Sethi A, Albarqouni S, Wang L, Baust M, et al. Structure-preserving color normalization and sparse stain separation for histological images. IEEE Trans Med Imaging. 2016;35:1962–71. doi:10.1109/TMI.2016.2529665.
[5] Shen Y, Luo Y, Shen D, Ke J. RandStainNA: learning stain-agnostic features from histology slides by bridging stain augmentation and normalization. In: Wang L, Dou Q, Fletcher PT, Speidel S, Li S, editors. Med Image Comput Comput Assist Interv – MICCAI 2022. Cham: Springer Nature Switzerland; 2022. p. 212–21. doi:10.1007/978-3-031-16434-7_21.
[6] Kim HH, et al. Stain augmentation via stain separation enhances classification and segmentation in H&E breast cancer histopathology: development and validation study. JMIR Cancer [Preprint]. Available from: https://preprints.jmir.org/preprint/68446 [accessed 2024 Dec 23].
[7] Liu R, Xiao H, Wang Y, Zhang M, Meng B, Long X, Liu J. Exploring domain generalization in semantic segmentation for digital histopathology: a comparative evaluation of deep learning models. In: Proceedings of the 2024 9th International Conference on Biomedical Signal and Image Processing. Suzhou, China: ACM; 2024. p. 110–6. doi:10.1145/3691521.3691537..